Non-Asymptotic Theory of Random Matrices
Lecture 12: Sudakov's Minoration

Lecturer: Roman Vershynin Scribe: Blake Hunter

Thursday, February 13, 2007

1 Sudakov's Minoration

Sudakov's Minoration provides a weak converse to Dudley's Inequality (there exists no strong converse).

Theorem 1. *Let $(x_t)_{t\in T}$ be a Gaussian Process on a metric space (T,d) s.t.*

$$\|x_t - x_s\|_2 = d(t,s) \quad \forall t,s \in T$$

Then

$$c\epsilon\sqrt{\log N(T,\epsilon)} \leq \mathbb{E}\sup_{t\in T} x_t \leq c\int_0^\infty \sqrt{\log N(T,\epsilon)}d\epsilon \quad \forall \epsilon > 0$$

Remark 2. *Note that we can start from an abtsract set T, and define the metric d as above.*

Proof. **1)** Discretize T: Recall the greedy algorithm to construct an ϵ-net. There exists $\mathcal{N} \subseteq T$, $|N| = N(T,\epsilon)$ an ϵ-net of T s.t.

$$d(t,s) > \epsilon \quad \forall t,s \in \mathcal{N} \quad t \neq s$$

Then

$$\mathbb{E}\sup_{t\in T} x_t \geq \mathbb{E}\sup_{t\in\mathcal{N}} x_t$$

2) Comparison using Slepian's Lemma: Compare $(x_t)_{t\in N}$ with $(y_t)_{t\in N}$ where $y_t = \frac{\epsilon}{\sqrt{2}}g_t$ and g_t are i.i.d. standard Gaussian random variables.
We look at the increments :

$$\|y_t - y_s\|_2 = \|g_t - g_s\|_2 = \sqrt{\|g_t\|_2^2 + \|g_s\|_2^2} = \sqrt{2} \quad s.t. \quad \frac{\epsilon}{\sqrt{2}} = \epsilon$$

Slepian's Lemma applies,

$$\mathbb{E}\sup_{t\in T} x_t \geq \mathbb{E}\sup_{t\in\mathcal{N}} y_t = \frac{\epsilon}{\sqrt{2}}\mathbb{E}\max_{k=1,2,\ldots,|\mathcal{N}|} g_k \leq c\epsilon\sqrt{\log|\mathcal{N}|}$$

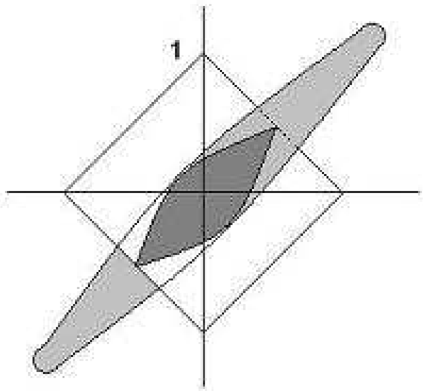
□

Lemma 3. *Let $g_1, ..., g_N$ be i.i.d. standard Gaussian random variables. Then*

$$c - 1\sqrt{\log N} \leq \mathbb{E} \max_{k=1,..N} |g_k| \leq \sqrt{\log N}$$

Sections of convex sets: Let $T \subseteq \mathbb{R}^n$ be a convex set. We can improve T by taking its random section $T \cap E$ where E is a random subspace of $\mathbb{R}^n$ uniform in the Grassmanian of a given dimension. Now we wish to improve the diameter of T.

Example 4. *Thin long sausage.*



Theorem 5 (Low M^*-estimate)**.** *Let T be a convex set in $\mathbb{R}^n$ and let E be a random subspace of $\mathbb{R}^n$ of codimension k.*
Then

$$diam(T \cap E) \leq c\frac{l(T)}{\sqrt{k}} \quad \text{with probability } \geq 1 - \mathrm{e}^{-k}$$

where $l(T)$ is the mean width of T, $l(T) = \mathbb{E}\sup_{t\in T}\langle g, t\rangle$, $g \in \mathbb{R}^n$ is a standard Gaussian vector.

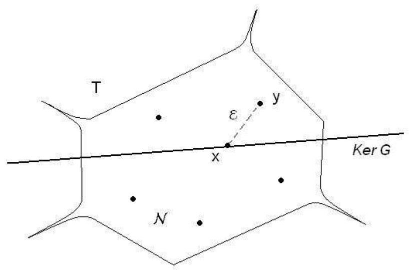
Exercise 6. *Compute $l(T)$ for the sausage in the example above.*

Proof : via random matrices. We can realize E as the kernel of an $k \times n$ Gaussian Matrix $G : \mathbb{R}^n \to \mathbb{R}^k$. Both E and $\ker G$ are uniformly distributed in the Grassmanian $G_{n,n-k}$ (by rotational invariance).

