L8: More Upper Bound Techniques

Wenbo V. Li

University of Delaware

wli@math.udel.edu

CBMS Lectures at UAH, June 4, 2012

The upper bound of small ball probabilities is much more challenging than the lower bound. This can be seen easily from the precise links with the metric entropy. The upper bound of small ball probabilities gives the lower estimate of the metric entropy and vice versa. The lower estimates for metric entropy are frequently obtained by a volume comparison, i.e. for suitable finite dimensional projections, the total volume of the covering balls is less than the volume of the set being covered. As a result, when the volumes of finite dimensional projections of K_μ do not compare well with the volumes of the same finite dimensional projection of the unit ball of E, sharp lower estimates for metric entropy (upper bounds for small ball probabilities) are much harder to obtain.

Independent and Increments

Let $S_\alpha(t)$ be the α-stable process with $S_\alpha(0) = 0$, $0 < \alpha \leq 2$. Then for any $0 = t_0 \leq t_1 < \cdots < t_n \leq 1$,

$$\mathbb{P}\left(\sup_{0 \leq t \leq 1} |S_\alpha(t)| \leq \varepsilon\right) \leq \mathbb{P}\left(\max_{1 \leq i \leq n} |S_\alpha(t_i)| \leq \varepsilon\right)$$

Independent and Increments

Let $S_\alpha(t)$ be the α-stable process with $S_\alpha(0) = 0$, $0 < \alpha \leq 2$. Then for any $0 = t_0 \leq t_1 < \cdots < t_n \leq 1$,

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\leq t\leq 1} |S_\alpha(t)| \leq \varepsilon\right) &\leq \mathbb{P}\left(\max_{1\leq i\leq n} |S_\alpha(t_i)| \leq \varepsilon\right) \\
&\leq \mathbb{P}\left(\max_{1\leq i\leq n} |S_\alpha(t_i) - S_\alpha(t_{i-1})| \leq 2\varepsilon\right) \\
&= \prod_{i=1}^{n} \mathbb{P}\left(|S_\alpha(t_i) - S_\alpha(t_{i-1})| \leq 2\varepsilon\right)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{P}\left(|S_\alpha(t_i) - S_\alpha(t_{i-1})| \leq 2\varepsilon\right) &= \mathbb{P}\left(|S_\alpha(t_i - t_{i-1})| \leq 2\varepsilon\right) \\
&= \mathbb{P}\left(|S_\alpha(1)| \leq 2\varepsilon(t_i - t_{i-1})^{-1/\alpha}\right).
\end{aligned}$$

Independent and Increments

Let $S_\alpha(t)$ be the α-stable process with $S_\alpha(0) = 0$, $0 < \alpha \le 2$. Then for any $0 = t_0 \le t_1 < \cdots < t_n \le 1$,

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\le t\le 1} |S_\alpha(t)| \le \varepsilon\right) &\le \mathbb{P}\left(\max_{1\le i\le n} |S_\alpha(t_i)| \le \varepsilon\right) \\
&\le \mathbb{P}\left(\max_{1\le i\le n} |S_\alpha(t_i) - S_\alpha(t_{i-1})| \le 2\varepsilon\right) \\
&= \prod_{i=1}^{n} \mathbb{P}\left(|S_\alpha(t_i) - S_\alpha(t_{i-1})| \le 2\varepsilon\right)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{P}\left(|S_\alpha(t_i) - S_\alpha(t_{i-1})| \le 2\varepsilon\right) &= \mathbb{P}\left(|S_\alpha(t_i - t_{i-1})| \le 2\varepsilon\right) \\
&= \mathbb{P}\left(|S_\alpha(1)| \le 2\varepsilon(t_i - t_{i-1})^{-1/\alpha}\right).
\end{aligned}$$

If we take $t_i = i/n$ and $\varepsilon n^{1/\alpha} = 1$ so that each term in the product is $\mathbb{P}\left(|S_\alpha(1)| \le 2\right)$, then $\mathbb{P}\left(\sup_{0\le t\le 2} |S_\alpha(t)| \le \varepsilon\right) \le \exp(-c\varepsilon^{-\alpha})$.

