Non-Asymptotic Theory of Random Matrices
Lecture 13: Sections of convex sets via entropy and volume

Lecturer: Roman Vershynin   Scribe: Yuji Nakatsukasa

Thursday, February 15th, 2007

1 Application of Low $M^\star$-estimate

Recall the Low $M^\star$-estimate from the previous lecture:

$$\text{Let } T \subseteq \mathbb{R}^n \text{be convex and symmetric,}$$

$$\text{and let } E \text{ be a random subspace of } \mathbb{R}^n, \text{with codimension } k.$$

Then,

$$\operatorname{diam}(T \cap E) \leq C\frac{l(T)}{\sqrt{k}}, \text{with high probability.}$$

Here, $l(T) = \mathbb{E}\sup\limits_{t\in T} < g, t >$, where g is a gaussian vector.
The following Example shows an application of the Low $M^\star$-estimate.

Example 1 (l_p-balls)**.** *Let* B_p^n $=$*Ball*$(l_p^n) = \{x \in \mathbb{R}^n; ||x||_p \leq 1\})$*, where* $1 < p \leq 2$ *(note that* p *cannot be 1 here). We want to evaluate how "spiky", or how "round"* B_p^n *is. We answer this question by examining the radii of the inscribed and circumscribed balls. The closer the two radii are, the rounder* B_p^n *is.*

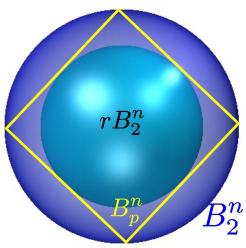


Figure 1: Inscribed (rB_2^n), Circumscribed (B_2^n) balls of B_p^n.

We can form the inequality

$$||x||_2 \leq ||x||_p \leq n^{\frac{1}{p}-\frac{1}{2}}. \text{ (Check)}$$

$$\Rightarrow r \cdot B_2^n \subseteq B_p \subseteq B_2^n.$$

Here, $r = n^{\frac{1}{2}-\frac{1}{p}} \to 0$ as $n \to \infty$.

This result implies that B_p^n is not round. However, it is easily seen that the "spikes" are in the coordinate directions, and our intuition is that there are only $2n$ spikes (which are not exponentially many), so they can be eliminated. In fact,

$$\sup_{t \in B_p^n} < g, t >= \sup_{t \in Ball(l_p^n)} < g, t >= ||g||_{(l_p^n)^*} = ||g||_{l_q^n},$$

$$\text{where} \quad \frac{1}{p} + \frac{1}{q} = 1.$$

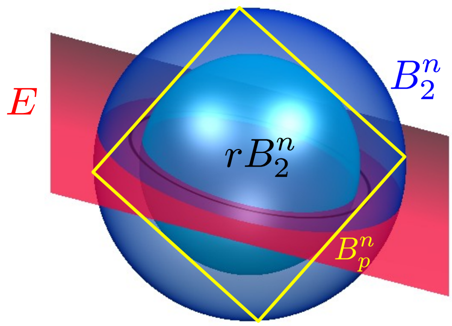
Here $(l_p^n)^*$ denotes the dual space of (l_p^n). Therefore,

$$\begin{aligned} l(B_p^n) &= \mathbb{E}||g||_q = \mathbb{E}(\sum_{i=1}^n |g_i|^q)^{1/q} \\ &\leq (\mathbb{E}\sum_{i=1}^n |g_i|^q)^{1/q} \quad (\text{Jensen's Inequality}) \\ &= n^{1/q}(\mathbb{E}|g_i|^q)^{1/q}. \end{aligned}$$

Now we apply Low $M^\star$-estimate: for $k = \delta n$ $(0 < \delta < 1)$,

$$\operatorname{diam}(B_p^n \cap E) \leq c_q \frac{n^{1/q}}{\sqrt{\delta n}} = C_{q,\delta} n^{\frac{1}{q}-\frac{1}{2}} = C_{q,\delta} r, \tag{1}$$

with high probability. The result implies that the intersection of E and B_p^n is likely to be small when n is large ($r \to 0$ as $n \to \infty$).



From this result follows the Corollary:

Corollary 2 (Almost round sections of $l_p - Balls$)**.**

$$\textit{Let} \quad 1 < p \leq 2, \quad 0 < \delta < 1.$$

Also let E be a random subspace of $\mathbb{R}^n$ with codimension δn.

$$\textit{Then} \quad r(B_2^n \cap E) \subseteq (B_p^n \cap E) \subseteq C_{q,\delta} r(B_2^n \cap E), \quad \textit{with high probability.}$$

This can be expressed equivalently as

$$C_{q,\delta}||x||_2 \leq r||x||_p \leq ||x||_2 \quad \forall x \in E.$$

This implies $||x||_p \sim ||x||_2$, that is those two norms are equivalent. Since this holds with high probability, the ball B_p^n is almost round.
However, in the case when $p = 1$, the Low $M^\star$-estimate turns out not to be good enough.

