Ten Lectures on Small Value Probabilities and Applications

L7: More Lower Bound Techniques: Chaining and Shifting

Wenbo V. Li

University of Delaware

http://www.math.udel.edu/~wli
wli@math.udel.edu

CBMS Lectures at UAH, June 4-8, 2012

We first establish a commonly used general lower bound estimate for the supremum of non-differentiable Gaussian process via the chain argument. Then we present a connection between small ball probabilities that can be used to estimate small ball probabilities under any norm via a relatively easier L_2-norm estimate.

The Generic Chaining Method

•M. Talagrand, The generic chaining: upper and lower bounds of stochastic processes, Springer Verlag, 2005, 222 pages.

“The fundamental question of characterizing continuity and boundedness of Gaussian processes goes back to Kolmogorov. After essential contributions by R. Dudley and X. Fernique, it was solved by the author in 1985. This advance was followed by a great improvement of our understanding of the boundedness of other fundamental classes of processes (empirical processes, infinitely divisible processes, etc.) This challenging body of work has now been considerably simplified through the notion of ”generic chaining”, a completely natural variation on the ideas of Kolmogorov. The entirely new presentation adopted here takes the reader from the first principles to the edge of current knowledge, and to the wonderful open problems that remain in this domain.”

•M. Talagrand, The generic chaining: upper and lower bounds of stochastic processes, Springer Verlag, 2005, 222 pages.

•M. Talagrand (2001), Majorizing measures without measures, Annals of Probability 29, (2001), 411417

•M. Talagrand (1996), Majorizing measures: the generic chaining, Ann. Probab. 24, 1049-1103.

•M. Talagrand (1992), , A simple proof of the majorizing measure theorem, Geometric and Functional Analysis 2 (1992), 119-125

•M. Talagrand (1987), Regularity of Gaussian processes, Acta Math. 159 (1987), 99-149.

•X. Fernique (1971, 1975), stationary setting and related formulation.

•R. Dudley (1967, 1973), using covering numbers to get Dudley integral.

The Generic Chaining Method: Upper Estimates

Consider a stochastic process $\{X_t\}$ with index set T. We want upper bound for $\sup_{t\in T}|X_t - X_{t_0}|$.

•Take $q > 1$ as a power of discretization. Let $n_0 = \max\{n : D(T) = diam(T) \leq 2q^{-n}\}$. Consider an increasing sequence of partition $\mathcal{A} = (\mathcal{A}_n)$ of T such that $D(A) \leq 2q^{-n}$ for $A \in \mathcal{A}_n$, $n \geq n_0$.

•For each $t \in T$, write $A_n(t) \in \mathcal{A}_n$ such that $t \in A_n(t)$.

•For each $A \in \mathcal{A}_n$, fix a point of T in A to represent A, and denote by T_n the collection of these points to represent the partition $\mathcal{A}_n$.

•For each $t \in T$, denote $s_n(t) \in T_n$ such that $s_n(t) \in A_n(t)$. Then $s_n(t) \in A_{n-1}(S_n(t)) = A_{n-1}(t)$, and $d(s_n(t), s_{n-1}(t)) \leq 2q^{-n+1}$.

•The fundamental relation is

$$X_t - X_{t_0} = \sum_{j=1}^{\infty} X_{s_j(t)} - X_{s_{j-1}(t)}$$

which decomposes the increments as one moves from t_0 to t along the increasing "chain" $s_j(t) \in T_j$ such that $s_j(t) = s_j(s)$ implies $s_{j-1}(t) = s_{j-1}(s)$ and $s_j(v) = v$ for any $v \in T_j$.

