The Random Version of Dvoretzky's Theorem in ℓ_∞^n

G. Schechtman*

Department of Mathematics, Weizmann Institute of Science, Rehovot, Israel
gideon.schechtman@weizmann.ac.il

Summary. We show that with "high probability" a section of the ℓ_∞^n ball of dimension $k \le c\varepsilon \log n$ ($c > 0$ a universal constant) is ε close to a multiple of the Euclidean ball in this section. We also show that, up to an absolute constant the estimate on k cannot be improved.

1 Introduction

Milman's version of Dvoretzky's theorem states that:
There is a function $c(\varepsilon) > 0$ *such that for all* $k \le c(\varepsilon) \log n$, ℓ_2^k $(1+\varepsilon)$*-embeds into any normed space of dimension* n.

See [Dv] for the original theorem of Dvoretzky (in which the dependence of k on n is weaker), [Mi] for Milman's original work, and [MS] and [Pi] for expository outlets of the subject (there are many others). It would be important for us to notice that the proof(s) of the theorem above actually give more: The vast majority of subspaces of the stated dimension are $(1+\varepsilon)$-isomorphic to ℓ_2^k.

The dependence of k on n in the theorem above is known to be best possible (for ℓ_∞^n) but the dependence on ε is far from being understood. The best known estimate is $c(\varepsilon) \ge c\varepsilon/(\log \frac{1}{\varepsilon})^2$ given in [Sc] (here and elsewhere in this paper c and C denote positive universal constants). However, the proof in [Sc] does not give the additional information that *most* subspaces are $(1+\varepsilon)$-isomorphic to ℓ_2^k. If one also want this requirement then the best estimate for $c(\varepsilon)$ that was known was $c(\varepsilon) \ge c\varepsilon^2$ ([Go]).

As an upper bound for $c(\varepsilon)$ one gets $C/\log \frac{1}{\varepsilon}$ for some universal C. Indeed, if ℓ_2^k $(1+\varepsilon)$ embed into ℓ_∞^n then $k \le C \log n / \log \frac{1}{\varepsilon}$. This is also the right order of k in the ℓ_∞ case: If $k \le c \log n / \log \frac{1}{\varepsilon}$ then ℓ_2^k $(1+\varepsilon)$ embed into ℓ_∞^n.

We show here that, in the ℓ_∞ case, if one is interested in the probabilistic statement of Dvoretzky theorem (i.e, that the vast majority of subspaces of

* Supported by the Israel Science Foundation

ℓ_∞^n of a certain dimension are $(1+\varepsilon)$-isomorphic to Euclidean spaces) then the right estimate for $c(\varepsilon)$ is $c\varepsilon$.

Theorem 1. *For $k < c\varepsilon \log n$, with probability $> 1 - e^{-ck}$, the ℓ_∞^n norm and a multiple of the ℓ_2^n norm are $1+\varepsilon$ equivalent on a k dimensional subspace. Moreover, this doesn't hold anymore for k of higher order. i.e., For every a there is an A such that if, with probability larger than $1-e^{-ak}$, a k dimensional subspace satisfies that the ratio between the ℓ_∞^n norm and a multiple of the ℓ_2^n norm are $1+\varepsilon$ equivalent for all vectors in the subspace, then $k \le A\varepsilon \log n$.*

2 Computation of the Concentration of the Max Norm

Let $g_1, g_2, \dots$ be a sequence of standard independent Gaussian variables. fix n and let M be the median of $\|(g_1, g_2, \dots, g_n)\|_\infty$. In this section we compute some fine estimates on the probability of deviation of $\|(g_1, g_2, \dots, g_n)\|_\infty$ from M.

Claim 1.

$$\big(1 - 2^{-1/n}\big)\frac{\sqrt{\pi}M}{\sqrt{2}} \le e^{-M^2/2} \le (1 - 2^{-1/n})\frac{\sqrt{\pi}(M+1)}{\sqrt{2}(1 - e^{-\frac{1}{2}}e^{-M})}. \tag{1}$$

Proof.

$$\frac{1}{2} = P\Big(\max_{1\le i\le n} |g_i| < M\Big) = \bigg(1 - \sqrt{\frac{2}{\pi}}\int_M^\infty e^{-s^2/2}ds\bigg)^n.$$