Remarks

•Here we intentionally use arbitrary partition points t_i and pick what we need later. There are problems that we do need to take uneven partition points.
•Scaling or self-similarity plays important role here.
•What about integrated or m-th integrated Stable process $I_m(t)$ defined by

$$I_m(t) = \int_0^t I_{m-1}(s)ds, \quad I_0(t) = S_\alpha(t), \quad , m \geq 1?$$

•For m-th integrated BM, we can use L_2 bound via KL expansion.

Integrated Stable Process

The first step is to “differentiate” the process by using higher order differences. We have

$$
\begin{aligned}
& \mathbb{P}\left(\sup_{0\le t\le 1} \left| \int_0^t S_\alpha(u)du \right| \le \varepsilon\right) \\
\le \ & \mathbb{P}\left(\max_{1\le i\le n} \left| \int_0^{t_i} S_\alpha(u)du \right| \le \varepsilon\right)
\end{aligned}
$$

Integrated Stable Process

The first step is to "differentiate" the process by using higher order differences. We have

$$\begin{aligned}
& \mathbb{P}\left(\sup_{0\le t\le 1} \left|\int_0^t S_\alpha(u)du\right| \le \varepsilon\right) \\
\le\ & \mathbb{P}\left(\max_{1\le i\le n} \left|\int_0^{t_i} S_\alpha(u)du\right| \le \varepsilon\right) \\
\le\ & \mathbb{P}\left(\max_{1\le i\le n} \left|\int_0^{t_{i+1}} S_\alpha(u)du - 2\int_0^{t_i} S_\alpha(u)du + \int_0^{t_{i-1}} S_\alpha(u)du\right| \le 4\varepsilon\right) \\
=\ & \mathbb{P}\left(\max_{1\le i\le n} \left|\int_{t_i}^{t_{i+1}} S_\alpha(u)du - \int_{t_{i-1}}^{t_i} S_\alpha(u)du\right| \le 4\varepsilon\right)
\end{aligned}$$

Integrated Stable Process

The first step is to "differentiate" the process by using higher order differences. We have

$$
\begin{aligned}
& \mathbb{P}\left(\sup_{0\le t\le 1} \left|\int_0^t S_\alpha(u)du\right| \le \varepsilon\right) \\
\le\ & \mathbb{P}\left(\max_{1\le i\le n} \left|\int_0^{t_i} S_\alpha(u)du\right| \le \varepsilon\right) \\
\le\ & \mathbb{P}\left(\max_{1\le i\le n} \left|\int_0^{t_{i+1}} S_\alpha(u)du - 2\int_0^{t_i} S_\alpha(u)du + \int_0^{t_{i-1}} S_\alpha(u)du\right| \le 4\varepsilon\right) \\
=\ & \mathbb{P}\left(\max_{1\le i\le n} \left|\int_{t_i}^{t_{i+1}} S_\alpha(u)du - \int_{t_{i-1}}^{t_i} S_\alpha(u)du\right| \le 4\varepsilon\right)
\end{aligned}
$$

If we take $t_i = i/n$ and $\varepsilon n^{1+1/\alpha} = 1$, then the last line above equals

$$
\mathbb{P}\left(\max_{1\le i\le n} \left|\int_0^1 \left(S_\alpha(i+u) - S_\alpha(i-1+u)\right) du\right| \le 4\right)
$$

by using scaling property of stable process.

Independents

In order to create independent terms inside the maximum, we keep only the odd terms (i=2j+1) to obtain the upper bound

$$
\begin{aligned}
& \mathbb{P}\left(\max_{0\le j\le (n-1)/2} \Big|\int_0^1 \left(S_\alpha(2j+1+u) - S_\alpha(2j+u)\right) du\Big| \le 4\right) \\
= & \prod_{j=0}^{[(n-1)/2]} \mathbb{P}\left(\Big|\int_0^1 \left(S_\alpha(2j+1+u) - S_\alpha(2j+u)\right) du\Big| \le 4\right) \\
= & \; r^{[(n+1)/2]} \le \exp(-c\varepsilon^{\alpha/(1+\alpha)})
\end{aligned}
$$

where $c = \mathbb{P}\left(|\int_0^1 \left(S_\alpha(1+u) - S_\alpha(u)\right) du| \le 4\right) < 1$. Note that we do need to use the fact that the processes

$$
X_j(u) = S_\alpha(2j+1+u) - S_\alpha(2j+u), \quad 0 \le u \le 1,
$$

are independent for $j \ge 0$. This can be checked by using the joint characteristic function.