$$
\begin{aligned}
\sup_{t\in T}\left|X_t - X_{t_0}\right| &\leq \sup_{t\in T}\sum_{j=1}^{\infty}\left|X_{s_j(t)} - X_{s_{j-1}(t)}\right| \\
&= \sup_{t\in T}\sum_{j=1}^{\infty}\left|X_{s_j(t)} - X_{s_{j-1}(s_j(t))}\right| \\
&\leq \sum_{j=1}^{\infty}\max_{v\in T_j}\left|X_v - X_{s_{j-1}(v)}\right| \\
&\leq \sum_{j=1}^{\infty}\sum_{v\in T_j}\left|X_v - X_{s_{j-1}(v)}\right|
\end{aligned}
$$

Majorizing measure

For $\sum_{j=1}^{\infty} \sum_{v \in T_j} w_j(v) \leq 1$

$$
\begin{aligned}
& \mathbb{P}\left(\sup_{t \in T} |X_t - X_{t_0}| \geq u\right) \\
\leq \ & \mathbb{P}\left(\sum_{j=1}^{\infty} \sum_{v \in T_j} \left|X_v - X_{s_{j-1}(v)}\right| \geq u\right) \\
\leq \ & \sum_{j=1}^{\infty} \sum_{v \in T_j} \mathbb{P}\left(\left|X_v - X_{s_{j-1}(v)}\right| \geq w_j(v) u\right)
\end{aligned}
$$

•Majorizing measure: Best partition $\mathcal{A} = (\mathcal{A}_n)$ and best weight or 'measure' $w_j(v)$.

•The Dudley entropy upper bound: For any $A \in \mathcal{A}_n$, cover A by $N(A, d, q^{-n-1}) \leq N(T, d, q^{-n-1})$.

A lower bound for small ball probability

For centered Gaussian process $X_t, t \in T$,

$$
\begin{aligned}
& \mathbb{P}\left(\sup_{t\in T} |X_t - X_{t_0}| \le \varepsilon\right) \\
\ge\ & \mathbb{P}\left(\sum_{j=1}^{\infty} \max_{v\in T_j} \left|X_v - X_{s_{j-1}(v)}\right| \le \varepsilon\right) \\
\ge\ & \mathbb{P}\left(\max_{v\in T_j} \left|X_v - X_{s_{j-1}(v)}\right| \le p_j(\varepsilon)\varepsilon \ \text{ for } \ j \ge 1\right) \\
\ge\ & \prod_{j\ge 1} \prod_{v\in T_j} \mathbb{P}\left(\left|X_v - X_{s_{j-1}(v)}\right| \le p_j(\varepsilon)\varepsilon\right) \\
=\ & \prod_{j\ge 1} \prod_{v\in T_j} \mathbb{P}\left(|\xi_v| \le p_j(\varepsilon)\varepsilon / d(v, s_{j-1}(v))\right)
\end{aligned}
$$

for $\sum_j p_j(\varepsilon) \le 1$, where we used Sidak's inequality for the last inequality.

Lower bound on supremum under entropy conditions

Assume $(X_t)_{t\in T}$ is a centered Gaussian process with entropy number $N(T, d; \varepsilon)$, the minimal number of balls of radius $\varepsilon > 0$, under the Dudley metric $d(s,t) = (\mathbb{E}\,|X_s - X_t|^2)^{1/2}, \quad s, t \in T$ that are necessary to cover T. Then a commonly used general lower bound estimate on the supremum is the following formulation of Talagrand:

Thm: Assume that there is a nonnegative function ψ on $\mathbb{R}_+$ such that $N(T, d; \varepsilon) \le \psi(\varepsilon)$ and such that $c_1\psi(\varepsilon) \le \psi(\varepsilon/2) \le c_2\psi(\varepsilon)$ for some constants $1 < c_1 \le c_2 < \infty$. Then, for some $C > 0$ and every $\varepsilon > 0$ we have

$$\log \mathbb{P}\left(\sup_{s,t\in T} |X_s - X_t| \le \varepsilon\right) \ge -C\psi(\varepsilon).$$

In particular, $\log \mathbb{P}\left(\sup_{t\in T} |X_t| \le \varepsilon\right) \ge -C'\psi(\varepsilon)$.

•See Ledoux (1996) for a detailed proof.

