Contents

To Anna, who always stands by me

Acknowledgements

I would like to thank Gideon Schechtman, my supervisor, for his guidance throughout this research, and especially for his great patience.

I am grateful to my parents Simon and Ludmila and for my brother Michael for their unlimited love and permanent support. Without them the work would never have come into existence.

I am thankful to Anna Rapoport, for an unconditional support.

I would also like to thank all my friends and especially Eliezer Schochat for many useful discussions and nice ideas; Dalia Krieger, without whom I wouldn't have study here; and certainly my dear friends Felix Sychougin, Denis Zilber, Vitaly Azarin for encouragement and friendship.

And finally, I want to thank Weizmann Institute of Science and the Faculty of Mathematics and Computer Science for all the knowledge I have got and for great conditions they provided to make the studies interesting.

Chapter 1

Introduction

> *The mathematical sciences particularly exhibit order, symmetry, and limitation; and these are the greatest forms of the beautiful.*
> Aristotle, *Metaphysica, 3-1078b.*

Man had admired symmetry since antiquity, and in the ordinary usage the words "symmetric" and "beautiful" are held to be almost synonymous. Among all geometrical figures, an Euclidean ball or simply a sphere is the most symmetric one. Looking at a general geometrical figure at high dimension one may ask the question of how far is it from being a ball. The next natural question to ask is how can we beautify it, or in other words make it closer to a sphere, by cutting it with some lower dimensional section or a subspace. The answer to this question was given in 1961 by A. Dvoretzky. His result, known as "Dvoretzky's Theorem" was really the starting point of the modern theory of convex geometry in high dimensions. It claims, that every convex, symmetric body in $\mathbb{R}^n$, has an "almost" spherical section of dimension of order $\log n$. Although $\log n$ is the best possible order of magnitude in the case of a general body (the worst case is the cube), there are particular interesting examples in which we can prove the existence of almost spherical section of order n. In this work we will concentrate on one of these examples - the cross-polytope B_1^n(see Figure 2.1a).

This geometrical problem could be formulated in a functional analytic form. Finding a spherical section of the cross-polytope is equivalent to finding a subspace of $\mathbb{R}^n$ on which the norms of L_1^n and L_2^n are equivalent. In other words, we can consider such a subspace as the embedding of L_2^k into L_1^n with the identity operator, justifying our use of the word "embedding" in the sequel. A general way of proving the existence of such an embedding is to choose k random vectors and to show, that the norms of L_1^n and L_2^n are equivalent on the linear span of these vectors with positive probability. A first natural candidate for random vectors are Gaussian vectors, and indeed, this case can be derived directly from the proof of Dvoretzky's Theorem. In [Sch1], G. Schechtman showed, that we can find such a good embedding by considering only subspaces spanned by vectors whose coordinates are Bernoulli ± 1 random variables. The only limitation in such an embedding is, that we could not make L_1^n and L_2^n norms as close as we want, but rather make it constantly equivalent (not depending on n).

After understanding the asymptotic behavior on n of the dimension k of the almost spherical section of the cross-polytope, we may ask how close k and n could be? The answer to this question is even more surprising and was given by B. Kashin in [K]. Kashin's theorem states, that for every proportion $0 < \delta < 1$, there exists a constant $C = C(\delta) > 0$, such that the cross-polytope of dimension n, for every n possesses a section of dimension $k = \delta n$, which is C-close to Euclidean ball. The last expression means, that for some section of B_1^n, the ratio of radiuses of circumscribed and inscribed Euclidean balls equals C. The proof of the theorem can be found in [P] and in [Sz2] it was shown, that the same is true for a larger class of symmetric convex bodies, the so called bodies with bounded volume-ratio.

A large part of the thesis was inspired by [An]. In his article, G. Anderson considered a random k-dimensional subspace of $\mathbb{R}^n$ and derived formulas to compute the equivalence constant between norms L_1^n and L_2^n on it. The first section of our work is devoted to deriving those formulas in a more general and accurate settings. We also discuss the geometrical meaning of those formulas. It is interesting, that the formulas provide an algorithm for determining the constant of the embedding of L_2^k into L_1^n. For k of order $\log n$ it is even a polynomial time algorithm.

The second section of the thesis concerns with the existence of L_2^k into L_1^n embeddings. We show that the formulas developed in section 2 can be used in the proof instead of the regular ϵ-net arguments. We also show, that the subspace of $\mathbb{R}^n$ on which the norms of L_1^n and L_2^n are equivalent can be constructed using a wide class of random variables, which also includes Gaussians and Bernoulli, generalizing a bit the result of [Sch1].

The third part of the thesis deals with Euclidean sections of dimension $k = \lfloor \delta n \rfloor$, where we allow any $0 < \delta < 1$. We show probabilistic proofs of existence using various random variables and techniques, which are based on the ideas of [An], [Sch3] and [LPRT].

Chapter 2

How well can we embed one space into another?

2.1 Solving maximization problem on subspaces of ℓ_1^n

We denote by L_1^n – the Banach space $\mathbb{R}^n$ equipped with the norm

$$\|x\|_{L_1^n} = \frac{1}{n}\left|\sum_{i=1}^{n} x_i\right|.$$

The unit ball of L_1^n, which we denote by $\bar{B}_1^n$, is the n-dimensional octahedron, or the cross-polytope: the convex-hull of the $2n$ points $(\pm n, 0, \ldots, 0)$, $(0, \pm n, 0, \ldots, 0)$, $\ldots$, $(0, 0, \ldots, 0, \pm n)$, or in other words n times multiple of the unit ball of ℓ_1^n, which is usually denoted by B_1^n. Let us ask the following question: consider some function $\varphi(x) : L_1^n \to \mathbb{R}$, which is convex and homogenous, namely

$$\forall a \in \mathbb{R}. \quad \varphi(ax) = a\varphi(x).$$

Where on $\bar{B}_1^n$ does this function attains its maximum? Well, this is not such a hard question. Our intuition suggests, that the maximum should be attained at the vertices of $\bar{B}_1^n$, which is indeed the case and we will see it later. Hence, we would like to answer a slightly more general question: given a subspace E of dimension $1 \leq k \leq n$, find a maximum of $\varphi(x)$ on $E \bigcap \bar{B}_1^n$. In fact, using the homogeneity of the function, we are actually looking for some finite set of points of E, which will be denoted by $\mathcal{M}_E$ which satisfies

$$\max_{x \in E} \frac{\varphi(x)}{\|x\|_{L_1^n}} = \max_{x \in \mathcal{M}_E} \frac{\varphi(x)}{\|x\|_{L_1^n}}.$$

At the end of the section, we will get a nice description of the set $\mathcal{M}_E$, using the special structure of $\bar{B}_1^n$.

It is well known, that the convex function $\varphi(x)$, defined on some set D, attains its maximum on the extreme points of D. For simplicity, assume, that D is a polytope, i.e. the convex hull of its vertices $\{y_i\}_{i=1}^l$. If x_0 is such that $\varphi(x_0) \geq \varphi(x)$ for all $x \in D$, then for some vertex y_j, $\varphi(x_0) = \varphi(y_j)$. To see this consider scalars $\{\lambda_i\}_{i=1}^l$, satisfying $\sum_{i=1}^l \lambda_i = 1$, such that

$$x_0 = \sum_{i=1}^{l} \lambda_i y_i,$$

and by the definition of convex function, we would get

$$\varphi(x_0) \leq \sum_{i=1}^{l} \lambda_i \varphi(y_i) \leq \max_{1 \leq i \leq l} \{\varphi(y_i)\}.$$

Moreover, if D is a polytope, then its extreme points, or vertices, can be characterized as the facets of dimension 0. Indeed, we are interested in maximizing $\varphi(x)$ on polytope $E \bigcap \bar{B}_1^n$, so our next task will be to characterize the extreme points, or facets of dimension 0 of it.

First, note, that the polytope, which is the convex-hull of the finite number of points, can also be characterized as the intersection of the finite number of hyperplanes (see [G], pp. 31-35). In case of the $\bar{B}_1^n$ these hyperplanes are easily seen to be

$$H_{v^i} = \{x \ : < x, v^i >= 1\},$$

where $< \cdot, \cdot >$ is an inner product, $v^i = (\pm\frac{1}{\sqrt{n}}, \ldots, \pm\frac{1}{\sqrt{n}})$ and the index i runs over all possible choices of signs. Each facet of the $\bar{B}_1^n$ is thus defined by the choice of the signs of the vector v^i. Instead of answering our problem on the whole $\bar{B}_1^n$, let us solve it at each of the 2^n facets separately. For the sake of simplicity, let us consider the facet, defined by the vector $v = (\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})$, which we will denote by F_v, as the rest are dealt similarly.

