Asymptotic Geometric Analysis, Fall 2006*

Gideon Schechtman

January 22, 2007

1 Introduction

Lecture 1, Oct 30, 2006

The course will deal with convex symmetric bodies in $\mathbb{R}^n$. In the first few lectures we will formulate and prove Dvoretzky theorem, Theorem 1.2.

Definition 1.1. *A convex, symmetric (around 0) body $K \subset \mathbb{R}^n$ is a compact set with non-empty interior which is:*

- *convex: $x, y \in K$ and $\lambda \in [0,1] \implies \lambda x + (1-\lambda)y \in K$*
- *symmetric: $x \in K \implies -x \in K$.*

Examples: Consider the family of normed linear spaces ℓ_p^n over $\mathbb{R}$ which are just $\mathbb{R}^n$ equipped with the norm:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \text{ for } 1 \le p < \infty$$

and

$$\|x\|_\infty = \max_{1 \le i \le n} |x_i|.$$

Then the unit balls of ℓ_p^n:

$$B_p^n = \{x \in \mathbb{R}^n \ : \text{ such that } \|x\|_p \le 1\}$$

are examples of convex, symmetric bodies.

(Examples of $p = 1, 2, \infty$ for $n = 2, 3$)

*These notes are based on notes that Boris Levant prepared following a similar course given by me a few years ago

Theorem 1.2. *(A. Dvoretzky, 1960) For every $\epsilon > 0$ there exists a constant $c = c(\epsilon) > 0$ such that for every $n \in \mathbb{N}$ and every convex symmetric body in $K \subset \mathbb{R}^n$ there exists a subspace $V \subseteq \mathbb{R}^n$ satisfying:*

1. $\dim V = k$, *where* $k \geq c \cdot \log n$.
2. $V \cap K$ *is "ϵ-euclidean", which means that there exists $r > 0$, such that:*

$$r \cdot V \cap B_2^n \subset V \cap K \subset (1+\epsilon) r \cdot V \cap B_2^n.$$

For example the unit ball of ℓ_∞^n - the n-dimensional cube - is far from the Euclidean ball. Its easy to see, that the ratio of radii of the bounding and the bounded ball is $\sqrt{n}$:

$$B_2^n \subset B_\infty^n \subset \sqrt{n} B_2^n$$

and $\sqrt{n}$ is the best constant. Yet, according to Dvoretzky theorem, we can find a subspace of $\mathbb{R}^n$ of dimension proportional to $\log n$ in which the ratio of bounding and bounded balls will be $1 + \epsilon$.

Remark: The constant c in the formulation of Dvoretzky theorem depends on the quality of approximation - ϵ. It is known, that: $c_1 \cdot \frac{\epsilon}{(\log \frac{1}{\epsilon})^2} \leq c \leq c_2 \cdot \frac{1}{\log \frac{1}{\epsilon}}$. Clearly, there is a big gap between the upper and lower bounds and the exact dependence is an important open question.

Proposition 1.3. *For a non-empty set K denote $\|x\|_K = \inf\{\lambda > 0 \ : \ \frac{x}{\lambda} \in K\}$*

1. $\|x\|_K$ *is a norm* $\iff$ K *is a convex symmetric body.*
2. *Let K, L be two convex, symmetric bodies.* $K \subset L \iff \forall x. \ \|x\|_K \geq \|x\|_L$.

The proof of the proposition is left as an exercise. Proposition 1.3 allows us to identify the class of convex symmetric bodies in $\mathbb{R}^n$ with the class of norms on $\mathbb{R}^n$. It justifies the following equivalent formulation of Dvoretzky theorem: exercise

Theorem 1.4. *For every $\epsilon > 0$ there exist a constant $c = c(\epsilon) > 0$ such that for every $n \in \mathbb{N}$ and every norm $\|\cdot\|$ in $\mathbb{R}^n$ there exist a subspace $V \subseteq \mathbb{R}^n$ satisfying:*

1. $\dim V = k$, *where* $k \geq c \cdot \log n$.
2. *There exists* $0 < M < \infty$ *such that for every* $x \in V$:

$$M \cdot \|x\|_2 \leq \|x\| \leq (1+\epsilon) M \cdot \|x\|_2.$$

In other words, the norms $\|\cdot\|_2$ and $\|\cdot\|$ are equivalent on V up to ϵ.

Very vague sketch of the proof: Consider the unit sphere of ℓ_2^n, the surface of B_2^n, which we will denote by $S^{n-1} = \{x \in \mathbb{R}^n \ : \ \|x\|_2 = 1\}$. Let $\|x\|$ be some arbitrary norm in $\mathbb{R}^n$. The first task will be to show that there exists a "large" set $S_{\text{good}} \subset S^{n-1}$ satisfying $\forall x \in S_{\text{good}}.\ |\|x\| - M| < \epsilon M$ where M is the average of $\|x\|$ on S^{n-1}. Moreover, we shall see that, dependeing on the Lipschitz constant of $\|\cdot\|$, the set S_{good} is "almost all" the sphere in the measure sense. This phenomenon is called *concentration of measure*.
The next stage will be to pass from the "large" set to a large dimensional subspace of $\mathbb{R}^n$ contained in it. Denote $O(n)$ - the group of orthogonal transformations from $\mathbb{R}^n$ into itself. Choose some subspace V_0 of appropriate dimension k and fix an ϵ-net N on $V_0 \cap S^{n-1}$. For some $x_0 \in N$, "almost all" transformations $U \in O(n)$ will send it into some point in S_{good}. Moreover, if the "almost all" notion is good enough, we will be able to find a transformation that sends all the points of the ϵ-net into S_{good}. Now there is a standard approximation procedure that will let us pass from the ϵ-net to all points in the subspace.

2 Concentration of Measure

Denote by μ the normalized Haar measure on S^{n-1} - the unique, probability measure which is invariant under rotations. In other words, for all $U \in O(n)$, the group of orthogonal transformations, and every measurable set $A \subseteq S^{n-1}$: $\mu(A) = \mu(UA)$. There are many equivalent ways to define this measure. Here is one: $\mu(A)$ is the n dimensional Lebesgue measure of the cone defined by A intersected with the ball, B_2^n, and normalized (by divided by the measure of the whole ball). The uniqueness of this measure will be important for us. The exact definition and proof of uniqueness can be found in [MS].

Theorem 2.1. *(P. Levy) Let, $f : S^{n-1} \longrightarrow \mathbb{R}$ be a Lipshitz function with a constant L:*

$$\forall x, y \in S^{n-1}.\ |f(x) - f(y)| \leq L \cdot \|x - y\|_2.$$

Then,

$$\mu\{x \in S^{n-1} \ : \ |f(x) - Ef| > \epsilon\} \leq Ce^{-\frac{\epsilon^2 n}{16L^2}}$$

for some specific absolute constant $0 < C < \infty$.

Remark: The theorem also holds with the expectation of f replaced by its median.

Theorem 2.2. *(Brunn–Minkowski inequality) Let, A, B be two measurable non-empty sets in $\mathbb{R}^n$. Then:*

$$Vol(A + B)^{1/n} \geq Vol(A)^{1/n} + Vol(B)^{1/n}$$

We can reformulate the inequality in an equivalent multiplicative form:

Theorem 2.3. *(Brunn–Minkowski inequality, multiplicative form) Let, A, B be two measurable non-empty sets in $\mathbb{R}^n$, $0 < \lambda < 1$. Then:*

$$Vol(\lambda A + (1-\lambda)B) \geq Vol(A)^{\lambda} \cdot Vol(B)^{1-\lambda}.$$

Claim 2.4. *The two inequalities in theorems 2.2 and 2.3 are equivalent.*

Proof. 2.2⇒ 2.3)

$$\begin{aligned} Vol(\lambda A + (1-\lambda)B)^{1/n} &\geq Vol(\lambda A)^{1/n} + Vol((1-\lambda)B)^{1/n} \\ &= \lambda Vol(A)^{1/n} + (1-\lambda)Vol(B)^{1/n} \\ &\geq Vol(A)^{\lambda/n} \cdot Vol(B)^{(1-\lambda)/n}. \end{aligned}$$

Where the last inequality follows from the inequality of arithmetic and geometric mean or equivalently by the concavity of the log function.

2.3⇒ 2.2) Assume, that $Vol(A), Vol(B) > 0$.

$$\begin{aligned} \frac{Vol(A+B)}{(Vol(A)^{1/n} + Vol(B)^{1/n})^n} &= Vol\left(\frac{A+B}{Vol(A)^{1/n} + Vol(B)^{1/n}}\right) = \\ &= Vol\left(\lambda\frac{A}{Vol(A)^{1/n}} + (1-\lambda)\frac{B}{Vol(B)^{1/n}}\right). \end{aligned}$$

Where, $\lambda = \frac{Vol(A)^{1/n}}{Vol(A)^{1/n}+Vol(B)^{1/n}}$. Applying theorem 2.3 we get:

$$\frac{Vol(A+B)}{(Vol(A)^{1/n} + Vol(B)^{1/n})^n} \geq Vol\left(\frac{A}{Vol(A)^{1/n}}\right)^{\lambda} \cdot Vol\left(\frac{B}{Vol(B)^{1/n}}\right)^{1-\lambda} = 1.$$

□

Remark: For convex bodies (with non-empty interiors), equality holds in (each of the two versions of) the Brunn–Minkowski inequality if and only if the two bodies are homothetic.