Application: Fractional Brownian Motion

Let $B^H(t)$, $t \geq 0$ be a standard real valued fractional Brownian motion with index $H \in (0,1)$. That is, $B^H(t)$ is a zero-mean Gaussian process with stationary increments and covariance function

$$\mathbb{E}\left[B^H(t)B^H(s)^\top\right] = \frac{1}{2}\left\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right\}$$

Then

$$\log \mathbb{P}(\sup_{0\leq t\leq 1} |B^H(t)| \leq \varepsilon) \geq -c\varepsilon^{1/H}$$

Remark

•Although the lower bound is relatively easy to use, it does *not* always provide sharp lower estimates even when $N(T, d; \varepsilon)$ can be estimated sharply. The simplest example is $X_t = \xi t$ for $t \in T = [0, 1]$, where ξ denotes a standard normal random variable. In this case,

$$\mathbb{P}\left(\sup_{t \in T} |X_t| < \varepsilon\right) = \mathbb{P}\left(|\xi| < \varepsilon\right) \sim (2/\pi)^{1/2} \cdot \varepsilon,$$

but the Theorem produces an exponential lower bound $\exp(-c/\varepsilon)$ for the above probability. More interesting examples are the integrated fractional Brownian motion and the fractional integrated Brownian motion.

Open: Find better probabilistic lower bound for smooth Gaussian processes.

General Approach

Let $(X_t)_{t\in T}$ be a centered Gaussian process. If there are a countable set T_c and a Gaussian process Y on T_c such that

$$\left\{\sup_{t\in T}|X_t|\le x\right\}\supset\left\{\max_{t\in T_c}|Y_t|\le x\right\},$$

then we have by Sidak's inequality

$$\mathbb{P}\left(\sup_{t\in T}|X_t|\le x\right)\ge\prod_{t\in T_c}\mathbb{P}(|Y_t|\le x).$$

Since Y_t is a normal random variable for each $t\in T_c$, the right hand side above can be easily estimated. So, the key step of estimating the lower bound of $\mathbb{P}\left(\sup_{t\in T}|X_t|\le x\right)$ is to find countable set T_c and Gaussian process Y.

Review: Inequality in Chen and Li (2003)

Let X and Y be any two centered independent Gaussian random vectors in a separable Banach space B with norm $\|\cdot\|$. We use $|\cdot|_{\mu(X)}$ to denote the inner product norm induced on the associated reproducing Hilbert space H_μ by $\mu = \mathcal{L}(X)$. Then for any $\lambda > 0$ and $\epsilon > 0$,

$$\mathbb{P}(\|Y\| \leq \epsilon) \geq \mathbb{P}(\|X\| \leq \lambda\epsilon) \cdot \mathbb{E}\,\exp\{-2^{-1}\lambda^2 |Y|^2_{\mu(X)}\}.$$

In particular, for any $\lambda > 0$, $\varepsilon > 0$ and $\delta > 0$,

$$\mathbb{P}\left(\|Y\| \leq \varepsilon\right) \cdot \exp\{-\lambda^2\delta^2/2\} \geq \mathbb{P}\left(\|X\| \leq \lambda\varepsilon\right) \mathbb{P}\left(|Y|_{\mu(X)} \leq \delta\right).$$

•This inequality provides a powerful way to estimate lower bound for the class of processes $Y_t = G(C_t), t \in T$ where $G(\cdot)$ is a Gaussian process and C_t is an independent 'clock'. More details in lecture 7.

Review: m-th Integrated BM

Let $X_0(t) = W(t)$ and

$$X_m(t) = \int_0^t X_{m-1}(s)ds, \quad t \geq 0, \quad m \geq 1,$$

which is the m'th integrated Brownian motion or the m-fold primitive.

•We have for $m \geq 0$

$$\lim_{\lambda \to \infty} \lambda^{-1/(2m+2)} \log \mathbb{E} \exp \left\{ -\lambda \int_0^1 X_m^2(t)dt \right\} = -C_m.$$

where $C_m = 2^{-(2m+1)/(2m+2)} \left(\sin \frac{\pi}{2m+2} \right)^{-1}$.