F_v (see Figure 2.1b) is an $(n-1)$-dimensional polytope – the convex-hull of n points $(n, 0, \ldots, 0)$, $(0, n, 0, \ldots, 0)$, $\ldots$, $(0, 0, \ldots, 0, n)$. Moreover, it can be described as the intersection of n halfspaces $\{x \ : \ x_i \geq 0\}$ with the hyperplane $H_v = \{x \ : < x, v >= 1\}$. Now, let us cut F_v with some subspace E of dimension $k \leq n$. Finally, we reduced the problem to finding the extremal points of $F_v \bigcap E$. The obtained figure will be bounded by the "hyperplanes" $H_i = \{x \ : \ x_i = 0\} \bigcap E$ and $H_v \bigcap E$. The quotation mark actually means, that we do not really know the dimension of the H_i-s, which may be even empty sets, however we know the trivial upper bound

$$\dim H_i \leq k.$$

All the facets of the polytope are intersections of the facets of higher dimensions. Therefore, all extreme points of $F_v \bigcap E$ – the facets of dimension 0 – are intersections of at least $(k-1)$

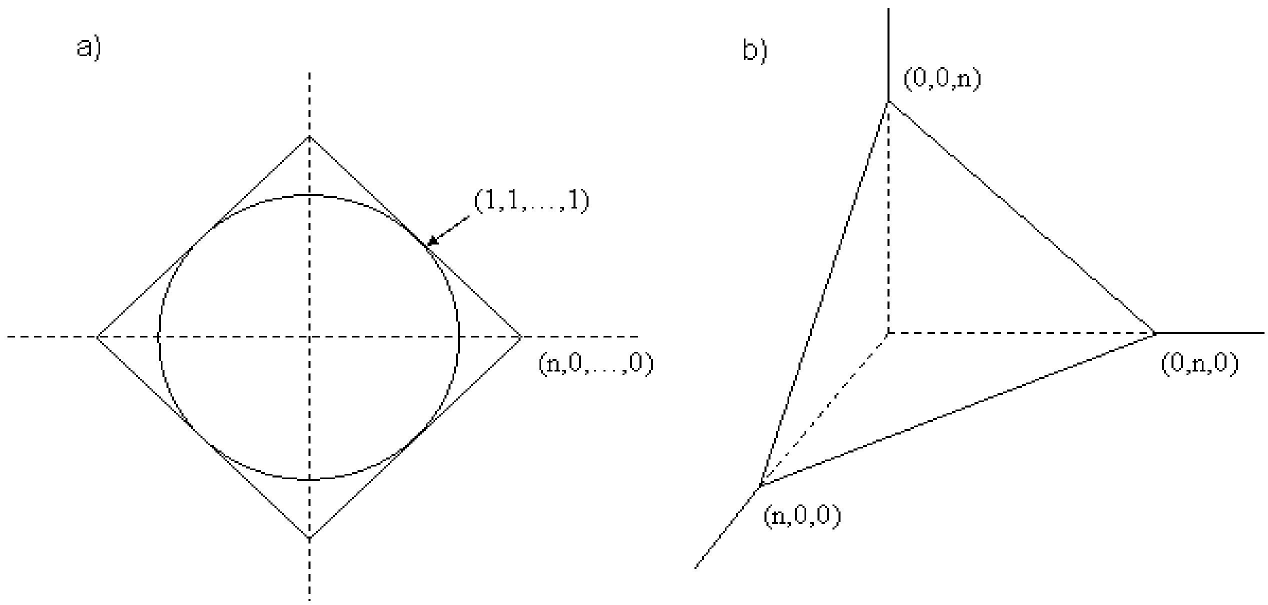

Figure 2.1: a) the cross-polytope and the Euclidean ball inscribed in it; b) one facet of the cross-polytope.

facets H_i – resulting in the facet of dimension at least 1 – with the hyperplane $H_v \bigcap E$. Recalling the definition of H_i-s we can formulate the result as follows. Let $I \subseteq \{1, \dots, n\}$ and define

$$E_I = \{x \in E \;:\; \forall i \in I \;\; x_i = 0\}.$$

and

$$\mathcal{M}_E(v) = \bigcup_{\substack{I \subset \{1,\dots,n\} \\ |I| = k-1 \\ \dim E_I = 1}} \{x \in E_I \;:\; \|x\|_{L_1^n} = 1\}.$$

Then the maximum of $\varphi(x)$ on $F_v \bigcap E$ is attained at some point of $\mathcal{M}_E(v)$:

$$\max_{x \in F_v \bigcap E} \frac{\varphi(x)}{\|x\|_{L_1^n}} = \max_{x \in \mathcal{M}_E(v)} \varphi(x).$$

As we already mentioned, all the facets of the $\bar{B}_1^n$ are similar and we can finally conclude:

Lemma 2.1.1. *Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be the convex homogenous function and define*

$$\mathcal{M}_E = \bigcup_{\substack{I \subset \{1,\dots,n\} \\ |I| = k-1 \\ \dim E_I = 1}} \{x \in E_I \;:\; \|x\|_{L_1^n} = 1\}. \tag{2.1.1}$$

Then

$$\max_{x\in E}\frac{\varphi(x)}{\|x\|_{L_1^n}}=\max_{x\in\mathcal{M}_E}\varphi(x). \tag{2.1.2}$$

Note, that for each E_I we choose at most 1 vector. To emphasize that, we chose the normalization $\|x\|_{L_1^n}=1$ in the definition of the $\mathcal{M}_E$, however in some cases it would be convenient to choose a different one. Eventually, we found a set of $\binom{n}{k-1}$ points on $\bar{B}_1^n$ where $\varphi(x)$ attains its maximum. Clearly, if $E=\mathbb{R}^n$, the set $\mathcal{M}_E$ will contain n points, having $(n-1)$ coordinates 0, which are exactly n out of $2n$ vertices of $\bar{B}_1^n$, just as suggested by our intuition. Most of the arguments above work for any polytope.

2.2 The case $\varphi(x)=\|x\|_{L_2^n}$

Up till now, we were trying to maximize some abstract function $\varphi(x)$. Now we will consider an interesting example. Let us denote by L_2^n – the Banach space $\mathbb{R}^n$ equipped with the norm

$$\|x\|_{L_2^n}=\left(\frac{1}{n}\sum_{i=1}^n x_i^2\right)^{1/2}.$$

The unit ball of L_2^n, which we denote by $\bar{B}_2^n$ is just the Euclidean ball of radius $\sqrt{n}$, or $\sqrt{n}\cdot B_2^n$. Clearly, the norm is always a homogeneous and a convex function, so we would like to apply the results of the previous section to $\|\cdot\|_{L_2^n}$. It is not hard to see that for all $x\in\mathbb{R}^n$

$$\|x\|_{L_1^n}\le\|x\|_{L_2^n}\le\sqrt{n}\|x\|_{L_1^n},$$

or in the geometrical form

$$\bar{B}_2^n\subseteq\bar{B}_1^n\subseteq\sqrt{n}\bar{B}_2^n.$$

Hence, the maximum of $\|x\|_{L_2^n}$ on the $\bar{B}_1^n$ is equal to $\sqrt{n}$ and is obtained at the vertices of the polytope. Imagine, that we cut the $\bar{B}_1^n$ with some k-dimensional subspace E, and see what is the smallest possible radius of the Euclidean ball, that we can inscribe $\bar{B}_1^n\bigcap E$ into. This radius equals exactly the reciprocal of the maximum of the function $\|x\|_{L_2^n}$ on $\bar{B}_1^n\bigcap E$. The current section is devoted to deriving formulas for that quantity, and studying their geometrical meanings.

Let us assume, that the k-dimensional subspace $E\subseteq\mathbb{R}^n$ is spanned by the rows of the $k\times n$ matrix $\mathbb{A}$. Similarly, the orthogonal complementary subspace $E^\perp$ can be identified with $\ker(\mathbb{A})$. Given a set of indices $I\subseteq\{1,\ldots,n\}$ we define $\mathbb{A}_I$ to be the $k\times|I|$ sub-matrix of $\mathbb{A}$, taking the columns with numbers in I. Using the new definitions we can write the set $\mathcal{M}_E$ as follows

$$\mathcal{M}_E = \bigcup_{\substack{I\subset\{1,\dots,n\}\\ |I|=k-1\\ \dim(\ker \mathbb{A}_I^\tau)=1}} \{x\mathbb{A} \ : \ x \in \mathbb{R}^k, \text{ and } \ker \mathbb{A}_I^\tau = \mathrm{span}(x)\}, \tag{2.2.1}$$

and do the same for the orthogonal complement $E^\perp$

$$\mathcal{M}_{E^\perp} = \bigcup_{\substack{I\subset\{1,\dots,n\}\\ |I|=k+1\\ \dim(\ker \mathbb{A}_I)=1}} \{x \in \mathbb{R}^n \ : \ \ker \mathbb{A}_I = \mathrm{span}(x_I)\}. \tag{2.2.2}$$

In order to understand the last expression, it is enough to note, that $\dim E^\perp = (n-k)$, so according to 2.1.1, $\mathcal{M}_{E^\perp}$ includes those x-s, which have $(n-k-1)$ zero coordinates.

Eventually, we are left to find the form of the vector, which spans the kernel of a given $(k-1)\times k$ matrix. Denote by $|A|$ the absolute value of the determinant of the square matrix A, then the following lemma gives us the answer.

Lemma 2.2.1. *Let $A = \{a_{ij}\}$ be an $(k-1)\times k$ matrix of a rank $k-1$. Then:*

$$\ker(A) = \mathrm{span}\{\begin{pmatrix} +|A_{[k]\setminus 1}| \\ \vdots \\ (-1)^{k-1}|A_{[k]\setminus k}| \end{pmatrix}\},$$

where $|A_{[k]\setminus i}|$ is a determinant of a matrix obtained from A by erasing i-th column.

Proof. Clearly, the dimension of $\ker(A)$ is 1. Denote a_i – the i-th row of A. Let $A^i = \begin{bmatrix} a_i \\ \dots \\ A \end{bmatrix}$ be the $k\times k$ singular matrix and $x = \begin{pmatrix} +|A_{[k]\setminus 1}| \\ \vdots \\ (-1)^{k-1}|A_{[k]\setminus k}| \end{pmatrix}$. Then

$$Ax = (< a_1, x >, \dots, < a_{k-1}, x >)^\tau = (\det(A^1), \dots, \det(A^{k-1}))^\tau = 0.$$

■

Combining the lemma with the characterizations of the sets $\mathcal{M}_E$ and $\mathcal{M}_{E^\perp}$ in 2.2.1 and 2.2.2 we can finally derive the following formulas

$$\max_{x\in E}\frac{\|x\|_{L_2^n}}{\|x\|_{L_1^n}}=\max_{x\in\mathcal{M}_E}\frac{\|x\|_{L_2^n}}{\|x\|_{L_1^n}}=\sqrt{n}\max_{\substack{I\subset\{1,\dots,n\}\\|I|=k-1\\\exists i\notin I\ |A_{I\cup\{i\}}|\neq 0}}\frac{\sqrt{\sum_{i\notin I}|A_{I\cup\{i\}}|^2}}{\sum_{i\notin I}|A_{I\cup\{i\}}|},\tag{2.2.3}$$

and

$$\max_{x\in E^\perp}\frac{\|x\|_{L_2^n}}{\|x\|_{L_1^n}}=\max_{x\in\mathcal{M}_{E^\perp}}\frac{\|x\|_{L_2^n}}{\|x\|_{L_1^n}}=\sqrt{n}\max_{\substack{I\subset\{1,\dots,n\}\\|I|=k+1\\\exists i\in I\ |A_{I\setminus\{i\}}|\neq 0}}\frac{\sqrt{\sum_{i\in I}|A_{I\setminus\{i\}}|^2}}{\sum_{i\in I}|A_{I\setminus\{i\}}|}.\tag{2.2.4}$$