Before proving the Brunn–Minkowski inequality we prove a corollary - The classical isoperimetric inequality. We first need to define the measure of the boundary of a body in $\mathbb{R}^n$.

Definition 2.5. *Let $A \subset \mathbb{R}^n$ with a smooth boundary. Then the volume of the boundary of A is defined as:*

$$Vol(\partial A) = \lim_{t\to 0} \frac{Vol(A + tB_2^n) - Vol(A)}{t},$$

given that the limit exists.

Proposition 2.6. *Denote by $v_n = Vol(B_2^n)$ - the volume of the n-dimensional Euclidean ball. Then B_2^n has the smallest volume of the boundary among all bodies of volume v_n in $\mathbb{R}^n$.*

Proof. Let $A \subset \mathbb{R}^n$ be some body, which has a finite volume of the boundary defined in definition 2.5 and $Vol(A) = v_n$. Define the following function: $f(t) = Vol(A+tB_2^n)$. According to Brunn-Minkowski inequality, the derivative of $f(t)^{1/n}$ at $t = 0$ satisfies:

$$(f(t)^{1/n})'_{t=0} = \lim_{t\to 0} \frac{Vol(A + tB_2^n)^{1/n} - Vol(A)^{1/n}}{t} \geq \lim_{t\to 0} \frac{Vol(tB_2^n)^{1/n}}{t} = Vol(B_2^n)^{1/n}$$

with equality for $A = B_2^n$. On the other hand:

$$(f(t)^{1/n})'_{t=0} = \frac{1}{n} \cdot f(0)^{1/n-1} \cdot f'(0) = \frac{1}{n} \cdot Vol(A)^{1/n-1} \cdot Vol(\partial A).$$

Combining these equalities we get the result:

$$Vol(\partial A) \geq n \cdot Vol(B_2^n)^{1/n} \cdot Vol(A)^{1-1/n} = n \cdot v_n = Vol(\partial B_2^n).$$

□

Instead of directly proving the Brunn–Minkowski inequality, we will formulate and prove a generalization - the Prekopa-Leindler inequality: Lecture 2, Nov 6, 2006

Theorem 2.7. *(Prekopa–Leindler) Let $f, g, m : \mathbb{R}^n \longrightarrow [0, \infty)$ be three measurable non-negative functions satisfying for some $0 < \lambda < 1$:*

$$\forall x, y \in \mathbb{R}^n.\ m(\lambda x + (1 - \lambda)y) \geq f(x)^\lambda \cdot g(y)^{1-\lambda}. \tag{2.0.1}$$

Then:

$$\int_{\mathbb{R}^n} m \geq \left(\int_{\mathbb{R}^n} f\right)^\lambda \cdot \left(\int_{\mathbb{R}^n} g\right)^{1-\lambda}$$

Remark: Before proving the theorem, we would like to note, that the Brunn–Minkowski inequality immediately follows from the Prekopa–Leindler inequality. To see this, let the functions f, g be the indicator functions of the sets A and B respectively, and let the function m be the indicator of the set $\lambda A + (1 - \lambda)B$. These functions satisfy condition (2.0.1) for every $0 < \lambda < 1$ (check). Hence, according to theorem 2.7 check

$$Vol(\lambda A + (1 - \lambda)B) = \int_{\mathbb{R}^n} m \geq \left(\int_{\mathbb{R}^n} f\right)^\lambda \cdot \left(\int_{\mathbb{R}^n} g\right)^{1-\lambda} = Vol(A)^\lambda \cdot Vol(B)^{1-\lambda}.$$

Exercise: Show that the Prekopa–Leindler inequality follows from the Brunn–Minkowski inequality. (The one dimensional case will be proved below.) exercise

Proof. The proof will be by induction on the dimension n. We delay checking the $n = 1$ case.
Assume, the theorem is true for $\mathbb{R}^n$. Consider three non-negative functions $f, g, m : \mathbb{R}^{n+1} \to [0, \infty)$, satisfying (2.0.1) for some $0 < \lambda < 1$. Fix $x_0, y_0 \in \mathbb{R}$ and denote $z_0 = \lambda x_0 + (1 - \lambda) y_0$. Define new functions $f_{x_0}, g_{y_0}, m_{z_0} : \mathbb{R}^n \to [0, \infty)$ by fixing the first coordinate in the respective original functions: $f_{x_0}(x) = f(x_0, x)$, $g_{y_0}(y) = g(y_0, y)$, $m_{z_0}(z) = m(z_0, z)$. The new functions also satisfy condition (2.0.1) of the Prekopa–Leindler inequality (check), hence by the induction hypothesis: check

$$\int_{\mathbb{R}^n} m_{z_0}(x)dx \geq \left(\int_{\mathbb{R}^n} f_{x_0}(x)dx \right)^{\lambda} \cdot \left(\int_{\mathbb{R}^n} g_{y_0}(x)dx \right)^{1-\lambda}. \tag{2.0.2}$$

Note, that the last inequality holds true for every x_0, y_0, z_0 satisfying the relationship $z_0 = \lambda x_0 + (1 - \lambda) y_0$. Now again define three new functions $\widetilde{f}, \widetilde{g}, \widetilde{m} : \mathbb{R} \to [0, \infty)$:

$$\widetilde{f}(u) = \int_{\mathbb{R}^n} f_u(x)dx$$
$$\widetilde{g}(u) = \int_{\mathbb{R}^n} g_u(x)dx$$
$$\widetilde{m}(u) = \int_{\mathbb{R}^n} m_u(x)dx$$

According to (2.0.2), the functions $\widetilde{f}, \widetilde{g}, \widetilde{m}$ satisfy condition (2.0.1) of the Prekopa–Leindler inequality in the one-dimensional case. Applying the conclusion of theorem 2.7 to those functions we get the desired result:

$$\int_{\mathbb{R}^{n+1}} m = \int_{\mathbb{R}} \widetilde{m} \geq \left(\int_{\mathbb{R}} \widetilde{f} \right)^{\lambda} \cdot \left(\int_{\mathbb{R}} \widetilde{g} \right)^{1-\lambda} = \left(\int_{\mathbb{R}^{n+1}} f \right)^{\lambda} \cdot \left(\int_{\mathbb{R}^{n+1}} g \right)^{1-\lambda}.$$

This concludes the induction step. In order to complete the proof we need to prove the theorem in the one-dimensional case. First, we will show the Brunn–Minkowski inequality in $\mathbb{R}$. The volume is invariant under translations along $\mathbb{R}$, a simple approximation procedure that we'll not reproduce in these notes (but was shown in class) shows that it's enough to consider $A \subseteq (-\infty, +\epsilon]$ and $B \subseteq [-\epsilon, +\infty)$. Moreover, assume that $0 \in A \cap B$. Then, clearly, $A \cup B \subseteq A + B$. Hence:

$$Vol(A+B) \geq Vol(A \cup B) = Vol(A) + Vol(B) - Vol(A \cap B) \geq Vol(A) + Vol(B) - 2\epsilon$$

Letting ϵ tend to 0 proves Theorem 2.2 in the one-dimensional case.

Now, let $f, g, m : \mathbb{R} \longrightarrow [0, \infty)$ satisfy condition (2.0.1). Note, that by Fubini's theorem, we can represent the integral of the positive real function:

$$\int_{\mathbb{R}} f = \int_0^{\infty} |\{x \ : \ f(x) \geq t\}| dt \tag{2.0.3}$$

(check). We may assume by simple approximation and normalization argument, that $\|f\|_\infty = \|g\|_\infty = 1$, we get that the sets $\{x : f(x) \geq t\}$ and $\{y : g(y) \geq t\}$ are non-empty for every $0 \leq t < 1$. Let $x, y \in \mathbb{R}$ and $t \in [0,1)$ such that $f(x) \geq t$ and $g(y) \geq t$. Then $\forall \lambda \in [0,1]$. $m(\lambda x + (1-\lambda)y) \geq t$. Hence: check

$$\{z : m(z) \geq t\} \supseteq \lambda\{x : f(x) \geq t\} + (1-\lambda)\{y : g(y) \geq t\}.$$

Applying the one dimensional Brunn–Minkowski inequality we get:

$$\begin{aligned} |\{z : m(z) \geq t\}| &\geq |\lambda\{x : f(x) \geq t\} + (1-\lambda)\{y : g(y) \geq t\}| \geq \\ &\geq \lambda|\{x : f(x) \geq t\}| + (1-\lambda)|\{y : g(y) \geq t\}|. \end{aligned}$$