Consequently,

$$\begin{aligned} 1 - 2^{-1/n} &= \sqrt{\frac{2}{\pi}}\int_M^\infty e^{-s^2/2}ds \ge \sqrt{\frac{2}{\pi}}\frac{1}{M+1}\int_M^{M+1} se^{-s^2/2}ds \\ &\ge \sqrt{\frac{2}{\pi}}\frac{1}{M+1}e^{-M^2/2}(1 - e^{-\frac{1}{2}}e^{-M}), \end{aligned} \tag{2}$$

or

$$e^{-M^2/2} \le (1 - 2^{-1/n})\frac{\sqrt{\pi}(M+1)}{\sqrt{2}(1 - e^{-\frac{1}{2}}e^{-M})}. \tag{3}$$

Similarly,

$$1 - 2^{-1/n} = \sqrt{\frac{2}{\pi}}\int_M^\infty e^{-s^2/2}ds \le \sqrt{\frac{2}{\pi}}\frac{1}{M}\int_M^\infty se^{-s^2/2}ds \le \sqrt{\frac{2}{\pi}}\frac{e^{-M^2/2}}{M},$$

or

$$e^{-M^2/2} \ge (1 - 2^{-1/n})\frac{\sqrt{\pi}M}{\sqrt{2}}. \tag{4}$$

□

Claim 2.

$$\frac{\log 2}{4+\log 2}e^{-3\varepsilon M^2/2} \le P\Big(\max_{1\le i\le n}|g_i| > (1+\varepsilon)M\big) \le \log 2(1+o(1)\Big)e^{-\varepsilon M^2} \quad (5)$$

where $o(1)$ means $a(n)$ with $a(n) \to 0$ as $n \to \infty$ independently of ε.

Proof. (3) implies

$$\begin{aligned} &P\Big(\max_{1\le i\le n}|g_i| > (1+\varepsilon)M\Big) \\ &\le \sqrt{\frac{2}{\pi}}\frac{n}{(1+\varepsilon)M}e^{-(1+\varepsilon)^2M^2/2} \\ &\le \frac{n}{(1+\varepsilon)M}(1-2^{-1/n})\frac{M+1}{1-e^{-\frac{1}{2}}e^{-M}}e^{-\varepsilon M^2}e^{-\varepsilon^2M^2/2} \end{aligned} \quad (6)$$

and, since M is of order $\sqrt{\log n}$, we get from this that

$$P\Big(\max_{1\le i\le n}|g_i| > (1+\varepsilon)M\Big) \le \log 2\big(1+o(1)\big)e^{-\varepsilon M^2}. \quad (7)$$

(For a fixed ε one can replace $\log 2(1+o(1))$ with a quantity tending to 0 with n.)

We now look for a lower bound on $P(\max_{1\le i\le n}|g_i| > (1+\varepsilon)M)$. Since for iid X_i-s,

$$\begin{aligned} P\Big(\max_{1\le i\le n}(X_i > t\Big) &= 1-\big(1-P(X_1>t)\big)^n \ge 1-e^{-nP(X_1>t)} \\ &\ge 1-\frac{1}{1+nP(X_1>t)} = \frac{nP(X_1>t)}{1+nP(X_1>t)}, \end{aligned} \quad (8)$$

$$P\Big(\max_{1\le i\le n}|g_i| > (1+\varepsilon)M\Big) \ge \frac{nP(|g_1|>(1+\varepsilon)M)}{1+nP(|g_1|>(1+\varepsilon)M)}. \quad (9)$$

The right hand side is an increasing function of $P(|g_1| > (1+\varepsilon)M)$ and, by (4),

$$\begin{aligned} &P(|g_1|>(1+\varepsilon)M) \\ &\ge \sqrt{\frac{2}{\pi}}\frac{1}{(1+\varepsilon)M+1}e^{-(1+\varepsilon)^2M^2/2}\big(1-e^{-\frac{1}{2}}e^{-(1+\varepsilon)M}\big) \\ &= \sqrt{\frac{2}{\pi}}\frac{1}{(1+\varepsilon)M+1}e^{-M^2/2}e^{-\varepsilon M^2-\varepsilon^2M^2/2}\big(1-e^{-\frac{1}{2}}e^{-(1+\varepsilon)M}\big) \\ &\ge \frac{M(1-2^{-1/n})}{(1+\varepsilon)M+1}e^{-\varepsilon M^2-\varepsilon^2M^2/2}\big(1-e^{-\frac{1}{2}}e^{-(1+\varepsilon)M}\big) \\ &\ge \frac{\log 2}{4n}e^{-\varepsilon M^2-\varepsilon^2M^2/2} \ge \frac{\log 2}{4n}e^{-3\varepsilon M^2/2}, \end{aligned} \quad (10)$$

for $\varepsilon \le 1$ and n large enough (independently of ε). Using (9), we get

$$P\Big(\max_{1\le i\le n}|g_i| > (1+\varepsilon)M\Big) \ge \frac{\frac{\log 2}{4}e^{-3\varepsilon M^2/2}}{1+\frac{\log 2}{4}e^{-3\varepsilon M^2/2}} \ge \frac{\log 2}{4+\log 2}e^{-3\varepsilon M^2/2}. \quad (11)$$