Remarks

•For general integer m, we can use $(m+1)$-th difference scheme.
•Lower bound for I_m under sup norm can be obtained by using the norm comparison inequality of Chen and Li.

Generic L_2 upper bound for sup-norm

We have for any index set T,

$$\begin{aligned}
\mathbb{P}(\sup_{t\in T}|X(t)|\leq\varepsilon) &\leq \mathbb{P}(\sup_{t\in T}|\sum_{j=1}^{m} b_j X(t)|\leq\varepsilon) \\
&\leq \mathbb{P}(\max_{1\leq i\leq n}|\sum_{j=1}^{m} b_i X(t_i)|\leq\varepsilon) \\
&\leq \mathbb{P}\left(\sum_{i=1}^{n} w_i(\sum_{j=1}^{m} b_i X(t_i))^2\leq\varepsilon^2\right)
\end{aligned}$$

for any $m, n \geq 1$,

$$\sum_{j=1}^{m}|b_j|\leq 1, \quad \sum_{i=1}^{n} w_i \leq 1$$

with $b_j \in \mathbb{R}$ and $w_i \geq 0$.

L_2 upper bound

We start with the following basic fact.

Lemma For any centered Gaussian sequence $\{\xi_i\}$ and for any $0 < x < \sum_{i\le n} \mathbb{E}\,\xi_i^2$, we have

$$\mathbb{P}\Big(\sum_{i\le n} \xi_i^2 \le x\Big) \le \exp\Big(-\frac{(\sum_{i\le n} \mathbb{E}\,\xi_i^2 - x)^2}{4\sum_{1\le i,j\le n} (\mathbb{E}\,\xi_i\xi_j)^2}\Big).$$

L_2 upper bound

We start with the following basic fact.

Lemma For any centered Gaussian sequence $\{\xi_i\}$ and for any $0 < x < \sum_{i\le n} \mathbb{E}\,\xi_i^2$, we have

$$\mathbb{P}\Big(\sum_{i\le n} \xi_i^2 \le x\Big) \le \exp\Big(-\frac{(\sum_{i\le n}\mathbb{E}\,\xi_i^2 - x)^2}{4\sum_{1\le i,j\le n}(\mathbb{E}\,\xi_i\xi_j)^2}\Big).$$

Pf: It is easy to see that there exists a sequence of independent mean zero normal random variables η_i such that

$$\sum_{i=1}^n \xi_i^2 = \sum_{i=1}^n \eta_i^2.$$

Then for any $0 < x < \sum_{i=1}^n \mathbb{E}\,\xi_i^2$

$$\begin{aligned}\mathbb{P}(\sum_{i=1}^n \xi_i^2 \le x) = \mathbb{P}(\sum_{i=1}^n \eta_i^2 \le x) &\le e^{\lambda x}\prod_{i\le n}\mathbb{E}\,e^{-\lambda\eta_i^2} \\ &= \exp\Big(\lambda x - \frac{1}{2}\sum_{i\le n}\log(1+2\lambda\mathbb{E}\,\eta_i^2)\Big).\end{aligned}$$

Let

$$\lambda = \frac{\sum_{i\le n}\mathbb{E}\,\xi_i^2 - x}{2\sum_{i\le n}(\mathbb{E}\,\eta_i^2)^2}.$$