Using the (2.0.3) we get

$$\begin{aligned} \int_{\mathbb{R}} m &\geq \int_0^1 |\{z : m(z) \geq t\}| dt \\ &\geq \lambda \int_0^1 |\{x : f(x) \geq t\}| dt + (1-\lambda) \int_0^1 |\{y : g(y) \geq t\}| dt \\ &= \lambda \int_{\mathbb{R}} f + (1-\lambda) \int_{\mathbb{R}} g \\ &\geq \left(\int_{\mathbb{R}} f\right)^{\lambda} \cdot \left(\int_{\mathbb{R}} g\right)^{1-\lambda} \end{aligned}$$

by the arithmetic geometric inequality. □

Brunn's inequality This is another application of the Brunn–Minkowski inequality: Consider K - a convex body in $\mathbb{R}^{n+1}$ and let u be a unit vector in $\mathbb{R}^{n+1}$. For each x let H_x be the hyperplane

$$\{y \in \mathbb{R}^{n+1} : (y, u) = x\},$$

where $(\cdot, \cdot)$ denotes the usual inner product in $\mathbb{R}^{n+1}$. Then the original inequality proved by Brunn stateds that the function

$$f(x) = Vol(K \cap H_x)^{1/n}$$

is concave on its support. To illustrate that statement pick $u = (1, 0, \ldots, 0)$. Then the slice of K perpendicular to u at some point x will be

$$K \cap H_x = K_x = \{y = (y_1, \ldots, y_n) : (x, y_1, \ldots, y_n) \in K\}.$$

Hence, according to Brunn's inequality, the function $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = Vol(K_x)^{1/n}$ is concave on the set $S = \{x : Vol(K_x) > 0\}$.
Note, that Brunn's inequality implies that if for each x we replace K_x with the Euclidean ball B_x of an appropriate radius, such that $Vol(K_x) = Vol(B_x)$, then the body we obtain in such a process is again convex (check!). This procedure is called Steiner check

symmetrization and is a tool in proving several inequalities in convex geometry.

The next step towards the proof of Levy's concentration inequality is to formulate and prove an *approximate isoperimetric inequality* for the sphere S^{n-1} equipped with the usual Euclidean metric $d(x, y) = \|x - y\|_2$. (The same result holds for the geodesic metric.) For a set A in S^{n-1} and $\epsilon > 0$ we denote $A_\epsilon = \{x \in S^{n-1} \ : \ d(x, A) \le \epsilon\}$.

Theorem 2.8. *There are constants $0 < c < C < \infty$ such that for all n and all measurable $A \subseteq S^{n-1}$.*

$$\mu((A_\epsilon)^c) \le \frac{C}{\mu(A)} \cdot e^{-c\epsilon^2 n}. \tag{2.0.4}$$

(μ denotes the unique rotational invariant probability measure on S^{n-1}.)

Proof. (Due to Arias-de-Reyna, Ball, and Villa.) We will first prove a similar statement not on the sphere, but for subsets of the unit ball B_2^n. Let A, B be two sets in B_2^n, such that the distance between them is positive: $d(A, B) = \sup_{x\in A,\, y\in B} d(x, y) \ge \epsilon$. For every $x \in A$, $y \in B$ we have:

$$\left\|\frac{x+y}{2}\right\|_2^2 + \frac{\epsilon^2}{4} \le \left\|\frac{x+y}{2}\right\|_2^2 + \left\|\frac{x-y}{2}\right\|_2^2 = \frac{\|x\|_2^2 + \|y\|_2^2}{2} \le 1,$$

where the second equality is just the parallelogram equality. Hence, we get:

$$\begin{aligned} &\left\|\frac{x+y}{2}\right\|_2 \le \left(1 - \frac{\epsilon^2}{4}\right)^{1/2} \\ \Longrightarrow\ &\frac{A+B}{2} \subseteq \left(1 - \frac{\epsilon^2}{4}\right)^{1/2} B_2^n \\ \Longrightarrow\ &Vol\left(\frac{A+B}{2}\right) \le \left(1 - \frac{\epsilon^2}{4}\right)^{n/2} Vol(B_2^n) \le e^{-\epsilon^2 n/8} \cdot Vol(B_2^n). \end{aligned}$$

On the other hand the Brunn-Minkowski inequality gives:

$$Vol\left(\frac{A+B}{2}\right) \ge Vol(A)^{1/2} \cdot Vol(B)^{1/2}.$$

Combining the last two inequalities we get:

$$\frac{Vol(B)}{Vol(B_2^n)} \le \frac{Vol(B_2^n)}{Vol(A)} \cdot e^{-\epsilon^2 n/4}. \tag{2.0.5}$$

If we take $B = (A_\epsilon)^c$, we get an inequality, for the normalized Lebesgue measure on the ball, analogues to the one we seek for the sphere. It is not hard to pass from one to the other: Let A be a set in S^{n-1} and denote $B = (A_\epsilon)^c$. Define the sets lecture 3, Nov 13, 2006

$\widetilde{A} = [\frac{1}{2}, 1] \times A$ and i$\widetilde{B} = [\frac{1}{2}, 1] \times B$. Easily, $d(\widetilde{A}, \widetilde{B}) \geq \frac{\epsilon}{2}$. The measure of A can be defined it terms of the volume of the cone:

$$\mu(A) = \frac{Vol([0,1] \times A)}{Vol(B_2^n)}.$$

Trivially we have: $Vol([0,1] \times A) = 2^n \cdot Vol([0, \frac{1}{2}] \times A)$. Hence:

$$\frac{Vol(\widetilde{A})}{Vol(B_2^n)} = \frac{Vol([\frac{1}{2}, 1] \times A)}{Vol(B_2^n)} = \left(1 - \frac{1}{2^n}\right) \frac{Vol([0,1] \times A)}{Vol(B_2^n)} = \left(1 - \frac{1}{2^n}\right) \cdot \mu(A).$$

Applying the inequality (2.0.5) to the last result we get:

$$\left(1 - \frac{1}{2^n}\right) \cdot \mu(B) = \frac{Vol(\widetilde{B})}{Vol(B_2^n)} \leq \frac{Vol(B_2^n)}{Vol(\widetilde{A})} \cdot e^{-\epsilon^2 n/16} = \frac{1}{\left(1 - \frac{1}{2^n}\right) \cdot \mu(A)} \cdot e^{-\epsilon^2 n/16}.$$

Thus setting $C = 4 \geq = \left(1 - \frac{1}{2^n}\right)^{-2}$ and $c = \frac{1}{16}$ we get the result. (The constants are not the best possible.) □

Remark: There is actually an isoperimetric inequality which implies Theorem 2.8: For all $\epsilon > 0$, among all measurable subsets of S^{n-1} of a given measure, a cap of that measure is the one for which $\mu(A_\epsilon)$ is minimal.

The next theorem is equivalent to Theorem 2.8 although we will only show one direction. The theorem states, that Lipschitz functions on the sphere are "almost constants" on "almost all" the sphere.

Theorem 2.9. *(P. Levy) There exist constants $0 < c, C < \infty$ such that if $f : S^{n-1} \longrightarrow \mathbb{R}$ is a Lipschitz function with constant L, i.e:*

$$\forall x, y \in S^{n-1} \;\; |f(x) - f(y)| \leq L\|x - y\|_2,$$

and M is the median of f, namely:

$$\mu(\{x \;:\; f(x) \leq M\}) \geq \frac{1}{2} \text{ and } \mu(\{x \;:\; f(x) \geq M\}) \geq \frac{1}{2}.$$

Then, for all $t \in \mathbb{R}$:

$$\mu(\{x \;:\; |f(x) - M| > t\}) \leq C \cdot e^{-\frac{ct^2 n}{L^2}}.$$

Proof. Assume first that $L = 1$. Denote $A = \{x \;:\; f(x) \leq M\}$. Clearly:

$$\{x \;:\; f(x) > M + t\} \subseteq A_t^c,$$

which follows from the fact that $f(x)$ is Lipschitz with constant 1. Hence, according to Theorem 2.8:

$$\mu(\{f(x)-M>t\}) \le \mu(A_t^c) \le \frac{4}{\mu(A)} e^{-t^2 n/16} \le 8e^{-t^2 n/16}.$$

Similarly, check

$$\mu(\{f(x)-M<-t\}) \le \mu(A_t^c) \le 8e^{-t^2 n/16}.$$

Combining the two last inequalities yields the result. For a general $L \neq 1$, we have:

$$\mu(\{|f(x)-M|>t\}) = \mu\left(\left\{\left|\frac{f(x)}{L}-\frac{M}{L}\right| > \frac{t}{L}\right\}\right) \le 16e^{-t^2 n/16L^2}.$$

□

Exercise: Deduce Theorem 2.8 from Theorem 2.9. exercise

Remark: It is easy to deduce from the proofs of Theorems 2.8 and 2.9 versions of these theorems which are valid not only for the Euclidean unit sphere, but for unit spheres of other norms, which are sufficiently *uniformly convex*. Let $K \subseteq \mathbb{R}^n$ be convex, centrally symmetric body, and let $\|\cdot\|_K$ be its induced norm. K is called *uniformly convex* with modulus $\delta(\cdot) > 0$, if for all $\epsilon > 0$ and all $x, y \in K$,

$$\|x-y\|_K \ge \epsilon \implies \|x+y\|_K \le 2(1-\delta(\epsilon)).$$

In the versions alluded to above $\delta(\epsilon)$ replaces ϵ^2.