The last formulas look a bit awkward, but nevertheless, they have quite a nice geometric interpretation that we will discuss in the next section. Meanwhile, we would like to point out an observation that is an immediate consequence of the formulas. Recall, that for all $x\in\mathbb{R}^n$

$$\frac{1}{\sqrt{n}}\le\frac{\|x\|_2}{\|x\|_1}\le 1.$$

Hence, plugging it into the equation 2.2.3, where the right hand side is the ratio of norms of vectors of dimension $(n-k+1)$, we get the following

Corollary 2.2.2. *Let $1\le k\le n$ and denote $\delta=\frac{k-1}{n}$. Then for every k-dimensional subspace E of $\mathbb{R}^n$, there exists an $x\in E$ satisfying*

$$\frac{\|x\|_{L_2^n}}{\|x\|_{L_1^n}}\ge\sqrt{\frac{1}{1-\delta}}=\sqrt{\frac{n}{n-k+1}}.$$

In particular, if for some $\epsilon>0$ and for all $x\in E$

$$\frac{\|x\|_{L_2^n}}{\|x\|_{L_1^n}}\le(1+\epsilon),$$

then there exists an absolute constant $c>0$, such that

$$k\le c\epsilon n.$$

The corollary suggests, that if we cut the $\bar{B}_1^n$ by some k-dimensional subspace E, the smallest possible radius of the Euclidean ball we can inscribe $\bar{B}_1^n\bigcap E$ into, is $\sqrt{\frac{n}{n-k+1}}$.

2.3 The geometric meaning

The problem we addressed in previous sections has a direct geometrical interpretation. In order to simplify the notation, let us define

$$\mathcal{K}_E = \max_{x\in E} \frac{\|x\|_{L_2^n}}{\|x\|_{L_1^n}}. \tag{2.3.1}$$

Alternatively, $\mathcal{K}_E$ is the smallest number satisfying for every $x \in E$

$$\|x\|_{L_1^n} \le \|x\|_{L_2^n} \le \mathcal{K}_E \|x\|_{L_1^n}. \tag{2.3.2}$$

The following easy claim gives another meaning to the definition of $\mathcal{K}_E$.

Claim 2.3.1. *Let u be a non-zero vector in $\mathbb{R}^n$. Then the size of the projection of the cube B_∞^n onto u, denoted by $P_{\{u\}}$ satisfies*

$$Length(P_{\{u\}}B_\infty^n) = 2\frac{\|u\|_1}{\|u\|_2} = \frac{2}{\sqrt{n}} \cdot \frac{\|u\|_{L_1^n}}{\|u\|_{L_2^n}}.$$

Proof. The proof is a simple observation, that the size of the projection of the vertex $v = (\pm 1, \ldots, \pm 1)$ of B_∞^n onto u equals $\frac{1}{\|u\|_2} < v, u >$ and is maximized for v satisfying $< v, u >= \|u\|_1$.

■

Hence, given the k-dimensional subspace E of $\mathbb{R}^n$ we have

$$\mathcal{K}_E = \frac{\sqrt{n}}{2\min_{x\in E} Length(P_{\{x\}}B_\infty^n)}.$$

In order to further understand the geometrical interpretations of the formulas 2.2.3 and 2.2.4 we need to introduce a few notions.

Definition 2.3.1. A *zonotope Z in $\mathbb{R}^k$ is the Minkowsky sum of a finite number of segments in $\mathbb{R}^k$. Let V_j, $j = 1, \ldots, n$ be the collection of n segments in $\mathbb{R}^k$, then*

$$Z = \sum_{j=1}^n V_j = \{\sum_{j=1}^n x_j \ : \ x_j \in V_j \ \ j = 1, \ldots, n\}.$$

In general the segments are no necessarily centered at zero, but we will restrict ourselves to the centrally symmetric case, namely the segments of the form $V_j = [-v^j, v^j]$, where

$v^j \in \mathbb{R}^k$. Therefore, the resulting zonotope Z would also be symmetric around zero. Any zonotope can be brought to such a position by a shift. Throughout our discussion the term zonotope will refer only to a centrally symmetric one. Centrally symmetric zonotopes can be associated with the projections of the unit ball of ℓ_∞^n – the cube B_∞^n – into $\mathbb{R}^k$. In order to see this, consider a $k \times n$ matrix M which columns are vectors v^j. Then

$$MB_\infty^n = \{\sum_{j=1}^n x_j v^j \ : \ x = (x_1, \dots, x_n) \in B_\infty^n\} = \sum_{j=1}^n [-v^j, v^j] = Z.$$

Looking at the formula (2.2.4) we see, that the nominator and denominator are sums of determinants (squared, in case of the nominator) of all $k \times k$ sub-matrices of a bigger matrix. In case of a nominator (sum of squares), we can use the Cauchy-Binet formula to show

Lemma 2.3.2. *Let M be some $k \times n$ matrix of a rank $k \leq n$. Then for every body $K \subseteq \mathbb{R}^n$*

$$Vol(MK) = \left(\sum_{|I|=k} \det(M_I)^2\right)^{1/2} \cdot Vol(P_{\ker(M)^\perp} K).$$

Proof. The claim is almost an immediate consequence of the Cauchy-Binet formula. The matrix M can be represented as a product $NP_{\ker(M)^\perp}$, were $P_{\ker(M)^\perp}$ is the orthogonal projection onto the orthogonal complement of $\ker(M)$ and N is some non-singular $k \times n$ matrix. Moreover,

$$\sum_{|I|=k} \det(N_I)^2 = \det(NN^*) = \det(MM^*) = \sum_{|I|=k} \det(M_I)^2.$$

Hence,

$$Vol(MK) = Vol(NP_{\ker(M)^\perp} K) = \left(\sum_{|I|=k} \det(N_I)^2\right)^{1/2} \cdot Vol(P_{\ker(M)^\perp} K) =$$
$$= \left(\sum_{|I|=k} \det(M_I)^2\right)^{1/2} \cdot Vol(P_{\ker(M)^\perp} K).$$

■

Let us now apply the last claim to $K = B_\infty^{k+1}$ and the matrix $M = A_I$, where $I \subseteq \{1, \dots, n\}$, $|I| = k+1$, as in the formula 2.2.4. Assume, that $\dim \ker(M) = 1$, then

$$Vol(MK) = \left(\sum_{|J|=k} \det(M_J)^2\right)^{1/2} \cdot Vol(P_{\ker(M)^\perp} K)$$

On the other hand, a vector u, whose coordinates are determinants of $k \times k$ minors of M, spans $\ker(M)$. Hence, according to Claim 2.3.1

$$Length(P_{\ker(M)}K)\Big(\sum_{|J|=k}\det(M_J)^2\Big)^{1/2} = 2\sum_{|J|=k}|\det(M_J)|.$$

Finally, for the volume of the zonotope Z, generated by the columns of the matrix M, we derive

$$Vol(Z) = Vol(MK) = 2\sum_{|J|=k}|\det(M_J)| \cdot \frac{Vol(P_{\ker(M)^\perp}K)}{Length(P_{\ker(M)}K)}. \tag{2.3.3}$$

Hence, we showed, that the following two statements are equivalent:

Lemma 2.3.3. *The volume of a zonotope Z generated by the $n \times (n+1)$ matrix M of rank n is*

$$Vol(Z) = Vol(B_\infty^n)\sum_{|J|=k}|\det(M_J)| = 2^n\sum_{|J|=k}|\det(M_J)|.$$

Lemma 2.3.4. *For an $n \times (n+1)$ matrix M of full rank the following holds*

$$Vol(P_{\ker(M)^\perp}K) = 2^{n-1}Length(P_{\ker(M)}K).$$

The first claim is a form of the more general statement, of [Sh], where the same formula was proved for every $n \times k$ matrix. The second was proved in more general case in [Mc].

Lemma 2.3.5. *Let M be a $k \times n$ dimensional matrix of rank $k < n$. Then*

$$\frac{Vol(M^*D)}{Vol(D)} = \Big(\sum_{|I|=k}\det(M_I)^2\Big)^{1/2},$$

for every k-dimensional body D.

Proof. First of all, for a given matrix M we can find an orthogonal $n \times n$ matrix O satisfying $MO = [N\ 0]$, where N is a non-singular $k \times k$ matrix and moreover $MM^* = NN^*$. Multiplication by orthogonal matrix doesn't change the volume of a body, hence $Vol(M^*D) = Vol(ON^*D) = Vol(N^*D)$. Thus we get

$$\begin{aligned} Vol(M^*D) &= Vol(N^*D) = Vol(D)\det(N) = \\ &= Vol(D)\sqrt{\det(NN^*)} = Vol(D)\sqrt{\det(MM^*)} = Vol(D)\Big(\sum_{|I|=k}\det(M_I)^2\Big)^{1/2}, \end{aligned}$$

where the last equality is the Cauchy-Binet formula.

∎

Using the last claim we can finally formulate the geometric interpretation of the formula 2.2.4. Recall, that the subspace E is spanned by the rows of the matrix A, then letting $D = B_\infty^k$

$$\mathcal{K}_{E^\perp} = \sqrt{n} \max_{|I|=k+1} \frac{Vol(A_I^*(B_\infty^k))}{Vol(A_I(B_\infty^{k+1}))}.$$

Chapter 3

Embedding ℓ_2^k into ℓ_1^n

In the rest of the thesis we would consider the spaces L_1^n and L_2^n and the problem of finding large subspaces of $\mathbb{R}^n$ on which those norms are equivalent. The current section would be devoted to constructing such a subspace of dimension proportional to n, using different random variables.