□

Claim 3. For some absolute positive constants c, C and for all $0 < \varepsilon < 1/2$,

$$\exp\big(-Ce^{\varepsilon M^2}\big) \le P\Big(\max_{1\le i\le n}|g_i| < (1-\varepsilon)M\Big) \le C\exp\big(-ce^{3\varepsilon M^2/4}\big). \quad (12)$$

Proof.

$$\begin{aligned}
&P\Big(\max_{1\le i\le n}|g_i| < (1-\varepsilon)M\Big)\\
&= \Big(1-\sqrt{\frac{2}{\pi}}\int_{(1-\varepsilon)M}^{\infty} e^{-s^2/2}\Big)^n \le \Big(1-\sqrt{\frac{2}{\pi}}\int_{(1-\varepsilon)M}^{M} e^{-s^2/2}\Big)^n\\
&\le \Big(1-\sqrt{\frac{2}{\pi}}\frac{1}{M}\int_{(1-\varepsilon)M}^{M} se^{-s^2/2}\Big)^n\\
&= \Big(1-\sqrt{\frac{2}{\pi}}\frac{1}{M}\big(e^{-(1-\varepsilon)^2M^2/2}-e^{-M^2/2}\big)\Big)^n\\
&= \Big(1-\sqrt{\frac{2}{\pi}}\frac{1}{M}e^{-M^2/2}\big(e^{\varepsilon M^2-\varepsilon^2M^2/2}-1\big)\Big)^n\\
&\le \Big(1-(1-2^{-1/n})\big(e^{\varepsilon M^2-\varepsilon^2M^2/2}-1\big)\Big)^n \quad \text{by (4)}\\
&\le \exp\big(-n(1-2^{-1/n})(e^{\varepsilon M^2-\varepsilon^2M^2/2}-1)\big)\\
&\le 2\big(1+o(1)\big)\exp\big(-\log 2\big(1+o(1)\big)e^{3\varepsilon M^2/4}\big)
\end{aligned}$$

which proves the right hand side inequality in (12). As for the left hand side,

$$\begin{aligned}
&P\Big(\max_{1\le i\le n}|g_i| < (1-\varepsilon)M\Big)\\
&= \Big(1-\sqrt{\frac{2}{\pi}}\int_{(1-\varepsilon)M}^{\infty} e^{-s^2/2}\Big)^n \ge \Big(1-\sqrt{\frac{2}{\pi}}\frac{1}{(1-\varepsilon)M}\int_{(1-\varepsilon)M}^{\infty} se^{-s^2/2}\Big)^n\\
&= \Big(1-\sqrt{\frac{2}{\pi}}\frac{1}{(1-\varepsilon)M}(e^{-(1-\varepsilon)^2M^2/2}\Big)^n = \Big(1-\sqrt{\frac{2}{\pi}}\frac{e^{-M^2/2}}{(1-\varepsilon)M}e^{\varepsilon M^2-\varepsilon^2M^2/2}\Big)^n\\
&\ge \exp\Big(-\frac{2n(1-2^{-1/n})}{1-e^{-\frac{1}{2}-M}}\frac{M+1}{M}e^{\varepsilon M^2-\varepsilon^2M^2/2}\Big) \quad \text{by (3)}\\
&\ge \exp\big(-2\log 2\big(1+o(1)\big)e^{\varepsilon M^2}\big).
\end{aligned}$$

□

We summarize Claims 2 and 3 in a form that will be useful for us later in the following Proposition.

Proposition 1. *For some positive absolute constants* c, C *and for all* $0<\varepsilon<1$ *and* $n \in \mathbb{N}$*, denoting* $g = (g_1, g_2, \ldots, g_n)$*,*

$$ce^{-C\varepsilon \log n} \le P\Big(\|g\|_\infty < \frac{(1-\varepsilon)M}{\sqrt{n}}\|g\|_2 \;\; \text{or} \;\; \|g\|_\infty > \frac{(1+\varepsilon)M}{\sqrt{n}}\|g\|_2\Big)$$
$$\le Ce^{-c\varepsilon \log n}.$$

Proof. This follows easily from Claims 2 and 3 and the facts that $e^x > x$ for all x, M is of order $\sqrt{\log n}$ and

$$P\Big(\|g\|_2 < (1-\varepsilon)\sqrt{n} \;\; \text{or} \;\; \|g\|_2 > (1+\varepsilon)\sqrt{n}\Big) < Ce^{-\varepsilon^2 n}.$$