Theorem 2.9 remains valid if we change the median of $f(x)$ to its expectation $\mathbb{E}f = \int f d\mu$:

Corollary 2.10. *For some absolute constants* $0 < c, C < \infty$, *if* $f : S^{n-1} \longrightarrow \mathbb{R}$ *is a Lipschitz function with constant* L, *then:*

$$\mu(\{x \,:\, |f(x)-\mathbb{E}f| > t\}) \le C \cdot e^{-\frac{ct^2 n}{L^2}}.$$

Proof. As before, we can assume that $L = 1$. Denote by $\mu \times \mu$ - the Haar measure on the product space $S^{n-1} \times S^{n-1}$. Then:

$$\begin{aligned} &\mu \times \mu\{x, \overline{x} \in S^{n-1} \,:\, |f(x)-f(\overline{x})| > t\} \le \\ &\le \mu \times \mu\{x, \overline{x} \in S^{n-1} \,:\, |f(x)-M| + |f(\overline{x})-M| > t\} \le \\ &\le 2\mu\{x \in S^{n-1} \,:\, |f(x)-M| > \frac{t}{2}\} \le 32 \cdot e^{-t^2 n/64}. \end{aligned}$$

Hence, using Fubini's theorem we can estimate the expectation of $e^{\lambda^2 |f(x)-\mathbb{E}f|^2}$: check

$$\mathbb{E}_x\mathbb{E}_{\overline{x}}e^{\lambda^2|f(x)-f(\overline{x})|^2} = \int_0^\infty 2\lambda^2 t e^{\lambda^2 t^2}\mu\times\mu\{|f(x)-f(\overline{x})|>t\}dt \le 64\lambda^2\int_0^\infty te^{\lambda^2t^2-\frac{t^2n}{64}}dt.$$

Letting $\lambda = \sqrt{\frac{n}{128}}$ we get by simple calculations

$$\mathbb{E}_x\mathbb{E}_{\overline{x}}e^{\frac{n}{128}|f(x)-f(\overline{x})|^2} \le 32.$$

Noting, that the function e^{t^2} is convex, we can apply Jensen inequality:

$$\mathbb{E}e^{\lambda|f(x)-\mathbb{E}f|^2} \le \mathbb{E}_x\mathbb{E}_{\overline{x}}e^{\lambda|f(x)-f(\overline{x})|^2} \le 32.$$

Finally, Chebyshev's inequality yields:

$$\mu\{|f(x)-\mathbb{E}f|>t\} = \mu\{e^{\frac{n}{128}|f(x)-\mathbb{E}f|^2} > e^{\frac{n}{128}t^2}\} = \mu\{e^{\frac{n}{128}|f(x)-\mathbb{E}f|-\frac{n}{128}t^2} > 1\} \le$$
$$\le \mathbb{E}e^{\frac{n}{128}|f(x)-\mathbb{E}f|-\frac{n}{128}t^2} \le 32e^{-\frac{n}{128}t^2}.$$

□

Exercise: Deduce Theorem 2.9 from Corollary 2.10. exercise

Versions of Levy's concentration inequality are known to hold for many natural metric probability spaces of "high dimension". Here are two discrete examples:

Example: Consider a space $\Omega = \{0,1\}^n$, where the measure of the set is just the normalized number of elements: $\forall A \subseteq \Omega,\ \mu(A) = \frac{\#A}{2^n}$. The metric we will consider is the normalized Hamming distance:

$$\forall x, y \in \Omega. \quad d(x,y) = \frac{1}{n}\cdot\#\{i\ :\ x_i \neq y_i\} = \frac{1}{n}\sum_{i=1}^n |x_i - y_i|.$$

In this case we know the exact isoperimetric inequality, namely for every $\epsilon > 0$ and all the sets A of a given volume a, the minimal volume of A_ϵ is attained for a ball of a suitable radius.

Example: As a second example, consider the space $\Omega = \Pi_n$ - the group of permutations on n elements. Again, the measure of a set is just the normalized number of elements in it: $\forall A \subseteq \Omega,\ \mu(A) = \frac{\#A}{n!}$. The metric will be again the normalized Hamming distance:

$$\forall \pi, \rho \in \Omega. \quad d(\pi,\rho) = \frac{1}{n}\cdot\#\{i\ :\ \pi(i) \neq \rho(i)\}.$$

In this case we do not know the exact solution for the isoperimetric problem. It is known that it is not always a ball.

In both examples a version of Theorems 2.8 and 2.9 hold. For example for all $f : S^{n-1} \longrightarrow \mathbb{R}$ Lipschitz with constant L,

$$\mu(\{x\ :\ |f(x)-\mathbb{E}f|>t\}) \le Ce^{-\frac{ct^2n}{L^2}}$$

for some absolute $0 < c, C < \infty$.

3 Proof of Dvoretzky's Theorem

Our next goal is to prove: lecture 4 Nov 20, 2006

Theorem 3.1. *(V. Milman) For every $\epsilon > 0$ there exists a constant $c = c(\epsilon) > 0$ such that for every $n \in \mathbb{N}$ and every norm $\|\cdot\|$ in $\mathbb{R}^n$ there exists a subspace $V \subseteq \mathbb{R}^n$ satisfying:*

1. $\dim V = k$, *where* $k \geq c \cdot \left(\frac{E}{b}\right)^2 n$.
2. *For every $x \in V$:*

$$(1-\epsilon)E \cdot \|x\|_2 \leq \|x\| \leq (1+\epsilon)E \cdot \|x\|_2.$$

Here $E = \int_S^{n-1} \|x\| dx$ and b is the smallest constant satisfying $\|x\| \leq b\|x\|_2$.

The idea of the proof is to apply Corollary 2.10 to the function $f(x) = \|x\|$. Note, that in that case, b will be the Lipschitz constant of f. In addition we will need two lemmas:

Definition 3.2. *ϵ-net on S^{n-1} is a set $N \subset S^{n-1}$ such that for every $y \in S^{n-1}$ there exists $x \in N$ satisfying $\|x - y\|_2 \leq \epsilon$.*

Lemma 3.3. *For every $0 < \epsilon < 1$ there exists an ϵ-net N on S^{n-1} of cardinality $\leq \left(1 + \frac{2}{\epsilon}\right)^n$.*

Proof. Denote by B_2 the unit ball of $\mathbb{R}^n$. Let $N = \{x_i\}_{i=1}^m$ be a maximal set on S^{n-1} such that for all $x, y \in N$ $\|x-y\|_2 \geq \epsilon$. The maximality of N implies that it is an ϵ-net on S^{n-1}. Consider $\{B(x_i, \frac{\epsilon}{2})\}_{i=1}^m$ - a collection of balls of radius $\frac{\epsilon}{2}$ around each x_i. They are mutually disjoint and completely contained in $\left(1 + \frac{\epsilon}{2}\right)B_2$. Hence:

$$mVol\left(B(x_1, \frac{\epsilon}{2})\right) = \sum Vol\left(B(x_i, \frac{\epsilon}{2})\right) = Vol\left(\bigcup B(x_i, \frac{\epsilon}{2})\right) \leq Vol\left(\left(1 + \frac{\epsilon}{2}\right)B_2\right).$$

We get $m \leq \left(\frac{1+\epsilon/2}{\epsilon/2}\right)^n = \left(1 + \frac{2}{\epsilon}\right)^n$. □

Exercise: Show that the same lemma holds for the sphere of the unit ball of an arbitrary norm on $\mathbb{R}^n$ (where the distance is with respect to the same norm). exercise

Exercise: Prove that cardinality of every ϵ-net on S^{n-1} is $\geq \left(\frac{1}{\epsilon}\right)^{n-1}$. exercise

Lemma 3.4. *Suppose, V is a finite dimensional Banach space equipped with norms $|\cdot|$ and $\|\cdot\|$ satisfying:*

$$(1-\epsilon) \le \|x\| \le (1+\epsilon)$$

for every $x \in N$, where N is an δ-net on $S_{|\cdot|} = \{x \ : \ |x| = 1\}$ for some $0 < \delta < 1$. Then for every $x \in V$

$$\frac{1-\epsilon-2\delta}{1-\delta}|x| \le \|x\| \le \frac{1+\epsilon}{1-\delta}|x|.$$