In the next section we shall consider subspaces whose dimensions are a proportion, arbitrary close to one, of n. It will be devoted to answering the following question. Assume, we are given a number $0 < \delta < 1$, k satisfying $\delta = \frac{k-1}{n}$, can we find a subspace $E \subseteq \mathbb{R}^n$ of dimension k, such that $L_1^n \bigcap E$ and $L_2^n \bigcap E$ would be close? As we pointed out in the introduction, the results we are proving here are well known, but our approach slightly differs from the known proofs. The proof we bring here, as the majority of the known ones, is a probabilistic one. Namely, we define the probability space of the k-dimensional subspaces of $\mathbb{R}^n$ and show that the measure of the subspaces we are interested in is positive. In the current section we will focus our attention on the first problem.

3.1 Embedding with Gaussians

Let us consider a k-dimensional subspace E of $\mathbb{R}^n$, spanned by k vectors $\{g_i\}_{i=1}^k$ of the form

$$g_i = \sum_{j=1}^{n} g_{ij} e_j, \tag{3.1.1}$$

where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ are vectors of the standard basis of $\mathbb{R}^n$ and g_{ij} are independent, identically distributed (i.i.d) Gaussian random variables (r.v) satisfying

$$\mathbb{E}g_{ij} = 0, \text{ and } Var(g_{ij}) = 1.$$

We would like to find the maximum of $\|x\|_{L_2^n}$ on $\bar{B}_1^n \bigcap E$. Our main tool will be the concentration inequality for functions of Gaussian variables, due, in the form below, to Maurey and Pisier. The proof of it can be found in [MS], Theorem V.1.

Theorem 3.1.1. *Let $F : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function with constant d. Let $g = (g_1, \ldots, g_n)$ be a vector of independent, mean zero, normalized in L_2 Gaussian variables. Then,*

$$\mathbf{P}\Big(F(g) > \mathbb{E}F(g) + C\Big) \le \exp(-\frac{2C^2}{\pi^2 d^2}),$$

and

$$\mathbf{P}\Big(F(g) < \mathbb{E}F(g) - C\Big) \le \exp(-\frac{2C^2}{\pi^2 d^2}).$$

For simplicity, let us denote $m = (n-k+1)$ and define two functions $f_p : \mathbb{R}^{mk} \to \mathbb{R}$, $p = 1, 2$, as

$$f_1(x) = \sum_{i=1}^{m} |\sum_{j=1}^{k} x_{ij} v_j|, \tag{3.1.2}$$

and

$$f_2(x) = \Big(\sum_{i=1}^{m} (\sum_{j=1}^{k} x_{ij} v_j)^2\Big)^{1/2}, \tag{3.1.3}$$

where $v = (v_1, \ldots, v_k)$ is an arbitrary vector in S^{k-1}, namely

$$\|v\|_2 = \Big(\sum_{i=1}^{k} v_i^2\Big)^{1/2} = 1.$$

The problem we are facing in this section is clearly related to the ratio of the just defined functions. Hence, we would like to show, that both of them are well concentrated (by the word "well" we mean with a high probability) near some absolute constants. To fulfill this, we will apply the inequalities of Theorem 3.1.1 to the functions $f_1(x)$ and $f_2(x)$, so our first task is to show that the inequality is applicable and to calculate the Lipschits constant of the functions. Let $x, y \in \mathbb{R}^{mk}$. Then, using subsequently the triangular and the Cauchy-Schwarz inequalities we get

$$\Big|f_1(x) - f_1(y)\Big| = \Big|\sum_{i=1}^{m}\Big(|\sum_{j=1}^{k} x_{ij}v_j| - |\sum_{j=1}^{k} y_{ij}v_j|\Big)\Big| \leq \sum_{i=1}^{m}\Big|\sum_{j=1}^{k}(x_{ij} - y_{ij})v_j\Big| \leq$$

$$\leq \sum_{i=1}^{m}\Big(\sum_{j=1}^{k}(x_{ij} - y_{ij})^2\Big)^{1/2} \leq \sqrt{m}\Big(\sum_{i=1}^{m}\sum_{j=1}^{k}(x_{ij} - y_{ij})^2\Big)^{1/2} = \sqrt{m}\|x - y\|_2^2.$$

Hence, $f_1(x)$ is $\sqrt{m}$-Lipschits. For $f_2(x)$ we use once again the Cauchy-Schwarz inequality to get

$$|f_2(x)| = \Big(\sum_{i=1}^{m}(\sum_{j=1}^{k} x_{ij}v_j)^2\Big)^{1/2} \leq \Big(\sum_{i=1}^{m}(\sum_{j=1}^{k} x_{ij}^2)(\sum_{j=1}^{k} v_j^2)\Big)^{1/2} = \|x\|_2,$$

and using the fact that it satisfies the triangular inequality (being a norm)

$$\Big|f_2(x) - f_2(y)\Big| \leq \Big|f_2(x - y)\Big| \leq \|x - y\|_2,$$

we conclude that its Lipschits constant is 1.

The next step will be to evaluate the expectation. To do so we will need an important property of Gaussian distribution – *the rotation invariance.* Let $g = (g_1, \dots, g_n)$ be a vector of independent, mean zero, normalized in L_2 Gaussian variables. Then the distribution of g is invariant under the group of orthogonal transformations of the sphere. It means, that for any $U \subseteq \mathbb{R}^n$ and any $A \in O(n)$ – an $n \times n$ orthogonal matrix, we have

$$\mathbf{P}\Big(g \in U\Big) = \mathbf{P}\Big(Ag \in U\Big). \tag{3.1.4}$$

This property of Gaussian vectors can be checked by direct computations. Now let $u = (u_1, \dots, u_n) \in S^{n-1}$ be some unit vector and consider it as a first row of the orthogonal matrix A. Then it follows directly from (3.1.4) that the distribution of g_1 equals to distribution of $\sum_{i=1}^{n} g_i u_i$.

Applying it for $f_1(x)$ we get

$$\mathbb{E}f_1(g) = \sum_{j=1}^{m}\mathbb{E}\Big|\sum_{i=1}^{k} g_{ij}v_i\Big| = m\mathbb{E}|g_{11}|\Big(\sum_{i=1}^{k} v_i^2\Big)^{1/2} = m\sqrt{\frac{2}{\pi}}.$$

In the case of f_2 we combine this with the Jensen inequality to get

$$\mathbb{E}f_2(g) = \mathbb{E}\Big(\sum_{i=1}^{m}(\sum_{j=1}^{k} g_{ij}v_j)^2\Big)^{1/2} \leq \Big(\sum_{i=1}^{m}\mathbb{E}g_{11}^2(\sum_{j=1}^{k} v_j)^2\Big)^{1/2} \leq \sqrt{m}.$$

Eventually, we can apply Theorem 3.1.1 to both functions $f_1(x)$ and $f_2(x)$:

$$\mathbf{P}\Big(f_1(g) < \Big(\sqrt{\frac{2}{\pi}} - C\Big)m\Big) = \mathbf{P}\Big(f_1(g) < \mathbb{E}f_1(g) - Cm\Big) \le \exp(-\frac{2C^2m}{\pi^2}),$$

and

$$\mathbf{P}\Big(f_2(g) > \Big(1 + C\Big)\sqrt{m}\Big) \le \mathbf{P}\Big(f_2(g) > \mathbb{E}f_2(g) + C\sqrt{m}\Big) \le \exp(-\frac{2C^2m}{\pi^2}),$$

Summarizing the last expressions we get

$$\mathbf{P}\Big(\frac{f_2(g)}{f_1(g)} > \frac{1+C}{\sqrt{2/\pi} - C}\frac{1}{\sqrt{m}}\Big) \le 2\exp\Big(-\frac{2C^2m}{\pi^2}\Big), \tag{3.1.5}$$

and we reached the point where we can formulate the main results of the section:

Theorem 3.1.2. *There exist absolute constants $c_0, c_1 > 0$, such that for every $0 < \delta < c_0$ and for $k = \lfloor \delta n + 1 \rfloor$ we can find a k-dimensional subspace E of $\mathbb{R}^n$ satisfying for all $x \in E$*

$$\|x\|_{L_1^n} \le \|x\|_{L_2^n} \le \Big(\sqrt{\frac{\pi}{2}} + c_1(\delta \log 1/\delta)^{1/2}\Big)\|x\|_{L_1^n}. \tag{3.1.6}$$

Remark: Using the very similar arguments, were instead of dealing with a finite set $\mathcal{M}_E$, one constructs a finite ϵ-net and uses the approximation procedure (see, for example, [MS]), we may actually show the stronger statement. Namely, that for every $\epsilon > 0$, there exists a k-dimensional subspace E of $\mathbb{R}^n$, such that

$$\Big(\sqrt{\frac{\pi}{2}} - \epsilon\Big)\|x\|_{L_1^n} \le \|x\|_{L_2^n} \le \Big(\sqrt{\frac{\pi}{2}} + \epsilon\Big)\|x\|_{L_1^n}.$$

Proof. Let $k = \delta n + 1$ for some $\delta > 0$ (then $m = n - k + 1 = n(1-\delta)$). Let E be the span of the Gaussian vectors $\{g_i\}_{i=1}^k$, that were defined in (3.1.1) and let A be the matrix, whose rows are the g_i-s. In order to prove the statement of the theorem, it is enough to show, that (3.1.6) holds with a positive probability on the choice of g_i-s. However, according to the results of the Section 2.2, it is enough to show that the inequality holds for $\binom{n}{k-1}$ points in $\mathcal{M}_E$ (see 2.2.1). Let us note, that every $y \in \mathcal{M}_E$ is completely defined by $I \subseteq \{1, \dots, n\}$, $|I| = n - k + 1$ and

$$y = vA_I,$$

where $v \in \mathbb{R}^k$ is independent of the entries of A_I. In fact, if we denote $I^c = \{1, \dots, n\} \setminus I$, then v is the basis of the kernel of A_{I^c}. Hence, for a single $y \in \mathcal{M}_E$, the formula (3.1.6) yields

$$\mathbf{P}\Big(\frac{\|y\|_2}{\|y\|_1} > \frac{1+C}{\sqrt{2/\pi} - C}\frac{1}{\sqrt{m}}\Big) = \mathbf{P}\Big(\frac{f_2(g)}{f_1(g)} > \frac{1+C}{\sqrt{2/\pi} - C}\frac{1}{\sqrt{m}}\Big) \le 2\exp\Big(-\frac{2C^2m}{\pi^2}\Big).$$