3 Proof of the Theorem

The first part of the Theorem follows easily from the, by now well exposed, proof of Milman's version of Dvoretzky's theorem (see e.g, [MS] or [Pi]) with the improved concentration estimate in (the right hand side of the inequality in) Proposition 1 replacing the classical estimates. For the proof of the second part we need:

Lemma 1. *Let* $\mathcal{A}$ *be a subset of* $G_{n,k}$ *of* $\mu_{n,k}$ *measure* a*. Put* $U_\mathcal{A} = \bigcup_{E\in\mathcal{A}} E$*, then*

$$P\big((g_1, g_2, \ldots, g_n) \in U_\mathcal{A}\big) \ge a^{1/k}.$$

Proof. Let $X_1, X_2, \ldots, X_k$ be k independent random vectors distributed according to P, the canonical Gaussian measure on $\mathbb{R}^n$. Note that, since $\mu_{n,k}$ is the unique rotational invariant probability measure on $G_{n,k}$, the distribution of $\mathrm{span}\{X_1, \ldots, X_k\}$ is $\mu_{n,k}$. Accordingly,

$$\begin{aligned} P(U_\mathcal{A})^k &= P(X_1, X_2, \ldots, X_k \in U_\mathcal{A}) \\ &\ge P\big(\mathrm{span}\{X_1, X_2, \ldots, X_k\} \in \mathcal{A}\big) \\ &= \mu_{n,k}(\mathcal{A}). \end{aligned}$$

□

Remark 1. As we'll see below we use only a weak form of Lemma 1. We actually believe there is a much stronger form of it.

Proof of the moreover part in Theorem 1. Let $\mathcal{A} \subset G_{n,k}$ be such that every $E \in \mathcal{A}$ there is an M_E such that

$$M_E\|x\|_2 \le \|x\|_\infty \le (1+\varepsilon)M_E\|x\|_2$$

for all $x \in E$. Let $\mathcal{B}$ be the subset of $\mathcal{A}$ of all E for which $\frac{(1-3\varepsilon)M}{\sqrt{n}} \le M_E \le \frac{(1+\varepsilon)M}{\sqrt{n}}$, and let $\mathcal{C} = \mathcal{A} \setminus \mathcal{B}$. By Lemma 1,

$$\mu_{n,k}(\mathcal{C})^{1/k} \le P\left(\left\{x;\ \|x\|_\infty < \frac{(1+\varepsilon)(1-3\varepsilon)M}{\sqrt{n}}\|x\|_2 \text{ or } \|x\|_\infty > \frac{(1+\varepsilon)M}{\sqrt{n}}\|x\|_2\right\}\right)$$

and, by Proposition 1, this last quantity is smaller than $Ce^{-c\varepsilon\log n}$. It follows that

$$\mu_{n,k}(\mathcal{B}) > 1 - e^{-ak} - Ce^{-c\varepsilon k\log n}.$$

We may assume that $\varepsilon\log n$ is much larger than a so that the last term above is dominated by e^{-ak}. Applying Lemma 1 once more we get

$$P\left(\left\{x;\ (\frac{(1-3\varepsilon)M}{\sqrt{n}}\|x\|_2 \le \|x\|_\infty \le \frac{(1+\varepsilon)^2M}{\sqrt{n}}\|x\|_2\right\}\right) \ge \mu_{n,k}(\mathcal{B}) > 1-2e^{-ak}.$$

Using now the other part of Proposition 1 we get that

$$C\varepsilon\log n > ak.$$

References

[Dv] Dvoretzky, A.: Some results on convex bodies and Banach spaces. Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960) pp. 123–160. Jerusalem Academic Press, Jerusalem; Pergamon, Oxford (1961)

[Go] Gordon, Y.: Some inequalities for Gaussian processes and applications. Israel J. Math., **50**, no. 4, 265–289 (1985)

[Mi] Milman, V.D.: A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies. (Russian) Funkcional. Anal. i Priložen., **5**, no. 4, 28–37 (1971)

[MS] Milman, V.D., Schechtman, G.: Asymptotic theory of finite-dimensional normed spaces, Lecture Notes in Mathematics, **1200**. Springer-Verlag, Berlin (1986)

[Pi] Pisier, G.: The volume of convex bodies and Banach space geometry. Cambridge Tracts in Mathematics, **94**. Cambridge University Press, Cambridge (1989)

[Sc] Schechtman, G.: Two observations regarding embedding subsets of Euclidean spaces in normed spaces. Advances in Math., **200**, no. 1, 125–135 (2006)