•By the Tauberian theorem, we have

$$\begin{aligned} & \log \mathbb{P} \left(\int_0^1 X_m^2(t)dt \leq \varepsilon^2 \right) \\ \sim \quad & 2^{-1}(2m+1) \left((2m+2) \sin \frac{\pi}{2m+2} \right)^{-(2m+2)/(2m+1)} \varepsilon^{-2/(2m+1)}. \end{aligned}$$

Application: m-th Integrated BM under sup-norm

Thm: We have for $m \geq 1$,

$$\lim_{\varepsilon \to 0} \varepsilon^{2/(2m+1)} \log \mathbb{P}\left(\sup_{0 \leq t \leq 1} |X_m(t)| \leq \varepsilon\right) = -\kappa_m$$

with $\frac{2m+1}{2}\left((2m+2)\sin\frac{\pi}{2m+2}\right)^{-\frac{2m+2}{2m+1}} \leq \kappa_m \leq \frac{2m+1}{2}\left(\frac{\pi}{2}\right)^{\frac{2}{2m+1}}\left(2m\sin\frac{\pi}{2m}\right)^{-\frac{2m}{2m+1}}$ and $\lim_{m\to\infty} m^{-1}\kappa_m = \pi^{-1}$. In addition, the following upper bound holds for κ_m based on κ_{m-1}.

$$(4\kappa_m/(2m+1))^{2m+1} \leq (4\kappa_{m-1}/(2m-1))^{2m-1}.$$

- The particular case of $m = 1$, or the so called integrated Brownian motion was studied in Khoshnevisan and Shi (1998) by using local time techniques. Here our upper bound $\kappa_1 \leq (2\pi)^{2/3} \cdot (3/8)$ is explicit. In general, the exact values of κ_m is unknown for $m \geq 1$. The value $\kappa_0 = \pi^2/8$ is well known.

$\bullet$The existence of the limiting constant κ_m will be give in lecture 9.
$\bullet$The probability upper bound follows easily from

$$\mathbb{P}\left(\sup_{0\le t\le 1}|X_m(t)|\le\varepsilon\right)\le\mathbb{P}\left(\int_0^1|X_m(t)|^2dt\le\varepsilon^2\right)$$

and the L_2-estimate.
$\bullet$The probability lower bound follows from another L_2-estimate and the well known estimate

$$\log\mathbb{P}\left(\sup_{0\le t\le 1}|W(t)|\le\varepsilon\right)\sim-(\pi^2/8)\varepsilon^{-2}.$$

An Application

Take $Y = X_m(t)$ and $X = W(t)$. Then for any norm $\|\cdot\|$ on $C[0,1]$ and any $\lambda = \lambda_\varepsilon > 0$

$$\mathbb{P}\left(\|X_m\| \le \varepsilon\right) \ge \mathbb{P}\left(\|W(t)\| \le \lambda\varepsilon\right) \cdot \mathbb{E} \exp\left\{ -\frac{\lambda^2}{2} \int_0^1 X_{m-1}^2(s)ds \right\}$$

since $|f|_\mu^2 = \int_0^1 (f'(s))^2 ds$ for Wiener measure $\mu = \mathcal{L}(W)$. Taking $\|\cdot\|$ to be the sup-norm on $C[0,1]$ and $\lambda = \lambda_\varepsilon = \alpha\varepsilon^{-2m/(2m+1)}$ with $\alpha > 0$, then