Now we have to ensure that the same inequality holds for all $y \in \mathcal{M}_E$ simultaneously:

$$\mathbf{P}\left(\frac{\|y\|_2}{\|y\|_1} > \frac{1+C}{\sqrt{2/\pi}-C}\frac{1}{\sqrt{m}} \;:\; \text{for some } y \in \mathcal{M}_E\right) \le \binom{n}{k-1} 2\exp\left(-\frac{2C^2 m}{\pi^2}\right) \le$$
$$\le \exp\left(-\frac{2C^2}{\pi^2}(1-\delta)n + \delta n \log\frac{1}{\delta}\right).$$

Eventually, there exists a $c_0 > 0$, such that for $0 < \delta < c_0$ we can choose some $0 < C < \sqrt{2/\pi}$ of order $(\delta \log 1/\delta)^{1/2}$, in such a way, that the last expression is less then e^{-n}. Therefore, it follows, that for $k = \lfloor \delta n + 1 \rfloor$, there exists a $k \times n$ matrix A with Gaussian entries, such that the subspace of $\mathbb{R}^n$ spanned by the rows of A satisfies

$$\max_{x\in E}\frac{\|x\|_{L_2^n}}{\|x\|_{L_1^n}} = \max_{x\in\mathcal{M}_E}\frac{\|x\|_{L_2^n}}{\|x\|_{L_1^n}} \le \frac{1+(\delta\log 1/\delta)^{1/2}}{\sqrt{2/\pi}-(\delta\log 1/\delta)^{1/2}}\sqrt{\frac{n}{m}} \le \sqrt{\frac{\pi}{2}}(1+c_1(\delta\log 1/\delta)^{1/2}).$$

■

3.2 Proportional embedding with general random variables

In [Sch1] G. Schechtman proved, that not only there exists a subspace E of $\mathbb{R}^n$ of dimension $k = cn$, for some constant $0 < c < 1$, such that the norms of L_1^n and L_2^n are equivalent on E, but that such a subspace can be spanned by vectors with ± 1 entries. In this section we will show how this result can be slightly generalized. Namely, instead of the Bernoulli random variables, we will consider ξ - an arbitrary mean zero random variable satisfying $\mathbb{E}e^{\alpha|\xi|} < \infty$ for some $\alpha > 0$. This is called the *Cramer condition, and it is equivalent to the following: there exist constants* $a, \alpha > 0$ *such that*

$$\mathbf{P}\Big(|\xi| > t\Big) \le ae^{-\alpha t} \quad \textit{for all } t. \tag{3.2.1}$$

This condition is satisfied by various random variables including Gaussian r.v., Bernoulli ± 1, *variables, which are uniformly distributed on some finite symmetric segment, e.g. Let us define for a general random variable* ξ

$$\|\xi\|_p = \Big(\mathbb{E}|\xi|^p\Big)^{1/p}.$$

Using the condition (3.2.1) we conclude that

$$\|\xi\|_1 = \mathbb{E}|\xi| = \int_0^\infty \mathbf{P}\Big(|\xi| > t\Big)dt \le \int_0^\infty ae^{-\alpha t}dt = a\alpha^{-1}, \tag{3.2.2}$$

and similarly

$$\|\xi\|_2^2 = \mathbb{E}|\xi|^2 = \int_0^\infty \mathbf{P}\Big(|\xi| > t\Big) dt^2 \le \int_0^\infty a e^{-\alpha t} dt^2 = 2a\alpha^{-2}. \tag{3.2.3}$$

Let us consider a k-dimensional subspace E of $\mathbb{R}^n$, spanned by k vectors $\{a_i\}_{i=1}^k$ of the form

$$a_i = \sum_{j=1}^n \xi_{ij} e_j, \tag{3.2.4}$$

ξ_{ij} are independent, i.i.d copies of ξ. Denote A - matrix, whose rows are a_i-s. Eventually, we would like to prove the following:

Theorem 3.2.1. *There exits constants $C, \delta > 0$, such that for $k = \lfloor \delta n + 1 \rfloor$ there exists a subspace E of $\mathbb{R}^n$, spanned by k vectors of the form (3.2.4), such that for every $x \in E$*

$$\|x\|_{L_1^n} \le \|x\|_{L_2^n} \le C\|x\|_{L_1^n}.$$

The following generalized concentration inequality for norms of sums of independent random variables, proved in [V] lies in the basis of our proof:

Lemma 3.2.2. *Let $X_1, \ldots, X_n$ be independent Banach space valued random variables satisfying $\mathbf{P}(\|X_i\| > t) \le a e^{-\alpha_i t}$ for all t and i. Let $d \ge \max_{1\le i\le n} \alpha_i^{-1}$ and $b \ge a \sum_{i=1}^n \alpha_i^{-2}$. Then setting $S_n = \sum_{i=1}^n X_i$ we have*

$$\mathbf{P}\Big(|\|S_n\| - \mathbb{E}\|S_n\|| > t\Big) \le \begin{cases} 2\exp(-t^2/32b), & \text{for } 0 \le t \le 4b/d; \\ 2\exp(-t^2/8d), & \text{for } t \ge 4b/d. \end{cases}$$

In order to apply the above inequality in the most general case, we need some way of estimating expectations. The following lemma is taken from [V]. In fact it is a generalization of a Khinchine inequality (see [MS]):

Lemma 3.2.3. *Let $\{\xi_i\}_{i=1}^n$ be a sequence of real valued i.i.d mean zero random variables. Then for every sequence of scalars $\{a_i\}_{i=1}^n$*

$$\frac{1}{2} A_p \|\xi_1\|_{\min(2,p)} \Big(\sum_i a_i^2\Big)^{1/2} \le \|\sum_i a_i \xi_i\|_p \le 2B_p \|\xi_1\|_{\max(2,p)} \Big(\sum_i a_i^2\Big)^{1/2},$$

where A_p, B_p are constants from the classical Khinchine inequality.

In particular, we will be interested in the cases $p = 1, 2$, for which $A_1 = 1/\sqrt{2}$, $B_1 = 1$ (see [Sz1]), and $A_2 = B_2 = 1$.

Finally we are ready to prove the Theorem:

Proof. Denote $m = n-k+1$. Fix $v = (v_1, \dots, v_k) \in S^{k-1}$, such that for some $I \subseteq \{1, \dots, n\}$, $|I| = k-1$, $vA_I = 0$. Consider a sequence of independent random variables

$$X_{ij} = v_i \xi_{ij} e_j, \quad i = 1, \dots, k, \quad j = 1, \dots, m,$$

and their sum $S_v = \sum_{i=1}^k \sum_{j=1}^m X_{ij}$, then clearly $S_v = vA_{\{1,\dots,n\}\setminus I}$.

First of all we need to compute the parameters b, d from Lemma 3.2.2. Note

$$\mathbf{P}(\|X_{ij}\|_2 > t) = \mathbf{P}(\|X_{ij}\|_1 > t) = \mathbf{P}(|v_i \xi| > t) \le a e^{-\alpha |v_i|^{-1} t},$$

hence we can choose

$$d = \alpha^{-1} \text{ and } b = a \sum_{i=1}^k \sum_{j=1}^m \alpha^{-2} |v_i|^2 = am\alpha^{-2}.$$

The next step will be to compute expectations. Applying Lemma 3.2.3 we have

$$\mathbb{E}\|S_v\|_1 = \sum_{j=1}^m \mathbb{E}\left| \sum_{i=1}^k v_i \xi_{ij} \right| \ge \frac{\|\xi\|_1}{2\sqrt{2}} m.$$

In order to compute the expectation of $\|S_v\|_2$ we will use Jensen inequality, Lemma 3.2.3 to get

$$\mathbb{E}\|S_v\|_2 = \mathbb{E}\left(\sum_{j=1}^m \left(\sum_{i=1}^k v_i \xi_{ij} \right)^2 \right)^{1/2} \le$$

$$\le \left(\sum_{j=1}^m \mathbb{E}\left(\left| \sum_{i=1}^k v_i \xi_{ij} \right| \right)^2 \right)^{1/2} \le 2\|\xi\|_2 \sqrt{m}.$$

Note, that $4b/d = 4am/\alpha$, hence letting $t = c_1 m$ and $c_1 \le \min\{\frac{\|\xi\|_1}{2\sqrt{2}}, \frac{4a}{\alpha}\}$, we ensure that $t \le 4b/d$, and applying Lemma 3.2.2 we get:

$$\mathbf{P}\left(\|S_v\|_1 \le \left(\frac{\|\xi\|_1}{2\sqrt{2}} - c_1 \right) m \right) \le \mathbf{P}\left(\|S_v\|_1 \le \mathbb{E}\|S_v\|_1 - t \right) \le e^{-c_1^2 \alpha^2 m / 32a}. \tag{3.2.5}$$

For the $\|S_v\|_2$ let $t = c_2\sqrt{m}$, where $c_2 > \frac{4a}{\alpha}$, satisfying $t > 4b/d$, and get

$$\mathbf{P}\left(\|S_v\|_2 \ge \left(2\|\xi\|_2 + c_2 \right) \sqrt{m} \right) \le \mathbf{P}\left(\|S_v\|_2 \ge \mathbb{E}\|S_v\|_2 + t \right) \le e^{-c_2^2 \alpha m / 8}. \tag{3.2.6}$$

Recalling the definition of S_v, we can combine thelast two statements to get that thereare constants $C_1, C_2 > 0$, such that for every $v \in S^{k-1}$, satisfying $vA \in \mathcal{M}_E$, and

$$\mathbf{P}\left(\frac{\|S_v\|_2}{\|S_v\|_1} \ge \frac{C_1}{\sqrt{m}} \right) \le e^{-C_2 m}.$$

Hence, by the formula (2.2.3) and the fact that$\mathcal{M}_E$ contains at most$\binom{n}{k-1}$ vectors, we get

$$\mathbf{P}\Big(\max_{x\in E}\frac{\|x\|_{L_2^n}}{\|x\|_{L_1^n}}\geq\frac{C_1}{\sqrt{1-\delta}}\Big)\leq\binom{n}{k-1}e^{-C_2 m}\leq\exp\Big((-C_2(1-\delta)+\delta\log 1/\delta)n\Big)\leq e^{-cn},$$

for some constant $c>0$, concluding the proof.