Proof. Let $x \in S_{|\cdot|}$ and take $x_1 \in N$ such that $|x - x_1| \le \delta$. In the next step we choose $x_2 \in N$ such that $\left|\frac{x-x_1}{|x-x_1|} - x_2\right| \le \delta$ and we get $|x - x_1 - |x - x_1|x_2| \le \delta|x - x_1| \le \delta^2$. By repeating this approximation procedure we can choose an infinite sequence $\{x_i\} \subseteq N$ and real numbers $0 < \delta_i \le \delta^i$ such that for every n:

$$\left|x - \sum_{i=1}^{n} \delta_i x_i\right| \le \delta^n.$$

Hence, $\sum_{i=1}^{n} \delta_i x_i$ converges to x both in $(V, |\cdot|)$ and in $(V, \|\cdot\|)$ (because in the finite-dimensional Banach space every two norms are equivalent). Moreover:

$$\|x\| \le \sup_n \left\|\sum_{i=1}^{n} \delta_i x_i\right\| \le \sum_{i=1}^{\infty} \delta_i \|x_i\| \le (1+\epsilon)\sum_{i=1}^{\infty} \delta_i \le \frac{1+\epsilon}{1-\delta}.$$

For the lower bound, let $x \in B_{|\cdot|}$ and choose $y \in N$ such that $|x - y| \le \delta$. Then

$$\|x\| = \|y + x - y\| \ge \|y\| - \|x - y\| \ge 1 - \epsilon - \frac{1+\epsilon}{1-\delta}|x - y| \ge \frac{1-\epsilon-2\delta}{1-\delta}.$$

□

Remark: Consider $O(n)$ - the set of all $n \times n$ orthogonal matrices over $\mathbb{R}$. It is a multiplicative group, which is compact in the metrics of $\mathbb{R}^{n^2}$. According to Haar theorem there exists a unique probability (normalized) measure ν which is invariant under multiplication by the member of the group: $\nu(A) = \nu(gA)$ for all $A \subseteq O(n)$ and all $g \in O(n)$. Consider μ - the Haar probability measure on S^{n-1}. Then for all $A \subseteq S^{n-1}$ and some point $x \in S^{n-1}$:

$$\mu(A) = \nu(\{U \in O(n) \ : \ Ux \in A\}).$$

We will stick to this notation in the sequel.

Proof. (of Theorem 3.1) Let $\epsilon < 1/5$ and let $V_0 \subseteq \mathbb{R}^n$ be a subspace of dimension k. Denote $S^{k-1} = S^{n-1} \bigcap V_0$ and consider N - an ϵ-net on S^{k-1} of cardinality $\le \left(\frac{3}{\epsilon}\right)^k$. Fix $x_0 \in N$ and consider two functions:

$$f : S^{n-1} \longrightarrow \mathbb{R}, \quad f(x) = \|x\|,$$

and

$$F : O(n) \longrightarrow \mathbb{R}, \quad F(U) = f(Ux_0) = \|Ux_0\|.$$

First note that b is the Lipschitz constant of f:

$$|f(x) - f(y)| = |\|x\| - \|y\|| \leq \|x - y\| \leq b \cdot \|x - y\|_2.$$

Considering the remark before the proof:

$$\nu\{U \in O(n) \ : \ |F(U) - \mathbb{E}_\nu F| > t\} = \mu\{x \in S^{n-1} \ : \ |f(x) - \mathbb{E}_\mu f| > t\}.$$

The same remark implies:

$$E = \mathbb{E}_\mu f = \int_S^{n-1} f(x) d\mu(x) = \int_{O(n)} F(U) d\nu(U) = \mathbb{E}_\nu F.$$

Now we are ready to apply Theorem 2.9:

$$\begin{aligned}
& \nu\{U \in O(n) \ : \ |\|Ux_0\| - E| > t\} \leq Ce^{-ct^2 n/b^2} \\
\text{let } t = \epsilon E \Rightarrow \quad & \nu\{U \in O(n) \ : \ |\|Ux_0\| - E| > \epsilon E\} \leq Ce^{-cn(\frac{\epsilon E}{b})^2} \\
\Rightarrow \quad & \nu\{U \in O(n) \ : \ |\|Ux\| - E| > \epsilon E \text{ for some } x \in N\} \leq Ce^{-n(\frac{\epsilon E}{b})^2} \cdot |N| \\
\Rightarrow \quad & \nu\{U \in O(n) \ : \ |\|Ux\| - E| > \epsilon E \text{ for some } x \in N\} \leq Ce^{-cn(\frac{\epsilon E}{b})^2 + k \ln \frac{3}{\epsilon}}.
\end{aligned}$$

Choosing $k < c' \cdot \frac{\epsilon^2}{\ln \frac{3}{\epsilon}} (\frac{E}{b})^2 n$ we assure, that there exists (with some positive probability) $U \in O(n)$ satisfying:

$$\forall x \in N \quad (1 - \epsilon)E \leq \|Ux\| \leq (1 + \epsilon)E.$$

Hence, according to Lemma 3.4:

$$\forall x \in V_0 \quad (1 - 5\epsilon)E\|x\|_2 \leq \|Ux\| \leq (1 + 5\epsilon)E\|x\|_2.$$

The subspace we are looking for is $V = UV_0$. □

Remark: We got $c(\epsilon) \sim \epsilon^2 / \log 1/\epsilon$ this can be improved to $c((\epsilon) \sim \epsilon^2$ and this is best possible.

Let us see what Theorem 3.1 gives for the spaces $\ell_p^n, 1 \leq p < \infty$. For $p > 2$ we have $\|x\|_p \leq \|x\|_2$. For $1 \leq p < 2$:

$$\|x\|_p^p = \sum_{i=1}^n |x|^p = \sum_{i=1}^n 1 \cdot |x|^p \leq (\sum_{i=1}^n 1^{\frac{2}{2-p}})^{\frac{2-p}{2}} \cdot (\sum_{i=1}^n |x|^2)^{\frac{p}{2}} = n^{\frac{2-p}{2}} \cdot \|x\|_2^p.$$

Hence, for $1 \leq p < 2$ we have $\|x\|_p \leq n^{\frac{1}{p} - \frac{1}{2}} \cdot \|x\|_2$. Now we have to evaluate E. Let $x = (g_1, \ldots, g_n) \in \mathbb{R}^n$ be a vector of independent gaussian variables. Denote $\overline{x} = \frac{1}{\|x\|_2} x$. Then:

justify the two equality signs!

$$\mathbb{E}_\mu\|\overline{x}\|_p = \mathbb{E}\frac{(\sum g_i^p)^{1/p}}{(\sum g_i^2)^{1/2}} = \frac{\mathbb{E}(\sum g_i^p)^{1/p}}{\mathbb{E}(\sum g_i^2)^{1/2}}.$$

To bound the last quantity from below we will use the following inequality:

$$\sqrt{2/\pi} \cdot n^{1/r} = (\sum (\mathbb{E}|g_i|)^r)^{1/r} \le \mathbb{E}(\sum g_i^r)^{1/r} \le (\mathbb{E}\sum g_i^r)^{1/r} = c_r \cdot n^{1/r}$$

Hence:

$$E = \mathbb{E}_\mu\|\overline{x}\|_p \ge c_p \cdot n^{\frac{1}{p}-\frac{1}{2}}.$$

Finally we have:

$$k \ge \begin{cases} c_p(\epsilon) \cdot n^{\frac{2}{p}}, & 2 < p < \infty \\ c_p(\epsilon) \cdot n, & 1 \le p < 2. \end{cases}$$

In order to prove Dvoretzky's theorem (Theorem 1.2) we need to estimate E and b. By multiplying our symmetric convex body (or the corresponding norm) by a constant we can always assume that $b = 1$ which exactly means that the Euclidean ball lies completely inside $B_{\|\cdot\|}$ - the unit ball of the new norm. But if the euclidian ball is much smaller than $B_{\|\cdot\|}$, then we get nothing, because in such a case E would also be very small and our estimate of the dimension would be meaningless. It is thus better to choose the multiple in such a manner that the Euclidean ball will touch the unit sphere of our norm from the inside. Even then it is possible that E is arbitrarily small. Lecture 5 Nov 27, 2006

The way to overcome this difficulty is to deal first with more general Euclidean norms than the canonical one, replacing the ℓ_2 norm with more general inner product norms, or equivalently replacing the Euclidean ball with a general ellipsoid and proving a similar theorem (Theorem 3.7 below). We shall see later that deducing Theorem 1.2 from the existence of large dimensional Ellipsoidal sections is easy.

Definition 3.5. *An ellipsoid n-dimensional ellipsoid $\mathcal{E} \in \mathbb{R}^n$ is a set of the form*

$$\mathcal{E} = \{\sum_{i=1}^n a_i x_i \;:\; \sum_{i=1}^n a_i^2 \le 1\},$$

for some linearly independent $x_1, \ldots, x_n \in \mathbb{R}^n$. If we define a linear automorphism T of $\mathbb{R}^n$ by $Te_i = x_i$ (where e_i is the standard basis of $\mathbb{R}^n$) we get that $\mathcal{E} = TB_2^n$. The inner product induced by the ellipsoid $\mathcal{E}$ is just

$$< \cdot, \cdot >_{\mathcal{E}} = < T^{-1}x, T^{-1}y > .$$

The next lemma shows that at least in some sense the unit spheres of the given norm and of some contained ellipsoid are not too far apart.