Then the exponent

$$\begin{aligned}\lambda x - \frac{1}{2}\sum_{i\le n}\log(1+2\lambda\mathbb{E}\,\eta_i^2) &\le -\Big(\sum_{i\le n}\mathbb{E}\,\xi_i^2 - x\Big)\lambda + \lambda^2\sum_{i\le n}(\mathbb{E}\,\eta_i^2)^2 \\ &= -\frac{(\sum_{i\le n}\mathbb{E}\,\xi_i^2 - x)^2}{4\sum_{i\le n}(\mathbb{E}\,\eta_i^2)^2}.\end{aligned}$$

Let

$$\lambda = \frac{\sum_{i\le n} \mathbb{E}\,\xi_i^2 - x}{2\sum_{i\le n}(\mathbb{E}\,\eta_i^2)^2}.$$

Then the exponent

$$\begin{aligned} \lambda x - \frac{1}{2}\sum_{i\le n}\log(1+2\lambda\mathbb{E}\,\eta_i^2) &\le -(\sum_{i\le n}\mathbb{E}\,\xi_i^2 - x)\lambda + \lambda^2\sum_{i\le n}(\mathbb{E}\,\eta_i^2)^2 \\ &= -\frac{(\sum_{i\le n}\mathbb{E}\,\xi_i^2 - x)^2}{4\sum_{i\le n}(\mathbb{E}\,\eta_i^2)^2}. \end{aligned}$$

Note further that

$$\sum_{i\le n}(\mathbb{E}\,\eta_i^2)^2 = \frac{1}{2}\mathsf{Var}(\sum_{i=1}^n \eta_i^2) = \frac{1}{2}\mathsf{Var}(\sum_{i=1}^n \xi_i^2) = \sum_{i,j}\left(\mathbb{E}\,(\xi_i\xi_j)\right)^2$$

for $\mathbb{E}\,\xi_i^2\xi_j^2 = (\mathbb{E}\,\xi_i^2)(\mathbb{E}\,\xi_j^2) + 2(\mathbb{E}\,\xi_i\xi_j)^2$. The lemma follows from the above inequalities.

An L_2 type upper bound for Gaussian Process

We start with the following general result.
Let $\{X_t, t \in [0,1]\}$ be a centered Gaussian process. Then
$\forall\ 0 < a \le 1/2, \varepsilon > 0$

$$\mathbb{P}\Big(\sup_{0 \le t \le 1} |X_t| \le \varepsilon \Big) \le \exp\Big(- \frac{\varepsilon^4}{16a^2 \sum_{2 \le i,j \le 1/a} (\mathbb{E}\,(\xi_i \xi_j))^2} \Big)$$

provided that

$$a \sum_{2 \le i \le 1/a} \mathbb{E}\,\xi_i^2 \ge 32\varepsilon^2,$$

where $\xi_i = X(ia) - X((i-1)a)$ or
$\xi_i = X(ia) + X((i-2)a) - 2X((i-1)a)$.

Consequences

Let $\{X_t, t \in [0,1]\}$ be a centered Gaussian process with stationary increments and $X_0 = 0$. Put $\sigma^2(|t-s|) = \mathbb{E}\,|X_t - X_s|^2, \quad s, t \in [0,1]$. Assume that there are $1 < c_1 \leq c_2 < 2$ such that

$$c_1\sigma(h) \leq \sigma(2h) \leq c_2\sigma(h) \quad \text{for } 0 \leq h \leq 1/2.$$

Then there exists a positive and finite constant C such that

$$\mathbb{P}\left(\sup_{0\leq t\leq 1} |X_t| \leq \sigma(\varepsilon)\right) \leq \exp(-C/\varepsilon)$$

if one of the following conditions is satisfied. (i) σ^2 is concave on $(0,1)$; (ii)There is $c_0 > 0$ such that $(\sigma^2(x))''' \leq c_0 x^{-3}\sigma^2(x)$ for $0 < x < 1/2$.