■

Remark: *It follows from the proof of Theorem 3.2.1, that one can choose*

$$C_1=\frac{\|\xi\|_2}{\sqrt{2}\|\xi\|_1}+\epsilon,$$

where ϵ *is of order* $(\delta\log 1/\delta)^{1/2}$.

Chapter 4

High proportional embeddings

4.1 Gaussian random variables

Examining the proof of Theorem 3.1.2 we note, that the reason we could not make δ to be closer to 1 than some small constant, was the inequality we used, that did not allow us to increase the constant C as much as we wanted. In order to fix that, we need another type of inequality than the one that was used in the previous section. The price that we will pay eventually, is the dependance of the equivalency constant on δ. Eventually we want to prove:

Theorem 4.1.1. *For every $0 < \delta < 1$, $k = \lfloor \delta n + 1 \rfloor$ and for $C \geq 2\delta^{-\delta/(1-\delta)}$, there exists a k-dimensional subspace E of $\mathbb{R}^n$, such that for every $x \in E$*

$$\|x\|_{L_1^n} \leq \|x\|_{L_2^n} \leq C\|x\|_{L_1^n}.$$

Let us once again observe the formula (2.2.3). Note, that if the entries of the matrix A are independent Gaussian random variables, then the vectors in $y \in \mathcal{M}_E$ are just Gaussians, namely, y_i-s are iid normally distributed random variables, of dimension $m = n - k + 1$. The reason for that is once again the rotation invariance, namely, the linear combination of the Gaussian is again Gaussian. That observation allows us to use slightly different norm-ratio inequality and get the desired result. Before proceeding with the proof let us state and prove the norm-ratio inequality:

Lemma 4.1.2. *Let $X \in \mathbb{R}^n$ be a Gaussian vector, then for every $\beta > 1$*

$$\mathbf{P}\left(\frac{\|X\|_{L_2^n}}{\|X\|_{L_1^n}} > \beta\sqrt{\frac{2e}{\pi}}\right) \leq \beta^{-n}.$$

Proof. First of all, let us recall the formula for the volume of the n-dimensional Eucledian ball. We can estimate gamma function using Stirling's formula

$$\Gamma(\frac{n}{2}+1) \leq \sqrt{2\pi} e^{-n/2} \left(\frac{n}{2}\right)^{(n+1)/2},$$

and thus,

$$Vol(B_2^n) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} \geq \left(\sqrt{\frac{2\pi e}{n}}\right)^n.$$

Hence, the volume of $\bar{B}_2^n$, which is just $\sqrt{n}B_2^n$, is greater than $(\sqrt{2\pi e})^n$. We already pointed out, that the unit ball of ℓ_1^n - B_1^n - consists of 2^n pieces, which are copies of the figure bounded by the hyperplanes $< x, e_i >= 0$ and $\{x \ : \ \sum_{i=1}^n x_i = 1\}$ (see figure 2.1(b)). That figure is a cone of height 1 over a base, which is the analogues piece in $\mathbb{R}^{n-1}$ and by induction its volume equals

$$\frac{1}{n} \cdot \frac{1}{n-1} \cdot \cdots \cdot \frac{1}{2} \cdot 1 = \frac{1}{n!},$$

hence the volume of B_1^n is $\frac{2^n}{n!}$ and the volume of $\bar{B}_1^n = nB_1^n$ is $\frac{(2n)^n}{n!}$. Once again, using the Strirling's formula we can estimate the ratio

$$\left(\frac{Vol(\bar{B}_1^n)}{Vol(\bar{B}_2^n)}\right)^{1/n} \leq \sqrt{\frac{2e}{\pi}}.$$

Next, we will show another way to compute the last expression. Let us first write a formula for the volume of $\bar{B}_1^n$ in spherical coordinates. Denote by $\bar{S}^{n-1}$ the surface of the unit ball $\bar{B}_2^n$. For each direction $\theta \in \bar{S}^{n-1}$, the radius of $\bar{B}_1^n$ in that direction is $\|\theta\|_{L_1^n}^{-1}$. Hence, the volume of $\bar{B}_1^n$ is

$$Vol(\bar{B}_1^n) = nVol(\bar{B}_2^n) \int_{\bar{S}^{n-1}} \int_0^{\|\theta\|_{L_1^n}^{-1}} s^{n-1} ds d\sigma(\theta) = Vol(\bar{B}_2^n) \int_{\bar{S}^{n-1}} \|\theta\|_{L_1^n}^{-n} d\sigma(\theta),$$

where $d\sigma(\theta)$ is the normalized rotational invariant measure on the surface of $\bar{S}^{n-1}$. Combining both results we get

$$\left(\frac{Vol(\bar{B}_1^n)}{Vol(\bar{B}_2^n)}\right) = \int_{\bar{S}^{n-1}} \|\theta\|_{L_1^n}^{-n} d\sigma(\theta) \leq \left(\frac{2e}{\pi}\right)^{n/2}.$$

The last step is to apply Markov's inequality. We have for every $\beta > 1$

$$\sigma\left(x \in \bar{S}^{n-1} \ : \ \|x\|_{L_1^n}^{-n} \geq \beta^n \left(\frac{2e}{\pi}\right)^{n/2}\right) \leq \beta^{-n},$$

and the inequality follows, because rotational invariant measure on the sphere is equivalent to the Gaussian measure in the sense that

$$\sigma(A \subseteq \bar{S}^{n-1}) = \mathbf{P}\left(\frac{X}{\|X\|_{L_2^n}} \in A\right),$$

where X is a Gaussian vector.

■

Finally we are ready to prove Theorem 4.1.1:

Proof. Fix $\delta > 0$ and denote $k = \delta n + 1$, $m = n - k + 1 = (1-\delta)n$. Consider E a k-dimensional subspace of $\mathbb{R}^n$ spanned by the vectors $\{g_i\}_{i=1}^k$, that were defined in 3.2.4, and a $k \times n$ matrix A whose rows are g_i-s. This is a matrix that every entry in it is a standard Gaussian variable. Let $x \in S^{k-1}$ be such, that $xA \in \mathcal{M}_E$, hence, for some $I \in \{1, \dots, n\}$, $|I| = k-1$,

$$xA_I = 0, \quad \text{and set} \quad y_x = xA_{\{1,\dots,n\}\setminus I}.$$

The rotational invariance property of Gaussian distribution implies, that $y_x \in \mathbb{R}^m$ is also a Gaussian vector. Hence, we are able to apply to it inequality of Lemma 4.1.2 to get for every $\beta > 1$

$$\mathbf{P}\left(\frac{\|y_x\|_{L_2^n}}{\|y_x\|_{L_1^n}} > \beta\sqrt{\frac{2e}{\pi}}\right) \le \beta^{-m}.$$

Now, in order to conclude, that the norms L_2^n and L_1^n are equivalent on the whole E, we need to apply the last inequality to all the $\binom{n}{k-1}$ different y-s in $\mathcal{M}_E$. Using Stirling's formula we get

$$\mathbf{P}\left(\exists y \in \mathcal{M}_E. \ \frac{\|y\|_{L_2^n}}{\|y\|_{L_1^n}} > \beta\sqrt{\frac{2e}{\pi}}\right) \le \binom{n}{k-1} \cdot \beta^{-m} \le$$
$$\le \left(\frac{n}{k-1}\right)^{k-1} \cdot \beta^{-m} = \delta^{-\delta n} \cdot \beta^{-(1-\delta)n} \le 2^{-n},$$

for $\beta \ge 2\delta^{-\delta/(1-\delta)}$, concluding the proof of the theorem.

■

Remark 1: *As we mentioned above, the proof is very similar to existing ones. The difference is in utilizing the (relatively small) set $\mathcal{M}_E$ rather than using an ϵ-net argument.*

Remark 2: *In [K] Kashin also proved, that we can split $\mathbb{R}^n$ into two mutually orthogonal subspaces, such that the norms of L_1^n and L_2^n are equivalent on both of them. In fact, we can derive this result from the proof of Theorem 4.1.1.*

Corollary 4.1.3. *For every $0 < \delta < 1$, $k = \lfloor \delta n + 1 \rfloor$, there exists a constant $C = C(\delta) > 0$ and a k-dimensional subspace E of $\mathbb{R}^n$, such that for every $x \in E \bigcup E^\perp$*

$$\|x\|_{L_1^n} \leq \|x\|_{L_2^n} \leq C\|x\|_{L_1^n}.$$

Proof. Fix $0 < \delta < 1$ and denote $k = \delta n$ and $m = n - k = (1-\delta)n$. There exist constants $C_1, C_2 > 0$ such that with probability at least $1 - 2 \cdot 2^{-n}$ we can find a k-dimensional subspace E such that both E and its m-dimensional orthogonal complementary subspace $E^\perp$ satisfy Theorem 4.1.1 with constants C_1 and C_2. This follows from the proof of Theorem 4.1.1 together with the fact that the distribution of the resulting subspace E is the same as that of $E^\perp$. This follows from the orthogonal invariance of the Gaussian distribution. Hence, $C = \max\{C_1, C_2\}$ will satisfy

$$\|x\|_{L_1^n} \leq \|x\|_{L_2^n} \leq C\|x\|_{L_1^n},$$

for every $x \in E \bigcup E^\perp$.

■

4.2 Matrices with entries in a small set

In the previous section we showed a probabilistic proof to the fact that there exists a subspace of $\mathbb{R}^n$ of any dimension proportional to n such that the norms of L_1^n and L_2^n are equivalent on it. We proved for subspaces spanned by Gaussian vectors. Moreover, we know that it is possible to use another, more simple kind of variables. G. Anderson in [An] proved the same fact using random variables taking integer values in some fixed interval. The idea of the proof was to show that the inequality similar to 4.1.2 holds for such a variables, using the fact that they are in some sense "close" to being Gaussians. His motivation was to try and find an explicit (rather than by probabilistic means) such spaces.