Lemma 3.6. *(Dvoretzky-Rogers) Let* $\|\cdot\|$ *be some norm on* $\mathbb{R}^n$ *and denote its unit ball by* $K = B_{\|\cdot\|}$. *Let* $\mathcal{E}$ *be the (unique) ellipsoid of maximal volume inscribed in* K *and* $|\cdot|$ *- the norm induced by* $\mathcal{E}$. *Then there exist* $x_1, \ldots, x_n \in \partial\mathcal{E}$ *(the boundary of* $\mathcal{E}$*) orthonormal with respect to the inner product* $< \cdot, \cdot >_{\mathcal{E}}$ *such that*

$$e^{-1}(1 - \frac{i-1}{n}) \leq \|x_i\| \leq 1, \quad \text{for all } 1 \leq i \leq n.$$

Remark: This is a weaker version of the original Dvoretzky-Rogers lemma. It shows in particular that half of the x_i-s have norm bounded from below: for all $1 \leq i \leq \lfloor \frac{n}{2} \rfloor$ $\|x_i\| \geq (2e)^{-1}$. This is what will be used in the proof of the main theorem.

Proof. First of all choose an arbitrary $x_1 \in \partial\mathcal{E}$ of maximal norm. Of course, $\|x_1\| = |x_1| = 1$. Suppose we have chosen $\{x_1, \ldots, x_{i-1}\}$ that are orthonormal with respect to $\mathcal{E}$. Choose x_i as the one having the maximal norm among all $x \in \partial\mathcal{E}$ that are orthogonal to $\{x_1, \ldots, x_{i-1}\}$. Our original ellipsoid is $\mathcal{E} = \{\sum_{i=1}^n a_i x_i \ : \ \sum_{i=1}^n a_i^2 \leq 1\}$. Define a new ellipsoid which is smaller in some directions and bigger in others:

$$\overline{\mathcal{E}} = \{\sum_{i=1}^n a_i x_i \ : \ \sum_{i=1}^{j-1} \frac{a_i^2}{a^2} + \sum_{i=j}^{n} \frac{a_i^2}{b^2} \leq 1\}.$$

Suppose, $\sum_{i=1}^n b_i x_i \in \overline{\mathcal{E}}$. Then for each $x \in span\{x_j, \ldots, x_n\} \bigcap \partial\mathcal{E}$ we have $\|x\| \leq \|x_j\|$. Moreover $\sum_{i=1}^{j-1} b_i x_i \in a\mathcal{E}$, hence $\|\sum_{i=1}^{j-1} b_i x_i\| \leq a$; and $\sum_{i=j}^n b_i x_i \in b\mathcal{E}$, hence $\|\sum_{i=j}^n b_i x_i\| \leq \|x_j\| b$. Thus

$$\|\sum_{i=1}^n b_i x_i\| \leq \|\sum_{i=1}^{j-1} b_i x_i\| + \|\sum_{i=j}^n b_i x_i\| \leq a + \|x_j\| \cdot b.$$

The relation between the volumes of $\mathcal{E}$ and $\overline{\mathcal{E}}$ is $Vol(\overline{\mathcal{E}}) = a^{j-1} b^{n-j+1} Vol(\mathcal{E})$. If $a + \|x_j\| \cdot b \leq 1$, then $\overline{\mathcal{E}} \subseteq K$. Using the fact that $\mathcal{E}$ is the ellipsoid of the maximal volume inscribed in K we conclude that

$$\forall a, b, j \text{ s.t. } a + \|x_j\| \cdot b = 1, \quad a^{j-1} b^{n-j+1} \leq 1.$$

Substituting $b = \frac{1-a}{\|x_j\|}$ and $a = \frac{j-1}{n}$ it follows that for every $j \geq 2$

$$\|x_j\| \geq a^{\frac{j-1}{n-j+1}}(1-a) = \left(\frac{j-1}{n}\right)^{\frac{j-1}{n-j+1}} \left(1 - \frac{j-1}{n}\right) \geq e^{-1}\left(1 - \frac{j-1}{n}\right).$$

□ check

Exercise: exercises

1*. Let $\|\cdot\|$ be some norm on $\mathbb{R}^n$ and $\mathcal{E}$ - the ellipsoid of a maximum volume inscribed in $B_{\|\cdot\|}$, then:

$$B_{\|\cdot\|} \subseteq \sqrt{n}\mathcal{E}.$$

2. The previous exercise may be hard; prove at least that there exists an absolute constant $C > 0$ such that:

$$B_{\|\cdot\|} \subseteq C\sqrt{n}\mathcal{E}.$$

Theorem 3.7. *(Dvoretzky) For every $\epsilon > 0$ there exists a positive constant $c(\epsilon)$ such that if $\|\cdot\|$ is some norm on $\mathbb{R}^n$ then there exists a subspace $V \subseteq \mathbb{R}^n$ and an ellipsoid $\overline{\mathcal{E}}$ satisfying:*

1. $\dim V = k$, *where* $k \geq c(\epsilon) \log n$.

2. *For every* $x \in V$*:*

$$(1-\epsilon)\|x\|_{\overline{\mathcal{E}}} \leq \|x\| \leq (1+\epsilon)\|x\|_{\overline{\mathcal{E}}}.$$

Proof. Let $\mathcal{E}$ be (the) ellipsoid of maximal volume inscribed in $K = B_{\|\cdot\|}$ and T - a linear automorphism of $\mathbb{R}^n$ such that $TB_2^n = \mathcal{E}$. Denote $H = T^{-1}K$. Then the Euclidean ball B_2^n will be the ellipsoid of the maximal volume inscribed in H. It is sufficient to prove the theorem for the norm $\|\cdot\|_H$ (the norm for which H serves as the unit ball). check

Suppose we have proved, that

$$E = \int_{S^{n-1}} \|x\|_H dx \geq c\sqrt{\frac{\log n}{n}}, \tag{3.0.6}$$

for some absolute constant $c > 0$. Then by letting $b = 1$ (because $B_2^n \subseteq H$) and applying Theorem 3.1 we get that there exists a subspace $V \subseteq \mathbb{R}^n$ of dimension $k \geq c \log n$ satisfying for every $x \in V$

$$(1-\epsilon)E \cdot \|x\|_2 \leq \|x\|_H \leq (1+\epsilon)E \cdot \|x\|_2,$$

which will finish the proof.

We now turn to prove the inequality 3.0.6. According to Dvoretzky-Rogers lemma 3.6 there exist orthonormal vectors $x_1, \dots, x_n$ such that for all $1 \leq i \leq \lfloor \frac{n}{2} \rfloor$ $\|x_i\|_H \geq e/2$.

Note that $S^{n-1} = \{\sum_{i=1}^n a_i x_i \ : \ \sum_{i=1}^n a_i^2 = 1\}$, hence,

$$\begin{aligned}
\int_{S^{n-1}} \|x\|_H dx &= \int_{S^{n-1}} \|\sum_{i=1}^n a_i x_i\|_H da = \\
&= \int_{S^{n-1}} \frac{1}{2}(\|\sum_{i=1}^{n-1} a_i x_i + a_n x_n\|_H + \|\sum_{i=1}^{n-1} a_i x_i - a_n x_n\|_H) da \geq \\
&\geq \int_{S^{n-1}} \max\{\|\sum_{i=1}^{n-1} a_i x_i\|_H, \|a_n x_n\|_H\} da \geq \\
&\geq \int_{S^{n-1}} \max\{\|\sum_{i=1}^{n-2} a_i x_i\|_H, \|a_{n-1} x_{n-1}\|_H, \|a_n x_n\|_H\} da \geq \cdots \geq \\
&\geq \int_{S^{n-1}} \max_{1\leq i\leq n}\{\|a_i x_i\|_H\} da \geq \frac{e}{2}\int_{S^{n-1}} \max_{1\leq i\leq \lfloor\frac{n}{2}\rfloor}\{|a_i|\} da
\end{aligned}$$

The measure da is the normalized measure on the sphere. So the last integral is exactly the expectation of the vector with the normally distributed independent coordinates, normalized in the ℓ_2 norm:

$$\int_{S^{n-1}} \max_{1\leq i\leq \lfloor\frac{n}{2}\rfloor}\{|a_i|\} da = \mathbb{E}\frac{\max_{1\leq i\leq \lfloor\frac{n}{2}\rfloor}\{|g_i|\}}{(\sum_{i=1}^n g_i^2)^{1/2}} = \frac{\mathbb{E}\max_{1\leq i\leq \lfloor\frac{n}{2}\rfloor}\{|g_i|\}}{\mathbb{E}(\sum_{i=1}^n g_i^2)^{1/2}} \tag{3.0.7}$$

why?