Determinant method

The basic idea is finding lower bounds on the determinant for the covariance matrix.

$$
\begin{aligned}
& \mathbb{P}(\sup_{t\in T} |X(t)| < \varepsilon) \\
\leq \; & \mathbb{P}(\max_{1\leq i\leq n} |X(t_i)| < \varepsilon) \\
= \; & (2\pi)^{-n/2}(\det\Sigma)^{-1/2} \int_{\max_{1\leq i\leq n}|x_i|\leq\varepsilon} \exp\left(-\langle x, \Sigma^{-1}x\rangle\right) dx_1\cdots dx_n \\
\leq \; & (2\pi)^{-n/2}(\det\Sigma)^{-1/2}(2\varepsilon)^n \\
\leq \; & (\det\Sigma)^{-1/2}\varepsilon^n
\end{aligned}
$$

where $\Sigma = (\mathbb{E}\, X(t_i)X(t_j))_{1\leq i,j\leq n}$ is the covariance matrix.
•It is important to pick 'good' point $t_i, 1 \leq i \leq n = n_\varepsilon$

Lower bound on determinant

I believe this is the right algebraic approach for the Sheet problem ($d \geq 3$)
$\bullet$Given jointly Gaussian random variables $X_1, \cdots, X_n$, we denote by $\det \mathrm{Cov}(X_1, \cdots, X_n)$ the determinant of their covariance matrix. Then

$$\frac{(2\pi)^{n/2}}{\det \mathrm{Cov}(X_1, \cdots, X_n)} = \int_{\mathbb{R}^n} \mathbb{E} \exp\left(-i \sum_{j=1}^{n} y_j X_j\right) dy_1 \cdots dy_n.$$

and

$$\begin{aligned} \det \mathrm{Cov}(X_1, \cdots, X_n) &= \mathrm{Var}(X_1) \prod_{j=2}^{n} \mathrm{Var}(X_j | X_1, \cdots, X_{j-1}) \\ &\geq \cdots \end{aligned}$$

$\bullet$Local-nondeterminism, see Xiao and Wu (2008).

Small ball probability for smooth Gaussian processes

•Aurzada, Gao, Kühn, Li and Shao (2011): Small deviation probability for a family of smooth Gaussian processes with

$$\mathbb{E}\, X_{\alpha,\beta}(t) X_{\alpha,\beta}(s) = \frac{2^{2\beta+1}(ts)^{\alpha}}{(t+s)^{2\beta+1}}$$

for $\alpha > 0$ and $\beta > -1/2$.

Thm: For $\alpha > \beta > -1/2$,

$$-\log \mathbb{P}\left(\int_0^1 |X_{\alpha,\beta}(t)|^2\, dt \le \varepsilon^2\right) \sim \kappa_{\alpha,\beta} |\log \varepsilon|^3,$$

where the constant is given by $\kappa_{\alpha,\beta} := \frac{1}{3(\alpha-\beta)\pi^2}$.

For $\alpha > \beta + 1/2 > 0$,

$$\tilde{\kappa}_{\alpha,\beta} |\log \varepsilon|^3 \lesssim -\log \mathbb{P}\left(\sup_{t\in[0,1]} |X_{\alpha,\beta}(t)| \le \varepsilon\right) \lesssim \kappa_{\alpha-1/2,\beta} |\log \varepsilon|^3.$$

•Aurzada (2011+): Path regularity and small deviations of smooth Gaussian processes.

Gaussian Fields via Riesz Product

This is a combination of techniques in probability and analysis for the upper bound under the sup-norm for various Gaussian fields, see Gao and Li (2007). The basic ideas are

• Choosing Basis: Use (multi-dim) series expansion $X(t) = \sum_{n=1}^{\infty} f_n(t)\xi_n$, where ξ_n are i.i.d. standard normal random variables, and $f_n \in C([0,1]^d)$.

• Choosing Partial Sum: By Andersen's inequality, $\mathbb{P}(\|X\| \le \varepsilon) \le \mathbb{P}(\|Y\| \le \varepsilon)$ where $Y(t)$ is any partial sum $X(t) = \sum_{n\in E} f_n(t)\xi_n$.

• Construct Riesz Product:

$$\mathbb{P}(\|Y\| \le \varepsilon) \le \mathbb{P}(\int Y(t)R(t) \le \varepsilon)$$

where the Riesz product $R(t) = \prod_{n\in F}(1+\varepsilon_n h_n)$ satisfying $R(t) \ge 0$, $\|R\|_1 = \int R(t)dt = 1$.