Example 3 (l_1-ball)**.**

$$\text{Let} \quad r = \frac{1}{\sqrt{n}}, \;\; q = \infty.$$

$$l(B_1) = \mathbb{E}||g||_\infty \sim \sqrt{\log n}.$$

Applying Low $M^\star$-estimate gives

$$\operatorname{diam}(T \cap E) \leq \frac{c\sqrt{\log n}}{\sqrt{n}} = c_q \sqrt{\log n} \cdot r \tag{2}$$

Compared with (1), we observe that there is some discrepancy in this estimate ($\sqrt{\log n}$). Therefore, Low $M^\star$-estimate fails to produce nice sections of B_1^n. To deal with this problem, we want to replace mean width by a better "size" of T. In order to do so, the Volume Ratio Theorem in the next section provides an estimate with respect to the Volume. This approach via Volume and entropy is originally seen in [1].

2 Volume Ratio Theorem and Entropy Theorem

Theorem 4 (Volume Ratio Theorem)**.** *Let $T \subset R^n$ be a convex, symmetric set, and $B_2^n \subseteq T$. The Volume Ratio is defined as*

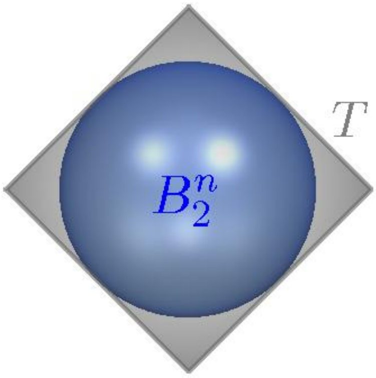
$$V(T) = \left(\frac{\mathsf{Vol}(T)}{\mathsf{Vol}(B_2^n)} \right)^{1/n}.$$

Also let

$$E \text{ be a random subspace of } \mathbb{R}^n \text{ with codimension } \delta\text{n}.$$

Then, with probability $\geq 1 - e^{-n}$,

$$\operatorname{diam}(T \cap E) \leq C(V(T), \delta).$$



Remarks:

1. $C(V(T), \delta) \leq (CV(T))^{1/\delta}$.

2. Compare with Low $M^\star$-estimate: From Uryson's Inequality, $V(T) \leq \frac{l(T)}{\sqrt{n}}$. Therefore this is a better estimate.

Notice that the Theorem enables us to eliminate the $\sqrt{\log n}$ term in (2). Details on Volume Ratio Theorem can be found in [2]. Instead of proving this result, we state a stronger result, which replaces volume with entropy;

Theorem 5 (Entropy)**.** *Let $T \subset R^n$ be a convex, symmetric set. Suppose*

$$N(T, B_2^n) \leq V^n \quad \textit{for some } V > 1.$$

Also let E be a random subspace of codimension δn. Then, with probability $1 - e^{-n}$,

$$\operatorname{diam}(T \cap E) \leq C(V, \delta).$$

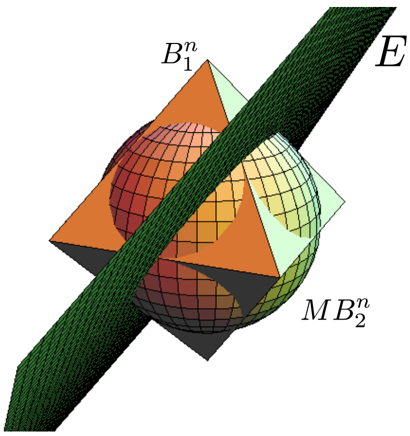
Remark that Volume Ratio Theorem follows from the Entropy Theorem:

$$\begin{aligned}\frac{\mathsf{Vol}(T)}{\mathsf{Vol}(B_2^n)} \leq N(T, B_2^n) &\leq \frac{\mathsf{Vol}(T + B_2^n)}{\mathsf{Vol}(B_2^n)} \quad \text{(lecture 6)} \\ &\leq \frac{\mathsf{Vol}(2T)}{\mathsf{Vol}(B_2^n)} \quad (B_2^n \subseteq T) \\ &= 2^n \frac{\mathsf{Vol}(T)}{\mathsf{Vol}(B_2^n)}.\end{aligned}$$

From this it follows that

$$(V(T))^n \leq N(T, B_2^n) \leq (2V(T))^n.$$

Therefore, we conclude that the Volume Ratio Theorem can be derived from the Entropy Theorem. We will therefore prove the Entropy Theorem. The idea is as follows;



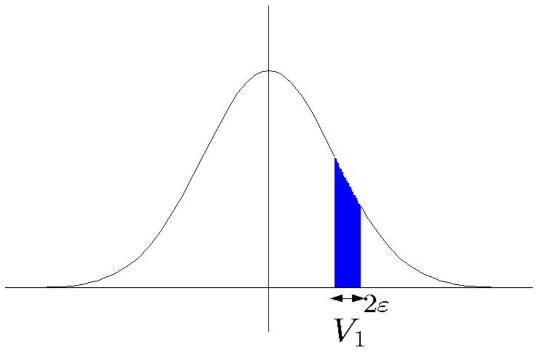
We want to repel the subspace E from the spikes. The idea is to discretize the tentacles and show that they can be eliminated (Lecture 14).