$$\begin{aligned}
-\kappa_m &= \lim_{\varepsilon\to 0} \varepsilon^{2/(2m+1)} \log \mathbb{P}\left(\sup_{0\le t\le 1} |X_m(t)| \le \varepsilon\right) \\
&\ge \lim_{\varepsilon\to 0} \varepsilon^{2/(2m+1)} \log \mathbb{P}\left(\sup_{0\le t\le 1} |W(t)| \le \alpha\varepsilon^{1/(2m+1)}\right) \\
&\quad + \lim_{\varepsilon\to 0} \varepsilon^{2/(2m+1)} \log \mathbb{E} \exp\left\{ -\frac{\alpha^2}{2}\varepsilon^{-4m/(2m+1)} \int_0^1 X_{m-1}^2(s)ds \right\} \\
&= -(\pi^2/8)\alpha^{-2} - \frac{1}{2}\alpha^{1/m}\left(\sin\frac{\pi}{2m}\right)^{-1}.
\end{aligned}$$

Picking the best $\alpha > 0$, we obtain the desired result.

Remarks

•It is of interest to note that both bounds rely on easier L_2-estimates and the constant bounds for κ_m are the sharpest known.

•We can also obtain similar lower estimates for iterated symmetric stable processes.

•The corresponding upper estimates are given in lecture 8.

Local Time

Let $L = \{L_t^x\,;\,(x,t) \in R^1 \times R_+\}$ denote the local time of a "nice" stochastic process $X(t), \geq 0$, i.e. we need at least the existence and joint continuity of L. Then we have the occupation density formula,

$$\int_0^t g(X_s)\,ds = \int g(x) L_t^x\,dx$$

for all continuous functions g of compact support. Therefore, for any bounded Borel set A,

$$\int_A L_t^y dy = \int_0^t 1_A(X_s) ds.$$

In particular, let $A = [x, x+\varepsilon]$. Using the continuity of L_t^x we have

$$L_t^x = \lim_{\varepsilon \to \infty} \frac{\text{measure}\{0 \leq s \leq t : x \leq W_s \leq x+\varepsilon\}}{\varepsilon} = \int_0^t \delta_x(X_s) ds.$$

This gives an intrinsic definition of L_t^x as the derivative of an occupation measure.

•One needs to show the existence and joint continuity by approximation.

Local Times via Fourier transform

For a fixed sample function and fixed time $t > 0$, the Fourier transform on space variable $x \in \mathbb{R}^d$ is the function of $\lambda \in \mathbb{R}^d$,

$$\int_{\mathbb{R}^d} e^{i\lambda\cdot x} L(t,x)dx = \int_0^t e^{i\lambda\cdot X(s)} ds.$$

Thus the local time $L(t,x)$ can be expressed as the inverse Fourier transform:

$$L(t,x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\lambda\cdot x} \int_0^t e^{i\lambda\cdot X(s)}\, ds d\lambda.$$

The m-th power of $L(t,x)$ is

$$L(t,x)^m = \frac{1}{(2\pi)^{md}} \int_{\mathbb{R}^{md}} e^{-ix\cdot\sum_{k=1}^m \lambda_k} \int_{[0,t]^m} \exp\Big(i\sum_{k=1}^m \lambda_k\cdot X(s_k)\Big) d\mathbf{s} d\lambda$$

Take the expected value under the sign of integration: the second exponential in the above integral is replaced by the joint characteristic function of $X(s_1), \cdots, X(s_m)$.

Moments of Local Time for Gaussian Process

In the Gaussian case, we obtain

$$\mathbb{E}\, L(t,x)^m$$
$$= \frac{1}{(2\pi)^{md}} \int_{\mathbb{R}^{md}} e^{-ix\cdot\sum_{k=1}^m \lambda_k} \int_{[0,t]^m} \exp\Big(-\frac{1}{2}\mathrm{Var}\big(\sum_{k=1}^m \lambda_k \cdot X(s_k)\big)\Big) d\mathbf{s} d\lambda.$$

•Interchanging integration and applying the characteristic function inversion formula, we can get more explicit (but somewhat less useful) expression in terms of integration associated with

$$\det(\mathbb{E}\, X(s_i)X(s_j))^{-1/2}.$$

•Estimates of the moments of local time $L(t,x)$ thus depend on the rate of decrease to 0 of $\det(\mathbb{E}\, X(s_i)X(s_j))$ as $s_j - s_{j-1} \to 0$ for some/all j.