G. Schechtman improved Anderson's result by finding even simpler vectors which do the job. He did it by overcoming the difficulties that we faced in our attempts to make the proportion close to 1 using concentration inequalities techniques. In [Sch3] he suggested some simple *kind of vectors to show the existence of the subspace E of dimension $n/2$ such that on both E and $E^{\perp}$ the norms L_1^n and L_2^n are equivalent, thus improving the result of [An]. We will show that this result could be extended to every proportion of n. In the next section we will show that a similar idea, with the help of some nice inequality, allowed authors of [LPRT] to prove the statement using vectors of Bernoulli random variables. Although both statements could be proved using the technique developed in previous sections, we would like to use the original ϵ-net arguments, which will make the presentation more clear. Before proceeding, let us remind the reader what ϵ-net argument is*

Lemma 4.2.1.

1. *For every $\epsilon > 0$ there exists an ϵ-net on S^{n-1} of cardinality at most $(\frac{3}{\epsilon})^n$.*

2. *Let $F : \mathbb{R}^n \to \mathbb{R}$ be a non-negative, convex homogeneous function and suppose that for some ϵ-net $\mathbb{N}$ of S^{n-1} one has $a \leq F(x) \leq b$ for every $x \in \mathbb{N}$. Then*
$$a - \frac{\epsilon}{1-\epsilon} b \leq F(x) \leq \frac{1}{1-\epsilon} b,$$
for every $x \in S^{n-1}$. In particular, if $\epsilon \leq a/3b$, then $\frac{1}{2}a \leq F(x) \leq \frac{3}{2}b$ for every $x \in S^{n-1}$.

For the proof of both statements see [MS] (4.1). This Lemma actually allows us, starting from some concentration inequality for norm (always a non-negative, convex and homogeneous function) of one vector, to extend it to the whole Euclidean sphere.

Let ξ be a random variable satisfying 3.2.1 and let E be a k-dimensional subspace of $\mathbb{R}^n$ spanned by vectors $\{a_i\}_{i=1}^k$, $a_i = (a_{i1}, \ldots, a_{in})$, where
$$a_{ij} = \begin{cases} \sqrt{k}, & j = i; \\ \xi_{ij}, & n-k+1 \leq j \leq n; \\ 0, & \text{otherwise.} \end{cases} \tag{4.2.1}$$
and ξ_{ij} are independent copies of ξ. The matrix, whose rows are a_i-s, has a form $A = [\sqrt{k}I, B]$, where I is a $k \times k$ identity matrix and B is a $k \times (n-k)$ matrix with entries - independent copies of ξ. We would like to prove the following:

Theorem 4.2.2. *For every $0 < \delta < 1$, $k = \delta n$, there exists a constant $C = C(\delta) > 0$ and a k-dimensional subspace $E = \{xA \ : \ x \in S^{k-1}\} = \{\sqrt{k}x + xB \ : \ x \in S^{k-1}\}$, were A as defined above, such that for every $x \in S^{k-1}$*
$$\|xA\|_{L_1^n} \leq \|xA\|_{L_2^n} \leq C\|xA\|_{L_1^n}.$$

Moreover, the statement holds with high probability on the choice of the matrix A (equivalently B).

The idea of the proof in [Sch3] is as follows: split the subspace E into two disjoint sets $E = E_\gamma \bigcup E_\gamma^c$, were

$$E_\gamma = \{xA \in E \ : \ \|x\|_1 \leq \sqrt{k}\gamma\},$$

for some prescribed constant $\gamma > 0$. The points which are in E_γ^c are "good" and all of them satisfy the statement of the theorem just by definition due to the left hand side of the matrix A. The main part of the proof deals with the "bad" points of E_γ. However, it turns out that those vectors can be "well approximated" by vectors having small support, which in turn could be treated by the concentration inequality for the right hand side of the matrix. Therefore, let us fill the details.

For $1 \leq l \leq k$ denote the set of vectors in S^{k-1} with support at most l by

$$F_l^k = \{x = (x_1, \ldots, x_k) \in S^{k-1} \ : \ x_i \neq 0 \text{ for at most } l \text{ values of } i\}.$$

Using Lemma 3.2.2 and the ϵ-net argument we can derive the following inequalities:

Lemma 4.2.3. *Let B be the $k \times m$ matrix of independent copies of ξ (the right hand side of A). Then the following holds*

1. *There exist constants $C_1, c_1 > 0$, such that*

$$\mathbf{P}\Big(\exists x \in S^{k-1} \ : \ \|xB\| > C_1\Big) \leq e^{-c_1 m},$$

   where $\|\cdot\|$ denotes both L_1^m and L_2^m norms.

2. *There exist constants $C_2, c_2 > 0$, such that for every $x \in S^{k-1}$*

$$\mathbf{P}\Big(\|xB\|_{L_1^m} < C_2\Big) \leq e^{-c_2 m}.$$

3. *There exist constants $C_3, c_3, \alpha > 0$, such that for $l \leq \alpha k$*

$$\mathbf{P}\Big(\exists x \in F_l^k \ : \ \|xB\|_{L_1^m} < C_3\Big) \leq e^{-c_3 m}.$$

The first inequality tells us, that for most of the choices of the $k \times m$ matrix of independent copies of ξ, its norm as the operator from ℓ_2^k to L_1^m is bounded from above by some constant. The second inequality is just the reformulation of Lemma 3.2.2.

Until now we did nothing new. But the following simple observation really makes the difference and allows us to prove Theorem 4.2.2.

Lemma 4.2.4. *Let $x = (x_1, \dots, x_k)$ be a norm one vector in ℓ_2^k and let $0 < \gamma < 1$. Assume $\sum_{i=1}^k |x_i| \le \gamma\sqrt{k}$ then*

$$\left(\sum_{i=l}^k (x_i^*)^2\right)^{1/2} \le l^{-1}\gamma\sqrt{k(k-l+1)}$$

for all $1 \le l \le k$, where $\{x_i^*\}$ denotes the decreasing rearrangement of $\{|x_i|\}$. In particular, if $l = \alpha k$, then

$$\left(\sum_{i=l}^k (x_i^*)^2\right)^{1/2} \le \mu\gamma,$$

where $\mu = \alpha^{-1}\sqrt{1-\alpha}$.

Proof. (Of Theorem 4.2.2)

First of all, fix $0 < \delta < 1$ and denote $k = \delta n$, $m = n - k = (1-\delta)n$. Throughout the proof we will denote by A the random $k \times n$ matrix of the form $[\sqrt{k}I, B]$, where I is a $k \times k$ identity matrix and B is a $k \times m$ matrix with entries - independent copies of ξ.

Let $\Omega = \{A \,:\, \forall x \in S^{k-1} \;\; \|xB\|_{L_1^n} \le C_1\}$. According to the first inequality of Lemma 4.2.3

$$\mathbf{P}(\Omega) = \mathbf{P}(\|xB\|_{L_1^n} \le C_1) \ge 1 - e^{-c_1 m}. \tag{4.2.2}$$

Let $0 < \gamma < 1$ be a constant, whose value will be determined later, and denote $E_\gamma = \{xA \in E \,:\, \|x\|_1 \le \sqrt{k}\gamma\}$. For some $0 < t$ define

$$\Omega(t) = \Omega \bigcap \{A \,:\, \exists x \in E_\gamma \quad \|xA\|_{L_1^n} \le t\},$$

and

$$\overline{\Omega}(t) = \Omega \bigcap \{A \,:\, \exists x \in S^{k-1} \setminus E_\gamma \quad \|xA\|_{L_1^n} \le t\}.$$

First of all note, that if $t < \gamma\delta$ and $A \in \overline{\Omega}(t)$, then by the properties of $x \in S^{k-1} \setminus E_\gamma$

$$\|xA\|_{L_1^n} = \delta k^{-1/2}\|x\|_1 + (1-\delta)\|xB\|_{L_1^n} > \gamma\delta,$$

hence, $\mathbf{P}(\overline{\Omega}(t)) = 0$ for all $0 < t < \gamma\delta$. Therefore, we need only to show that the measure of the complementary set $\Omega(t)$ is also small.

Let $x \in E_\gamma$ and $l = \alpha k$ from the third inequality of Lemma 4.2.3. Then, according to Lemma 4.2.4 we can write $x = y + z$, where

$$y \in F_l^k, \text{ and } \|z\|_2 \leq \mu\gamma.$$

Let us choose a random matrix $A \in \Omega(t)$. Then

$$\begin{aligned}\|xA\|_{L_1^n} &= \delta k^{-1/2}\|x\|_1 + (1-\delta)\|xB\|_{L_1^m} \geq \\ &\geq (1-\delta)(\|yB\|_{L_1^m} - \|zB\|_{L_1^m}) \geq (1-\delta)(\|yB\|_{L_1^m} - \|z\|_2\|B\|_{\ell_2^k \to L_1^m}).\end{aligned}$$

The fact that $A \in \Omega$ implies that the norm of B as an operator from ℓ_2^k to L_1^m is bounded by C_1. Moreover, $y \in F_l^k$ and $\|y\|_2 \geq 1 - \mu\gamma$, hence we can apply the third inequality of Lemma 4.2.3 to get

$$\mathbf{P}\Big(\|xA\|_{L_1^n} \leq (1-\delta)(C_3(1-\mu\gamma) - C_1\mu\gamma)\Big) \leq \mathbf{P}\Big(\|yB\|_{L_1^n} \leq C_3\|y\|_2\Big) \leq e^{-c_3 m}.$$

Choosing γ in such a way that $(C_3(1-\mu\gamma) - C_1\mu\gamma) \geq C_3/2$ we get

$$\mathbf{P}(\Omega(t)) \leq e^{-c_3 m},$$

for $t < C_3(1-\delta)/2$. Hence letting $t = \min\{C_3(1-\delta)/2, \delta\gamma\}$ we conclude that

$$\mathbf{P}\Big(\exists x \in S^{k-1} \;:\; \|xA\|_{L_1^n} \leq t\Big) \leq e^{-c_3 m}. \tag{4.2.3}$$