In order to prove the Entropy Theorem, we begin with the Small Ball Probability:

Lemma 6 (Small Ball Probability)**.** *A standard Gaussian Vector g in $\mathbb{R}^n$ is unlikely to be in a Euclidean ball with radius $\ll \sqrt{n}$.*

- Dim 1. $\mathbb{P}(|g - v_1| < \varepsilon) \sim \varepsilon, \quad \forall \varepsilon > 0, v_1 \in \mathbb{R}$. (Exercise)
- Dim n. $\mathbb{P}(||g - V||_2 < \varepsilon\sqrt{n}) \leq (c'\varepsilon)^n, \quad \forall \varepsilon > 0, V \in \mathbb{R}^n$.

To prove Dimension n, we state the following Lemma;



Lemma 7 (Tensorization Lemma)**.** *Let X be a random variable. Assume*

$$\mathbb{P}(|X - V| < \varepsilon) \leq C\varepsilon, \quad \forall \varepsilon > 0, V \in \mathbb{R}.$$

Let $X_1, X_2, \cdots, X_n$ be independent, identically distributed copies of X. Then

$$\mathbb{P}(\sum_{j=1}^{n} |X_j - V_j| < \varepsilon^2 n) \leq (C'\varepsilon)^n, \quad \forall \varepsilon > 0, V_i \in \mathbb{R}.$$

Note that this Lemma immediately implies the Dimension n of Small Ball probability.
Proof : Let $|X_j - V_j| = Y_j$.

$$\begin{aligned}
\mathbb{P}(\sum_j Y_j^2 < \varepsilon^2 n) &= \mathbb{P}(n - \frac{1}{\varepsilon^2}\sum_j Y_j^2 > 0) \\
&= \mathbb{P}(\exp(n - \frac{1}{\varepsilon^2}\sum_j Y_j^2) > 1) \\
&\leq \mathbb{E}\exp(n\frac{1}{\varepsilon^2}\sum_j Y_j^2) \quad \text{(Markov's Inequality)} \\
&= e^n \prod_{j=1}^{n} \exp(-Y_j^2/\varepsilon^2). \qquad (\divideontimes)\text{(Independence)}
\end{aligned}$$

Here, since

$$\text{if } X \geq 0, \quad \text{then} \quad \mathbb{E}\mathrm{X} = \int_0^\infty \mathbb{P}(\mathrm{X} > \mathrm{s})\mathrm{ds},$$

$$\begin{aligned}\exp(-Y_j^2/\varepsilon^2) &= \int_0^\infty \mathbb{P}(\exp(-Y_j^2/\varepsilon^2) > s)ds \\ &= \int_0^1 \mathbb{P}(\exp(-Y_j^2/\varepsilon^2) > s)ds. \quad (\exp(-Y_j^2/\varepsilon^2) \leq 1)\end{aligned}$$

Applying change of variables $s = e^{-u^2}$, then $x \in [0,1] \Rightarrow u \in [0,\infty]$, and $ds = -2ue^{-u^2}$. Therefore we have

$$\exp(-Y_j^2/\varepsilon^2) = \int_0^\infty \mathbb{P}(Y_j < \varepsilon u)2ue^{-u^2}du \leq C''\varepsilon.$$

Here we used used Dimension 1 in the last inequality ($\mathbb{P}(Y_j < \varepsilon u) = \mathbb{P}(|X - V_i| < \varepsilon u) \leq C\varepsilon u$). Using this inequality in ($\circledast$) yields

$$\mathbb{P}(\sum_j Y_j^2 > \varepsilon^2 n) \leq e^n(C''\varepsilon)^n = (C''e\varepsilon)^n,$$

which completes the proof of the Tensorization Lemma. In the next lecture we will apply the small ball probability to complete the proof of the Entropy Theorem.

References

[1] V. D. Milman and A Pajor. Eigenvalues and condition numbers of random matrices. *Studia Math*, 159(2):247–261, 2003.

[2] Gilles Pisier. *The Volume of Convex Bodies and Banach Space Geometry.* Cambridge University Press, 1999.