To evaluate the denominator observe, that by concavity of the root function and Jensen inequality:

$$\mathbb{E}(\sum_{i=1}^n g_i^2)^{1/2} \leq (\mathbb{E}\sum_{i=1}^n g_i^2)^{1/2} = \sqrt{n}.$$

Next we use the properties of the normal distribution to get $\mathbf{P}(|g_i| > t) \geq ce^{-t^2/2}$ for some constant $c > 0$. Let $m = \lfloor\frac{n}{2}\rfloor$.

$$\mathbf{P}(\max_{1\leq i\leq m}|g_i| > t) = 1 - \prod_{i=1}^m \mathbf{P}(|g_i| \leq t) = 1 - \prod_{i=1}^m (1 - \mathbf{P}(|g_i| > t)) \geq 1 - (1 - ce^{-t^2/2})^m.$$

Substituting $t = \sqrt{2\log m}$ we get:

$$\mathbf{P}(\max_{1\leq i\leq m}|g_i| > \sqrt{2\log m}) = 1 - (1 - \frac{c}{m})^m \approx 1 - e^{-c} \geq c',$$

for some absolute constant $c' > 0$. Apply Chebyshev's inequality to get the estimate on expectation:

$$\mathbb{E}\max|g_i| \geq \sqrt{2\log m}\mathbf{P}(\max|g_i| > \sqrt{2\log m}) \geq c''\sqrt{\log n}.$$

Plugging in the last estimate into 3.0.7 we finish the proof:

$$E = \int_{S^{n-1}} \|x\|_H dx \geq c'' \sqrt{\frac{\log n}{n}}.$$

□

To deduce the **The Proof of Dvoretzky's theorem 1.2** from Theorem 3.7 it is enough to prove the following claim.

Claim 3.8. *Let $\mathcal{E}$ be some ellipsoid in $\mathbb{R}^k$. Then there exists a subspace $W \subseteq \mathbb{R}^k$ of dimension $l = \lfloor \frac{k}{4} \rfloor$ and $0 < r < \infty$ such that:*

can improve to $\lceil \frac{k}{4} \rceil$ and even to $\lceil \frac{k}{2} \rceil$

$$r(B_2^k \bigcap W) = \mathcal{E} \bigcap W.$$

THIS WILL MOST PROBABLY BE THE LAST LECTURE TO BE TYPED AND PUT ON THIS SITE FOR A WHILE. YOU MAY WANT TO TAKE NOTES DURING FUTURE LECTURES.

Lecture 6
Dec 4, 2006

VOLUNTEERS FOR TeXing NOTES ARE MOST WELCOME.

Proof. First of all we will find $\lfloor \frac{k}{2} \rfloor$ vectors that are orthogonal both with respect to the usual inner product $< \cdot, \cdot >$ and with respect to the inner product $< \cdot, \cdot >_{\mathcal{E}}$ induced by $\mathcal{E}$. Let E be some subspace of $\mathbb{R}^k$, then we denote by

$$E^{\perp} = \{x \ : \ \forall y \in E \ \ < x, y >= 0\}$$

the orthogonal complement of E and by

$$E^{\perp}_{\mathcal{E}} = \{x \ : \ \forall y \in E \ \ < x, y >_{\mathcal{E}}= 0\}$$

the complement subspace of E, orthogonal to E with respect to $\mathcal{E}$.

Choose some $v_1 \in \mathbb{R}^k$ such that $\|v_1\|_2 = 1$. Notice, that the dimension of the subspace $(span\{v_1\})^{\perp} \bigcap (span\{v_1\})^{\perp}_{\mathcal{E}}$ is at least $k-2$. Hence, we can choose a vector v_2 satisfying $\|v_1\|_2 = 1$ and in addition $< v_1, v_2 >= 0$ and $< v_1, v_2 >_{\mathcal{E}}= 0$. We can repeat this iterative procedure $\lfloor \frac{k}{2} \rfloor$ times and get a set of vectors $\{v_i\}_{i=1}^{\lfloor \frac{k}{2} \rfloor}$ which are orthonormal with respect to the euclidean ball and orthogonal with respect to $\mathcal{E}$.

Now using the above sequence of vectors, we are going to construct $\lfloor \frac{k}{4} \rfloor$ vectors which are of the same length and orthogonal both in the euclidean norm and in the norm induced by $\mathcal{E}$. Suppose we arranged the vectors in such a way, that $\|v_1\|_{\mathcal{E}} \geq \|v_2\|_{\mathcal{E}} \geq \cdots \geq \|v_{\lfloor \frac{k}{2} \rfloor}\|_{\mathcal{E}} > 0$. Pick some number $\|v_{\lfloor \frac{k}{4} \rfloor}\|_{\mathcal{E}} \geq a \geq \|v_{\lfloor \frac{k}{4} \rfloor + 1}\|_{\mathcal{E}}$.

Now for every $1 \leq i \leq \lfloor \frac{k}{4} \rfloor$ we choose $0 < \lambda_i \leq 1$ in such a way that

$$\lambda_i \|v_i\|^2_{\mathcal{E}} + (1 - \lambda_i) \|v_{\lfloor \frac{k}{2} \rfloor - i + 1}\|^2_{\mathcal{E}} = a^2.$$

Now construct a sequence of $\lfloor \frac{k}{4} \rfloor$ vectors $u_i = \sqrt{\lambda_i} v_i + \sqrt{(1-\lambda_i)} v_{\lfloor \frac{k}{2} \rfloor - i + 1}$. Of course those vectors are again orthogonal both in euclidian norm and in the norm induced by $\mathcal{E}$. Moreover

$$\|u_i\|_2^2 = \lambda_i \|v_i\|_2^2 + (1-\lambda_i)\|v_{\lfloor \frac{k}{2} \rfloor - i+1}\|_2^2 = 1,$$

and

$$\|u_i\|_{\mathcal{E}}^2 = \lambda_i \|v_i\|_{\mathcal{E}}^2 + (1-\lambda_i)\|v_{\lfloor \frac{k}{2} \rfloor - i+1}\|_{\mathcal{E}}^2 = a^2.$$

Hence, we can choose $W = span\{u_1, \ldots, u_{\lfloor \frac{k}{4} \rfloor}\}$ and

$$\begin{aligned} B_2^n \bigcap W &= \{\sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i u_i \ : \ \|\sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i u_i\|_2 = (\sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i^2)^{1/2} \le 1\} = \\ &= \{\sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i u_i \ : \ \|\sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i u_i\|_{\mathcal{E}} = a \cdot (\sum_{i=1}^{\lfloor \frac{k}{4} \rfloor} a_i^2)^{1/2} \le a\} = a \cdot \mathcal{E} \bigcap W. \end{aligned}$$

□

We have seen in previous lectures, that ℓ_p^n has a $(1+\epsilon)$-euclidian section of dimension

$$k \ge \begin{cases} c_p(\epsilon) \cdot n^{\frac{2}{p}}, & 2 < p < \infty \\ c_p(\epsilon) \cdot n, & 1 \le p < 2 \end{cases}$$

Exercise: Prove, that for $2 < p < \infty$ one can get an estimation $k \ge c(\epsilon) p n^{\frac{2}{p}}$, where $c(\epsilon)$ depends only on ϵ. exercise

Hint: in the proof of the case $2 < p < \infty$ show that $\mathbb{E}(\sum_{i=1}^n |g_i|^p)^{1/p} \ge c\sqrt{p} n^{1/p}$.

In the next claim we prove that the dependance on n in the estimation of k for the case $2 < p < \infty$ is the best possible, namely:

Claim 3.9. *Let $2 < p < \infty$ and suppose we have that ℓ_2^k $(1+\epsilon)$-embeds into ℓ_p^n, meaning that there exists a linear operator $T : \mathbb{R}^k \to \mathbb{R}^n$ such that*

$$\|x\|_2 \le \|Tx\|_p \le (1+\epsilon)\|x\|_2,$$

then $k \le c_p(\epsilon) n^{2/p}$.

Proof. Let $T : \mathbb{R}^k \to \mathbb{R}^n$, $T = (a_{ij})_{i=1}^{n}{}_{j=1}^{k}$ be the linear operator from the statement of the claim. Then for every $x \in \mathbb{R}^k$:

$$(\sum_{j=1}^k x_j^2)^{1/2} \le (\sum_{i=1}^n |\sum_{j=1}^k a_{ij} x_j|^p)^{1/p} \le (1+\epsilon)(\sum_{j=1}^k x_j^2)^{1/2}. \tag{3.0.8}$$

In particular, for every $1 \leq l \leq n$, substituting instead of x the l-th row of T we get:

$$(\sum_{j=1}^{k} a_{lj}^2)^p \leq \sum_{i=1}^{n} |\sum_{j=1}^{k} a_{ij} a_{lj}|^p \leq (1+\epsilon)^p (\sum_{j=1}^{k} a_{lj}^2)^{p/2}.$$

Hence, for every $1 \leq l \leq n$:

$$(\sum_{j=1}^{k} a_{lj}^2)^{p/2} \leq (1+\epsilon)^p.$$

Let $g_1, \ldots, g_k$ be independent standard normal random variables. Then using the fact that $\sum_{j=1}^{k} g_i a_j$ has the same distribution as $(\sum_{j=1}^{k} a_j^2)^{1/2} g_1$ and the left hand side of the inequality 3.0.8 we have

$$\mathbb{E}(\sum_{j=1}^{k} g_j^2)^{p/2} \leq \mathbb{E}(\sum_{i=1}^{n} |\sum_{j=1}^{k} g_j a_{ij}|^p) = \sum_{i=1}^{n} \mathbb{E}(|g_1|^p (\sum_{j=1}^{k} a_{ij}^2)^{p/2}) =$$

$$= \mathbb{E}|g_1|^p \cdot \sum_{i=1}^{n} (\sum_{j=1}^{k} a_{ij}^2)^{p/2} \leq (1+\epsilon)^p \mathbb{E}|g_1|^p n.$$