Now we turn to estimating the upper bound on the L_2^n norm of xA for some $x \in S^{k-1}$. Note

$$\|xA\|_{L_2^n} = (\delta\|x\|_2^2 + (1-\delta)\|xB\|_{L_2^m}^2)^{1/2}.$$

Hence, applying the first inequality of Lemma 4.2.3 we get

$$\begin{aligned}\mathbf{P}\Big(\exists x \in S^{k-1} \;:\; \|xA\|_{L_2^n} \geq (\delta + (1-\delta)C_1^2)^{1/2}\Big) &= \\ = \mathbf{P}\Big(\exists x \in S^{k-1} \;:\; \|xB\|_{L_2^m} \geq C_1\Big) &\leq e^{-c_1 m}.\end{aligned}$$

Finally, combining the last inequality with 4.2.3 we prove the statement of the theorem

$$\mathbf{P}\left(\exists x \in S^{k-1} \;:\; \frac{\|xA\|_{L_2^n}}{\|xA\|_{L_1^n}} \geq \frac{(\delta + (1-\delta)C_1^2)^{1/2}}{\min\{C_3(1-\delta)/2, \delta\gamma\}}\right) \leq e^{-cm},$$

for some constant $c > 0$. ■

Remark: *In the previous section we showed that with high probability we can find a subspace E such that on both E and on its orthogonal complement $E^{\perp}$ the norms of L_1^n and L_2^n are equivalent. Let us consider the case of $\delta = \frac{1}{2}$ and $n = 2k$. If E is spanned by rows of the matrix $A = [\sqrt{k}I \ ; \ B]$, were I is a $k \times k$ identity matrix and B is a $k \times k$) matrix with entries ξ, then $E^{\perp}$ is spanned by the rows of $\overline{A} = [-B^{\tau} \ ; \ \sqrt{k}I]$. Since $-\xi$ has the same properties as ξ, applying the arguments above to $-B^{\tau}$ we can conclude:*

Corollary 4.2.5. *There exists a constant $C > 0$ and a k-dimensional subspace E of $\mathbb{R}^{2k}$, spanned by rows of the matrix $A = [\sqrt{k}I \ ; \ B]$, such that for every $x \in E \bigcup E^{\perp}$*

$$\|x\|_{L_1^n} \leq \|x\|_{L_2^n} \leq C\|x\|_{L_1^n}.$$

4.3 Embeddings with Bernoulli random variables

As we already mentioned in the introduction, in his paper [Sch1], G. Schechtman showed the existence of the proportional embedding of L_2^k into L_1^n using vectors with ± 1 entries. It is, then, a natural question to ask whether for every proportion $0 < \delta < 1$ there exists a subspace $E \subseteq \mathbb{R}^n$ of dimension $k = \delta n$ spanned by the vectors with ± 1 entries, such that the norms of L_2^k and L_1^n are equivalent on it. The positive answer was given by G. Schechtman in [Sch3], where he outlined the idea of the proof. However, the proof is not probabilistic. Namely, it only promises the existence of such a subspace and does not shows what is the probability to find one. In fact we cannot even be sure, that the set of such "good" subspaces, realizing the embedding, have a positive measure among all the subspaces spanned by the vectors with ± 1 entries. This gap was filled in [LPRT]. In the current section we would like to show both proofs of the following:

Theorem 4.3.1. *For every $0 < \delta < 1$ there exits a constant $C = C(\delta) > 0$, such that for $k = \lfloor \delta n \rfloor$ there exists a subspace E of $\mathbb{R}^n$, spanned by the vectors of ± 1, such that for every $x \in E$*

$$\|x\|_{L_1^n} \leq \|x\|_{L_2^n} \leq C\|x\|_{L_1^n}.$$

First, let us present the non-probabilistic proof that was outlined in [Sch3]. The proof is based on the following lemma, which is essentially Proposition 1 in [JS].

Lemma 4.3.2. *Let E be a k-dimensional subspace of ℓ_p^n, $1 \leq p \leq 2$, with $k \geq \lfloor \delta_0 n \rfloor$. Then for every $0 < \delta < 1$ and $m \geq \lfloor \delta n \rfloor$ there exists a constant $C = C(\delta, \delta_0) > 0$, such that E is C-isomorphic to a subspace of ℓ_p^m. More specifically, there exists a map $T : X \to \ell_p^m$, which* *is a restriction of X into m coordinates, such that for every $x \in X$*

$$\|Tx\|_p \leq \|x\|_p \leq C\|Tx\|_p.$$

Proof. (Of 4.3.1) The proof of the Theorem follows almost immediately from Theorem 3.2.1 and the previous lemma. Fix n, $0 < \delta < 1$ and denote $k = \delta n + 1$. Then, according to Theorem 3.2.1 there exists a subspace E' of $\mathbb{R}^m$ of dimension k, such that for every $x \in E'$

$$\|x\|_{L_1^n} \leq \|x\|_{L_2^n} \leq C_1\|x\|_{L_1^n},$$

were $C_1 > 0$ is some absolute constant. If $m \leq n$, then, clearly, E' can also be considered as the subspace of $\mathbb{R}^n$ and we are done. However, in the case of $m > n$, according to Lemma 4.3.2, there exists a constant $C_2 > 0$ and a restriction E of E' onto n coordinates, such that for every $x' \in E'$ and a corresponding $x \in E$

$$\|x'\|_{L_1^n} \leq \|x\|_{L_1^n} \leq \|x\|_{L_2^n} \leq C_1\|x\|_{L_1^n} \leq C_1 \cdot C_2\|x'\|_{L_1^n}.$$

Noting, that $\|x'\|_{L_1^n} \leq \|x'\|_{L_2^n} \leq \|x\|_{L_2^n} \leq C_1 \cdot C_2\|x'\|_{L_1^n}$ finishes the proof.

■

The reader may notice, that our construction of the subspace E was not probabilistic. First we constructed some embedding, and then, using the particular realization, chose the restriction onto n coordinates. Fortunately, utilizing an inequality of [H], the very same technique, that allowed us to prove the main theorem of the previous section, allowed authors of [LPRT] to construct a probabilistic proof of Theorem 4.3.1. Instead of repeating the whole proof, we will just sketch its main idea.

Fix some $0 < \delta < 1$ and $k = \delta n$. Let E be a span of k linearly independent vectors $\{a_i\}_{i=1}^k$ with coordinates ± 1. Denote by A a matrix, whose rows are a_i-s. Then $E = \{xA \,:\, x \in \mathbb{R}^k\}$. For $y \in \mathbb{R}$ and $a \in \mathbb{R}$ define $y \wedge a = y \cdot \mathbf{1}_{[-a,a]}(y)$. Similarly for $x = (x_1, \ldots, x_n)$ let $x \wedge a = (x_1 \wedge a, \ldots, x_n \wedge a)$.

The general line of the proof is similar to that of Theorem 4.2.2. We also split the subspace E into two subsets. The "bad" set

$$E_\gamma = \{xA \in E \,:\, \|x \wedge \gamma\|_2 \leq \sqrt{k}\gamma\},$$

and the "good" one

$$E_\gamma^c = E \setminus E_\gamma.$$

Assume, $y = xA \in E_\gamma$ is in the "bad" set. Define $z = y \wedge \gamma$ and $w = y - z$. Then we can represent y in the form

$$y = z + w,$$

where $\|z\|_2 \leq \gamma$ and w has a bounded support, namely $supp(w) \leq \gamma^{-2}$. We are left to choose the right parameter γ and the rest of the proof for the "bad" points is the same as in the case of Theorem 4.2.2. But what about the "good" points? The following nice inequality, based on the result of Howard and Oskolkov in [H] solves the problem:

Lemma 4.3.3. *For every $0 < t < 1$ and for any $x \in \mathbb{R}^n$ such that $\|x \wedge t^2\|_2 > 1$*

$$\mathbf{P}(\|\Gamma x\|_1 < tN) \leq (4t)^N.$$

Bibliography

[An] G. W. Anderson, *Integral Kashin splittings*, to appear in the Israel Journal of Math.

[E] H. G. Eggleston, *Convexity*, Cambridge Tracts in Mathematics, 47, Cambridge University Press, 1958

[G] B. Grunbaum, *Convex Polytopes*, Interscience Publishers, 1967

[H] R. Howard, *Estimates on the concentration function of sets in $\mathbb{R}^d$: Notes on lectures of Oskolkov*, http://www.math.sc.edu/~howard/Notes/concentration.pdf

[JS] W. B. Johnson and G. Schechtman, *Very tight embeddings of subspaces of L_p, $1 \leq p < 2$, into ℓ_p^n*, Geometric and Funct. Anal., to appear.

[K] B. S. Kashin, *Diameters of some finite-demensional sets and classes of smooth functions*, Math. USSR Izvestija, Vol 11 (1977), No. 2

[L] M. Ledoux, *The concentration of measure phenomenon*, AMS Math. Surveys and Monographs, Vol. 89, 2001

[LPRT] S. Litvak, A. Pajor, M. Rudelson, N. Tomczak-Jaegermann, *Invertibility of Rademacher matrices and random 0-1 polyotpes*, in preparation

[Mc] P. McMullen, *Volumes of projection of unit cubes*, Bull. London Math. Soc., 16, 1984, pp. 278-280

[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

[Sch1] G. Schechtman, *Random embedding of Euclidean spaces in sequence spaces*, Israel J. Math. 40 (1981), 187-192

[Sch2] G. Schechtman, Concentration, results and applications, available on-line at: http://www.wisdom.weizmann.ac.il/ gideon/papers/concentrationNov19.ps

[Sch3] G. Schechtman, Special orthogonal splittings of L_1^{2k}*, Israel J. Math. to appear*

[Sh] G. C. Shephard, Combinatorial properties of associated zonotopes, Can. Journal of Math., 26(2), 1974, pp. 302-321

[Sz1] S. J. Szarek, On the best constant in Khinchine inequality, Studia Math. 58 (1976), 197-208

[Sz2] S. J. Szarek, On Kashin almost Euclidean orthogonal decomposition of ℓ_1^n*, Bull. Acad. Polon. Sci., 26, 1978, pp. 691-694*

[V] R. Vershynin, On large random almost euclidean bases, Acta Mathematica Universitatis Comenianae 69 (2000), no. 2, 137-144