A Symmetrization Inequality for Two Norms

Let $K \subset \mathbb{R}^d$ and $L \subset \mathbb{R}^d$ be two origin symmetric convex bodies, $\|\cdot\|_K$ and $\|\cdot\|_L$ be the corresponding gauges on $\mathbb{R}^d$, that is the norms for which K and L are the unit balls.
Let $C_+ = C_+(\|\cdot\|_K, \|\cdot\|_L, d, a, b,)$ be the optimal constant such that, for all $\mathbb{R}^d$-valued i.i.d. random variables X and Y, and $a, b > 0$,

$$\mathbb{P}(\|X+Y\|_L \le b) \le C_+ \cdot \mathbb{P}(\|X-Y\|_K \le a).$$

•For $d = 1$, it is not hard to show $C_+ \le \lceil 2b/a \rceil + 1$.
•Schultze and Weizsäcker (2007): For $d = 1$ and $a = b$, $C_+ = 2$ which answers an open problem for about 10 years.
•Dong, J. Li and Li (2011+):

$$C_+ \le N(B_L(b), B_K(a/2)),$$

and the bound are optimal for $\|\cdot\|_K = \|\cdot\|_L = \|\cdot\|_\infty$ with $C_+ = \lceil 2b/a \rceil^d$.

A Symmetrization Inequality for Two Norms

Let $K \subset \mathbb{R}^d$ and $L \subset \mathbb{R}^d$ be two origin symmetric convex bodies, $\|\cdot\|_K$ and $\|\cdot\|_L$ be the corresponding gauges on $\mathbb{R}^d$, that is the norms for which K and L are the unit balls.

Let $C_+ = C_+(\|\cdot\|_K, \|\cdot\|_L, d, a, b,)$ be the optimal constant such that, for all $\mathbb{R}^d$-valued i.i.d. random variables X and Y, and $a, b > 0$,

$$\mathbb{P}(\|X+Y\|_L \le b) \le C_+ \cdot \mathbb{P}(\|X-Y\|_K \le a).$$

•For $d = 1$, it is not hard to show $C_+ \le \lceil 2b/a \rceil + 1$.

•Schultze and Weizsäcker (2007): For $d = 1$ and $a = b$, $C_+ = 2$ which answers an open problem for about 10 years.

•Dong, J. Li and Li (2011+):

$$C_+ \le N(B_L(b), B_K(a/2)),$$

and the bound are optimal for $\|\cdot\|_K = \|\cdot\|_L = \|\cdot\|_\infty$ with $C_+ = \lceil 2b/a \rceil^d$.

•For i.i.d r.v $X_i \in \mathbb{R}^d$, $i \ge 1$, we have as an application

$$\mathbb{P}(\|\sum^{2n} X_i\|_L \le b) \le C_+ \cdot \mathbb{P}(\|\sum^{n}(X_i - X_{n+i})\|_K \le a).$$

An Extension of 123 Theorem

Let $C_- = C_-(\|\cdot\|_K, \|\cdot\|_L, d, a, b,)$ be the optimal constant such that, for all $\mathbb{R}^d$-valued i.i.d. random variables X and Y, and $b > a > 0$,

$$\mathbb{P}(\|X - Y\|_L \leq b) \leq C_- \cdot \mathbb{P}(\|X - Y\|_K \leq a).$$

•Alon and Yuster (1995) and (independently) Kotlov: For $d = 1$, $C_- \leq 2\lceil b/a \rceil - 1$. In particular, for $a = 1$, $b = 2$, we have $C_- = 3$.
•Alon and Yuster (1995): For $\|\cdot\|_K = \|\cdot\|_L = \|\cdot\|_2$, $C_- \leq M$ if there is no set F of $M + 1$ points in a ball of radius b so that the center belongs to F and the distance between any two pints of F exceeds a. In addition, $C_- = M$ in special settings.
•Dong, J. Li and Li (2011+):

$$C_- \leq N(B_L(b) \setminus B_K(a), B_K(a/2)) + 1,$$

and the bound is optimal for $d = 1$.