On the other hand we can evaluate $\mathbb{E}(\sum_{j=1}^{k} g_j^2)^{p/2}$ from below using the convexity of the exponent function for $p/2 > 1$:

$$\mathbb{E}(\sum_{j=1}^{k} g_j^2)^{p/2} \geq (\mathbb{E}\sum_{j=1}^{k} g_j^2)^{p/2} = k^{p/2}.$$

Combining the last two inequalities we finally get the upper bound for k:

$$k \leq (1+\epsilon)^2 (\mathbb{E}|g_1|^p)^{2/p} n^{2/p}.$$

□

Remark:

1. There exist absolute constants $0 < c \leq C < \infty$ such that $c\sqrt{p} \leq (\mathbb{E}|g_1|^p)^{1/p} \leq C\sqrt{p}$. This can be proved by integrating by parts. Hence, for $p = \log n$ we have: exercise

   $$k \leq (1+\epsilon)^2 \log(n) n^{2/\log n}.$$

   Note, that $\log(n^{1/\log n}) = 1$, thus:

   $$k \leq (1+\epsilon)^2 e^2 \log(n).$$

   Hence, if we $(1+\epsilon)$-embed ℓ_2^k into $\ell_{\log n}^n$, then $k \leq (1+\epsilon)^2 e^2 \log(n)$, which means that the $\log n$ bound in the Dvoretzky's theorem is sharp.

2. However, the exact dependence on ϵ is an open question. From the proof of Dvoretzky's theorem we got an estimation $k \geq \frac{c\epsilon^2}{\log(1/\epsilon)} \log n$, see the Open Problems below for the best that is known.

Now we will see another way of obtaining an upper bound on k, which, unlike the estimate in Remark 1, tend to 0 as $\epsilon \to 0$. It still leaves a big gap with the lower bound above.

Claim 3.10. *If ℓ_2^k $(1+\epsilon)$-embeds into ℓ_∞^n, then*

$$k \leq \frac{C \log n}{\log(1/c\epsilon)},$$

for some absolute constants $0 < c, C < \infty$.

Proof. Assume we have $(1+\epsilon)$-embedding of ℓ_2^k into ℓ_∞^n using the linear operator $T = (a_{ij})_{i=1}^{n}{}_{j=1}^{k}$. Then every $x \in \mathbb{R}^k$ satisfies

$$(1-\epsilon)(\sum_{j=1}^{k} x_j^2)^{1/2} \leq \max_{1 \leq i \leq n} |\sum_{j=1}^{k} a_{ij} x_j| \leq (\sum_{j=1}^{k} x_j^2)^{1/2}. \tag{3.0.9}$$

This means that there exist vectors $v_1, \ldots, v_n \in \mathbb{R}^k$ such that for every $x \in \mathbb{R}^k$:

$$(1-\epsilon)\|x\|_2 \leq \max_{1 \leq i \leq n} < v_i, x > \leq \|x\|_2. \tag{3.0.10}$$

In particular, $\|v_i\|_2 \leq 1$ for every $1 \leq i \leq n$.

Suppose $x \in S^{k-1}$, then the left hand side of 3.0.10 states that there exists an $1 \leq i \leq n$ such that $< v_i, x > \geq (1-\epsilon)$, hence:

$$\|x - v_i\|_2^2 = \|x\|_2^2 + \|v_i\|_2^2 - 2 < v_i, x > \leq 2 - 2(1-\epsilon) = 2\epsilon.$$

Thus, the vectors $v_1, \ldots, v_n$ form a $\sqrt{2\epsilon}$-net on the S^{k-1}, which means that n is much larger (exponentially) then k.

In order to formalize the last statement we are going to use volume arguments, just as in proof of the cardinality of ϵ-nets. We have

$$\begin{aligned} & \bigcup_{i=1}^{n} B(v_i, 2\sqrt{2\epsilon}) \supseteq B_2^k \setminus (1-\sqrt{2\epsilon}) B_2^k \\ \Rightarrow \quad & nVolB(0, 2\sqrt{2\epsilon}) \geq VolB(0,1) - VolB(0, 1-\sqrt{2\epsilon}) \\ \Rightarrow \quad & n(2\sqrt{2\epsilon})^k \geq 1 - (1-\sqrt{2\epsilon})^k \geq \sqrt{2\epsilon} k (1-\sqrt{2\epsilon})^{k-1}. \end{aligned}$$

This gives for $\epsilon < \frac{1}{32}$ and $k \geq 12$

$$n \geq \frac{k}{2}(\frac{1}{4\sqrt{2\epsilon}})^{k-1} \geq (\frac{1}{4\sqrt{2\epsilon}})^{k/2},$$

or

$$k \leq \frac{4 \log n}{\log \frac{1}{32\epsilon}}.$$

□

Open problems

1. The best known estimates for $c(\epsilon)$ in Dvoretzky's Theorem 1.2 are

$$\frac{c\epsilon}{(\log 1/\epsilon)^2} \leq c(\epsilon) \leq \frac{C}{\log 1/\epsilon}.$$

   There is still a big gap between the two estimates. One can also ask for the best dependence on ϵ of ϵ Euclidean sections of specific bodies, in particular the ℓ_p balls.

2. Construction of explicit embeddings of close to the best possible dimensions. The only satisfactory results are for ℓ_4^n and ℓ_∞^n.

Exercise: Show that ℓ_2^k $(1+\epsilon)$-embeds into ℓ_∞^n for exercise

$$k \geq \frac{c}{\log(1/\epsilon)} \log(n).$$

Hint: Follow the proof of the claim 3.10 in opposite direction starting with ϵ-net in S^{k-1}.

Lecture 7 lecture 7 Dec 11, 2006

• Construction of Haar measure on homogeneous spaces of compact groups. (The first few pages in [MS].)

• The Johnson–Lindenstrauss Lemma. ([S])

Lecture 8 lecture 8 Dec 18, 2006

• Continuation of the J–L Lemma.

• Survey of tight embedding results in ℓ_p^n spaces. ([JS])

• Finite dimensional subspaces of L_1 and zonoids.

- ℓ_p^n embeds in L_1 if $1 \le p \le 2$.

Lecture 9 lecture 9 Dec 25, 2006

- Change of density and Lewis' lemma.

Lecture 10 lecture 10 Jan 1, 2007

- Splitting of atoms.

- Concentration of Rademacher sums.

- Embedding n-dimensional subspace of L_p in $\ell_p^{Cn^2}$, $1 \le p \le 2$.

Exercise: Extend to $p > 2$.
Exercise: Show that any n-point set in L_1 $(1+\epsilon)$ embeds into $\ell_1^{Cn\log n/\epsilon^2}$.

- Introduction to K convexity.

Lecture 11 lecture 11 Jan 8, 2007

- More on K covexity: $K(L_p)$ is of the order of $\max\{p, p/(p-1)\}$, $K(X)$ is of the order of $\sqrt{\dim(\mathrm{X})}$ for subspaces of L_1.

- Khinchine's inequality with asymptotically best constant.

- Statement of Th: X n-dimensional subspace of L_1, then X $(1+\epsilon)$ embeds into ℓ_1^m with m at most $\frac{C}{\epsilon^2}K(X)n$.

- Beginning of proof: Reduction to Gaussians, Consequence of Slepian's lemma (delaying full statement and proof of Slepian for week after next), Statement of the two propositions from [JS] with indication how to finish the proof of the theorem based on them.

Lecture 12 lecture 12 Jan 15, 2007

- Proof of the two propositions.
- Statement of Slepian's Lemma. Introduction to Gaussian processes.

Lecture 12 lecture 13 Jan 22, 2007

- Continuation of the introduction to Gaussian processes/vectors. Proof of Slepian's lemma.

- Take-home exam given.

References

[JS] W.B. Johnson and G. Schechtman, *Finite dimensional subspaces of L_p*, Handbook of the geometry of Banach spaces, available on-line at:
http://www.wisdom.weizmann.ac.il/∼gideon/papers/finiteLpJan01.ps

[MS] V. Milman and G. Schechtman, *Asyptotic theory of finite-dimensional normed spaces*, Lecture Notes in Mathematics, 1200, Springer-Verlag, Berlin, 1986

[P] G. Pisier, *The volumes of convex bodies and Banach space geometry*, Cambridge University Press, Cambridge 1989

[S] G. Schechtman, *Concentration, results and applications*, Handbook of the geometry of Banach spaces, available on-line at:
http://www.wisdom.weizmann.ac.il/∼gideon/papers/concentrationNov19.ps

[Schn] R. Schneider, *Convex bodies: the Brunn-Minkowski theory*, Encyclopedia of Mathematics and its Applications, 44. Cambridge University Press, Cambridge, 1993.