1) Discretize T using Sudakov's Minoration : There exists ϵ-net $\mathcal{N}$ of T with $\epsilon\sqrt{\log |\mathcal{N}|} \leq l(T)$. Thus

$$|\mathcal{N}| \leq \mathrm{e}^{cl(T)^2/\epsilon^2}.$$

2) Take $x \in T \cap \ker G$. We want to show $\|x\|_2 \lesssim \frac{l(T)}{\sqrt{k}}$.
3) Approximate x with $y \in \mathcal{N}$ $\|x-y\|_2 \le \epsilon$



4) We know $Gx = 0$, thus Gy has to be small too,

$$\|Gy\|_2 = \|G(x-y)\|_2 \le A$$

(one can always bound $A \le \|G\| \cdot \|x-y\|_2 \lesssim \sqrt{n} \cdot \epsilon$)
5) On the other hand, for all $y \in \mathcal{N}$

$$\|Gy\|_2 \ge B\|y\|_2$$

6) Hence $\|y\|_2 \le \frac{A}{B}$
7) Since $\|x-y\|_2 < \epsilon$, we have

$$\|x\|_2 \le \frac{A}{B} + \epsilon$$

□

Now choose $\epsilon = \frac{l(T)}{\sqrt{k}}$ so that $|\mathcal{N}| \le \mathsf{e}^c$.
A) We will give a better bound on A. (See [1].) Set z, A, and the convex set K to be:

$$z := x - y \ \in \ \epsilon B_2^n \cap 2T =: K$$

$$A := \max_{z \in K} \|Gz\|_2$$

Lemma 7. *$A \le l(T)$ with probability $1 - \mathsf{e}^{-k}$.*

Proof. Discretize B_2^k : find an $\frac{1}{2}$-net M of B_2^k with $|M| \leq 5^k$. Then,

$$A \leq 2 \max_{u \in M} \langle Gz, u \rangle$$

(see Lecture 3.) And so,

$$\mathbb{P}(A > t) \leq 5^k \mathbb{P}(\max_{z \in K} \langle Gz, u \rangle \ > t) \, (*)$$

where a unit vector u is fixed. Then also,

$$\langle Gz, u \rangle = \langle z, G^* u \rangle = \langle z, g \rangle$$

where g is a standard Gaussian in $\mathbb{R}^n$. So the above expression $(*)$ is less than or equal to $5^k \mathbb{P}\left(\max_{z \in K} \langle g, z \rangle > t\right)$. We next use Gaussian concentration of measure. The map $g \longmapsto f(g) = \max_{z \in K} \langle g, z \rangle$ is an ϵ-Lipschitz function ($z \in K \subseteq \epsilon B_2^n$). Thus

$$\mathbb{P}(f - \mathbb{E}f > t) \leq \mathsf{e}{-t^2/2\epsilon^2}$$

$$\mathbb{E}f \leq \max_{x \in 2T} \langle g, z \rangle = 2l(T)$$

$$\mathbb{P}(A > 2l(T) + t) \leq 5^k \mathsf{e}^{-t^2/2\epsilon^2}$$

Choosing $t = 10\, l(T)$ yields the result. □

B)

Lemma 8. *With probability at least* $1 - \mathsf{e}^{-k}$,

$$\|Gy\|_2 \geq c\sqrt{k}\|y\|_2 \quad \forall y \in \mathcal{N}$$

This shows that $B \approx \sqrt{k}$, which would complete the proof since by the above lemma,

$$\|x\|_2 \leq \frac{A}{B} + \epsilon \lesssim \frac{l(T)}{\sqrt{k}}$$

Proof. 1) Fix $y \in \mathcal{N}$. $G\left(\frac{y}{\|y\|_2}\right)$ is distribution identically with g, standard Gaussian in $\mathbb{R}^k$. Then

$$\begin{aligned}\mathbb{P}(\|Gy\|_2 < \delta\sqrt{k}\|y\|_2) &= \mathbb{P}(\|g\|_2 \leq \delta\sqrt{k}) \\ &\leq (cd)^k\end{aligned}$$

which we will prove later.
2) Finally, we apply the union bound:

$$\mathbb{P}(\exists y \in \mathcal{N} : \|Gy\| \leq \delta\sqrt{k}\|y\|) \leq |\mathcal{N}| \cdot c\delta^k \leq \mathsf{e}^{-k}$$

by choosing δ small. □

References

[1] A. Litvak ; A. Pajor; M. Rudelson; N. Tomczak-Jaegermann. Smallest singular values of random matrices and geometry of random polytopes. *Advances in Mathematics*, 195:491, 2005.