Our approach for both problems (C_+ and C_-) extends techniques developed in Schultze and Weizsäcker (2007) which starts with the following fact:

Lemma: The following two statement are equivalent for a given symmetric matrix $A = (a_{ij})_{n\times n}$:

(i) For all probability measure $p \in \mathcal{P} := \{p : \sum p_i = 1, p_i \geq 0\}$,

$$p^T Ap = \sum_{i,j} a_{ij} p_i p_j > 0;$$

(ii) For all $p \in \mathcal{P}$,

$$\max_i \sum_j a_{ij} p_i p_j > 0.$$

Our approach for both problems (C_+ and C_-) extends techniques developed in Schultze and Weizsäcker (2007) which starts with the following fact:

Lemma: The following two statement are equivalent for a given symmetric matrix $A = (a_{ij})_{n\times n}$:

(i) For all probability measure $p \in \mathcal{P} := \{p : \sum p_i = 1, p_i \geq 0\}$,

$$p^T Ap = \sum_{i,j} a_{ij} p_i p_j > 0;$$

(ii) For all $p \in \mathcal{P}$,

$$\max_i \sum_j a_{ij} p_i p_j > 0.$$

•Idea of Pf: Consider Lagrange multiplier $L(p) = p^T Ap + \lambda(1 - p^T \mathbf{1})$ for the minimum in (i). Then

$$\frac{\partial L(p)}{\partial p} = 2Ap^* - \lambda p^* = 0.$$

Thus (ii) implies $\lambda > 0$ and $\sum_j a_{ij} p_j^* = \lambda p_i^*/2 > 0$ for all i.

•The above fact can be reformulated in the infinite dimensional setting for product measure.

Fact: Let $(\Omega, \mathcal{B})$ be a measurable space and let $f : \Omega \times \Omega \to \mathbb{R}$ be a $\mathcal{B} \otimes \mathcal{B}$ measurable bounded symmetric function. Let $\mathcal{P}$ be the set of all probability measures on $\mathcal{B}$.

Then the following are equivalent:

(1). For all $\mu \in \mathcal{P}$

$$\int_{\Omega \times \Omega} f(x, y)\mu(dx)\mu(dy) > 0.$$

(2). For all $\mu \in \mathcal{P}$

$$\mu\left(\int_{\Omega} f(\cdot, y)\mu(dy) > 0\right) > 0.$$

Proof of $C_+ \le N(B_L(b), B_K(a/2)) = N_+$

We need to show that, for any constant $C > N_+$,

$$\mathbb{P}(\|X + Y\|_L \le b) < C \cdot \mathbb{P}(\|X - Y\|_K \le a)$$

for any two i.i.d. $\mathbb{R}^d$-valued random variables X and Y. The above inequality can be rewritten as $\int_{\mathbb{R}^d \times \mathbb{R}^d} g(x, y)\mu(dx)\mu(dy) > 0$ where the function $g(x, y)$ is defined as

$$g(x, y) = C \cdot 1_{\{\|x-y\|_K \le a\}} - 1_{\{\|x+y\|_L \le b\}}, \quad \text{for } x, y \in \mathbb{R}^d$$

and μ is the probability measure on $\mathbb{R}^d$ induced by X. By the key Lemma, we need to show

$$\mathbb{P}\left(\int_{\mathbb{R}^d} g(\cdot, y)\mu(dy) > 0\right) > 0 \quad \text{for all} \ \ \mu \in \mathcal{P}.$$

Assume otherwise, then there exists a probability measure $\mu \in \mathcal{P}$ (i.e. random variables X, Y) such that $\mu(D) = 1$, where the set D is defined by

$$\begin{aligned} D &= \{x \in \mathbb{R}^d : \int_{\mathbb{R}^d} g(x, y)\mu(dy) \le 0\} \\ &= \{x \in \mathbb{R}^d : \mu(-x + B_L(b)) \ge C \cdot \mu(x + B_K(a))\} \end{aligned}$$