•When considering a random fields $X(\mathbf{t})$ taking values in $\mathbb{R}^d$, where $\mathbf{t} = (t_1, \ldots, t_p) \in (\mathbb{R}^+)^p$, suitable adjustment via approximation for local time are needed.

Moment Comparison for Local Times of GP

Here are some simplest forms of comparison which are based on the standard Fourier analytic approach but go far beyond, motivated mainly by similar small deviation estimates.

•For two independent nice Gaussian fields (centered) X and Y with index set A,

$$\mathbb{E}\left[L_{X+Y}(A,0)^m\right] \le \mathbb{E}\left[L_X(A,0)^m\right]$$

which follows from $\mathrm{Var}(X+Y) \ge \mathrm{Var}(X)$. This is an analogy of Anderson's inequality for local time.

•For two independent nice Gaussian fields (centered) X and Y with index set A,

$$\mathbb{E}\left[L_{X+Y}(A,0)^m\right] \ge \mathbb{E}\, e^{-\frac{1}{2}|Y|^2_{\mu(X)}} \mathbb{E}\left[L_X(A,0)^m\right],$$

which follows from Cameron-Martin formula.

General Functional Analytic Facts

Lemma: Let μ be a centered Gaussian measure in a separable Banach space B. Let $g : B \mapsto \mathbb{R}_+$ be a measurable function. Then (i) if $\{x \in B : g(x) \geq t\}$ is symmetric and convex for every $t > 0$, then for every $y \in B$

$$\int_B g(x+y)\,\mu(dx) \leq \int_B g(x)\,\mu(dx);$$

(ii) if g is symmetric ($g(-x) = g(x)$, $x \in B$), then for every y in the RKHS $\mathcal{H}_\mu$ of μ

$$\int_B g(x+y)\,\mu(dx) \geq \exp\left\{-\frac{1}{2}\|y\|_\mu^2\right\} \int_B g(x)\,\mu(dx),$$

where $\|y\|_\mu$ denotes the norm in $\mathcal{H}_\mu$.

Pf: Part *(i)* follows from Anderson's inequality

$$\begin{aligned}\int_B g(x+y)\,\mu(dx) = & \int_0^\infty \mu\{x \in B : \, g(x+y) \geq t\}\,dt \\ \leq & \int_0^\infty \mu\{x \in B : \, g(x) \geq t\}\,dt = \int_B g(x)\,\mu(dx).\end{aligned}$$

Part *(ii)* uses Cameron-Martin formula and the convexity of exponential function

Chen, Li, Rosinski and Shao (2011)

Thm: Let $B^H(t)$ be a standard d-dimensional fractional Brownian motion with index H such that $Hd < 1$. Then the limit

$$\lim_{a\to\infty} a^{-1/(Hd)} \log \mathbb{P}\{L_1^0(B^H) \geq a\} = -\theta(H,d)$$

exists and $\theta(H,d)$ satisfies the following bounds

$$\left(\pi c_H^2/H\right)^{1/(2H)} \theta_0(Hd) \leq \theta(H,d) \leq (2\pi)^{1/(2H)}\theta_0(Hd)$$

and

$$\theta_0(\kappa) = \kappa \left(\frac{(1-\kappa)^{1-\kappa}}{\Gamma(1-\kappa)}\right)^{1/\kappa}$$

Thm: Let $\alpha^H(\cdot)$ be the intersection local time of p-independent standard d-dimensional fractional Brownian motions $B_1^H(t), \cdots, B_p^H(t)$, where $Hd < p^*$. Then the limit

$$\lim_{a\to\infty} a^{-p^*/(Hdp)} \log \mathbb{P}\left\{\alpha^H\left([0,1]^p\right) \geq a\right\} = -K(H,d,p)$$

exists with

$$K(H,d,p) = c_H^{1/H} \tilde{K}(H,d,p).$$