Define

$$\alpha = \sup_{x \in D} \mu(x + B_K(a)).$$

For $0 < \epsilon < (1 - N_+/C)\alpha$, pick $x_0 \in D$ such that

$$\mu(-x_0 + B_L(b)) \geq C \cdot \mu(x_0 + B_K(a)) > C \cdot (\alpha - \epsilon) > 0.$$

Using the fact that the set $-x_0 + B_L(b)$ can be covered by N_+ balls of size $B_k(a/2)$, we have

$$N_+ \cdot \sup_{x \in \mathbb{R}^d} \mu(x + B_K(a/2)) \geq \mu(-x_0 + B_L(b)) \geq C \cdot (\alpha - \epsilon) > 0.$$

Since $\mu(D) = 1$, we have for any set $x + B_K(a/2)$ with positive measure under μ, there is

$$x^* \in (x + B_K(a/2)) \cap D.$$

Thus we have the covering

$$(x + B_K(a/2)) \subset x^* + B_K(a)$$

which implies $N_+ \cdot \sup_{x \in D} \mu(x + B_K(a) \geq C(\alpha - \epsilon) > N_+\alpha$. This contradicts the definition of $\alpha = \sup_{x \in D} \mu(x + B_K(a)$.

1-Dimensional Case

Let X and Y be i.i.d. real random variables, then for positive reals a, b, we have

$$\mathbb{P}(|X+Y| \leq b) \leq \lceil 2b/a \rceil \cdot \mathbb{P}(|X-Y| \leq a).$$

Moreover, the constant $\lceil 2b/a \rceil$ is sharp.

Proof: It is easy to see that the interval $[-x-b, -x+b]$ can be covered by $\lceil 2b/a \rceil$ intervals in the form $[y-a/2, y+a/2]$, for $x, y \in \mathbb{R}$ and thus $N_+ = \lceil 2b/a \rceil$. To show the constant $\lceil 2b/a \rceil$ is the best possible, we only need to consider the following example. Let X, Y be independent and have the same distribution $\mathbb{P}(X = x_i) = \frac{1}{2n}$ with

$$x_i = \begin{cases} i(1+\epsilon)\delta, & i = 1, 2, \cdots, n \\ i(1+\epsilon)\delta - r & i = 0, -1, \cdots, -n+1, \end{cases}$$

where $\delta = a/b$, $\epsilon > 0$ is sufficiently small and $0 \leq r \leq \frac{1}{2}(1+\epsilon)\delta$. Then, it is clear

$$\mathbb{P}(|X-Y| \leq \delta) = \mathbb{P}(X = Y) = 1/2n.$$

For fixed $\delta > 0$ and n large enough, we have

$$\begin{aligned}\mathbb{P}(|X+Y| \le 1) &= \frac{1}{2n}\sum_{i=-n+1}^{n}\mathbb{P}(-1-x_i \le X \le 1-x_i) \\ &= \frac{1}{2n}\Big(\sum_{i\in I_1}+\sum_{i\in I_2}+\sum_{i\in I_3}\Big)\mathbb{P}(-1-x_i \le X \le 1-x_i)\end{aligned}$$

where I_1, I_2, I_3 defined below are three disjoint subsets of the summation index set $\{i : -n+1 \le i \le n\}$.

$$\begin{aligned}I_1 &= \{i : x_{-n+1} \le -1-x_i < 1-x_i \le x_0\} \\ I_2 &= \{i : x_1 \le -1-x_i < 1-x_i \le x_n\} \\ I_3 &= \{i : -1-x_i < x_{-n+1},\ \text{or}\ -1-x_i < x_0 < 1-x_i,\ \text{or}\ 1-x_i > x_n\}\end{aligned}$$

•We then need a careful counting of each index set, depending on various range of δ and r (very technical).
•Finally we can see that

$$C_+(\mathbb{R}, a/b, 1, |\cdot|, |\cdot|) \ge \lceil 2/\delta \rceil = \lceil 2b/a \rceil.$$

and So $C_+(\mathbb{R}, a, b, |\cdot|, |\cdot|) = \lceil 2b/a \rceil$.