Expectation of p -norm of random matrices with heavy tails

A Dissertation
presented to
the Faculty of the Graduate School
University of Missouri

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
CHANDRASEKAR VAIDYANATHAN
Dr. Stephen Montgomery-Smith, Dissertation Supervisor

JULY 2014

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled

Expectation of p -norm of random matrices with heavy tails

presented by Chandrasekar Vaidyanathan, a candidate for the degree of Doctor of Philosophy of Mathematics, and hereby certify that in their opinion it is worthy of acceptance.

Professor Stephen Montgomery-Smith

Professor Peter Casazza

Professor Nakhle Asmar

Professor William Banks

Professor David Retzloff

DEDICATION

This work is but a small offering of gratitude to my parents and grandparents, for whose tireless efforts, sacrifices and love, I remain indebted forever.

ACKNOWLEDGEMENTS

This therefore, is mathematics: she reminds you of the invisible forms of the soul; she gives life to her own discoveries; she awakens the mind and purifies the intellect; she brings to light our intrinsic ideas; she abolishes oblivion and ignorance which are ours by birth ... **Proclus, the Greek philosopher**

It is indeed a privilege to be doing mathematics and to be blessed with so many people who have helped me in my journey . My foray into Random Matrix Theory was due to Dr.Mark Rudelson. Much of what I know about Random Matrices is due to the exposure he gave me. He showed through his personal example an extraordinary commitment to doing high quality research, which I hope I can emulate in my life. Dr.Pete Casazza, agreed to be my adviser at a critical time in my personal life. The two semesters of research funding he provided, helped me do research while being close to my son. Dr.Stephen Montgomery-Smith, my thesis adviser, helped me define my thesis problem and met with me without fail, whenever I requested and carefully went over many of my arguments. Thanks are also due to Dr.William Banks, for the many games of fighting chess and sharing many a thoughts on life, which provided a welcome change from doing math. I am thankful to Dr.Carmen Chicone, for the many long discussions on the ethical issues that come up during research.

The blog of Terry Tao, Bill Thurston's essays, the many lucid lecture notes of Roman Vershynin and many more mathematicians, who were kind enough to share their work, to all of them I owe a lot of intellectual debt. Wei He, a student at Mizzou, helped me on a variety of mathematical topics, be it commutative algebra or functional analysis. We spent many hours working through many a problems. The

impact of this interaction is immeasurable.

Dr.Gang Zhou, helped me during my transition from engineering to mathematics at Notre Dame. I probably would not be doing mathematics, if not for his timely help. Dr.François Ledrappier, was instrumental in helping me switch from optimization to analysis. The working seminar we had on expander graphs, is one of the more memorable experiences of my academic life. I hope to pay forward this gift of knowledge I received from a wide variety of sources.

There are many people whose company enriched my life in untold ways, during my time at Columbia. I am incredibly grateful to Ms.Lynn Rossi, for introducing me to mindfulness meditation, a life changing experience for me. The Zendo of Prof. Ray Ronci, provided a comfortable environment to meditate. Scott Douglass, Gina Wright, Suzzane Carron, Kevin Leary and Venky Nagar, all helped us during various points of my academic life here. To John Meyer, Ela Celikbas, Shibi Vasudevan, John Fresen, Daniel Fresen and all the students who had to endure my teaching and my personality: Thank you!

Dr. Brad Lehman, of Northeastern University was my first graduate advisor. It was on his research project with Analog Devices Inc., that I began my stint at Analog. Gordon Sterling, my supervisor there, and Dr.Guolin Pan, fostered an environment of constant questioning and creativity that has become a part of me now. It was there that I realized that nothing is impossible when the working environment is right and one is willing to work hard on problems with dedication.

Dr. Oliver Collins, formerly of the University of Notre Dame, was kind enough to take me as his graduate student. He also introduced me to the beautiful books

of Michael Spivak and William Feller. These books changed the course of my life. Mizzou campus has been my home away from home. Ellis library, medical library, math library, business school study center, the many classrooms in the classroom buildings of the college of arts and science, the seminar hall of the math department, have all been places where I have enjoyed doing research uninterrupted, during regular hours and late into the night. I will miss Ms. Sandy Cunnighams warm welcome for lunch at the dining hall and the Mango Peach Smoothie that Marla makes at Infusion. Much of this thesis was typed in the comfortable settings of the computer lab at the college of education. It was a pleasure to use writelatex.com, for typesetting the thesis in the LATEX environment developed by them.

I would also like to thank the front office staff of the math department. Amy Crews, Gwen Gilpin and Kyle Groshong, have all helped me with my administrative needs promptly, despite other pressing matters.

I owe a great deal to my science and math teachers at high school, Ms.Shreelekha Kumaran, Mr.P.Balasubramaniam, Mr.R.Lakshmanan and Ms.C.Jayalakshmi, who saw in a young kid, the burning quest for knowledge and did their best to stoke the sparks of curiosity into a raging fire. Through their selfless act of spending many hours outside the school hours, they not only showed me what real teachers are made of but also nurtured in me the habit of working on problems not just during class hours but also outside of it.

It is hard to imagine life without my son and my wife. To them, I owe a lot more than an acknowledgment here.

Contents

ACKNOWLEDGEMENTS	ii
ABSTRACT	vii
1 History and Motivation	1
2 Introduction	6
2.1 Preliminaries	6
2.2 Marcinkiewicz-Zygmund inequalities	10
2.3 Rosenthal's Inequality	12
2.4 Brunn Minkowski Inequality	14
2.5 Concentration of measure on the sphere	18
3 Gaussian Concentration and its consequences	24
3.1 Preliminaries	24
3.2 Computing norm on a subset of the unit ball of l_2	26
4 Extending the result for p-norm	35
5 Putting it all together	44
BIBLIOGRAPHY	47

Expectation of p-norm of random matrices with heavy tails

Chandra Vaidyanathan

Dr. Stephen Montgomery Smith, Dissertation Supervisor

ABSTRACT

The p-norm ($p > 2$) of a random matrix whose entries are gaussian, subgaussian and log concave have been studied previously. We conjecture the following generalization of the above results for heavy tailed random matrices:

Conjecture 0.1. *Let $p > 2$. Fix $n > 0$. $A = (X_{ij}), i = 1, 2, \dots, n, j = 1, 2, \dots, N$, be a random matrix whose entries are independent random variables with bounded $2p^{th}$ moments, where $N \geq n^{\frac{p}{2}}$. Then, with high probability, $\forall x \in \mathbb{S}^{n-1}$:*
 $cN^{1/p} \leq \|Ax\|_p \leq CN^{1/p}$, for some constants, $c, C > 0$.

We establish the following upper bound, generalising the work of [Latala05]:

Theorem 0.2. *Let $p > 2$. Fix $n > 0$. $A = (X_{ij}), i = 1, 2, \dots, n, j = 1, 2, \dots, N$, be a random matrix whose entries are independent, identically distributed random variables with bounded $2p^{th}$ moments, where $N \geq n^{p-1}$. Then, with high probability, (close to 1): $\forall x \in \mathbb{S}^{n-1}$:*

$$\|Ax\|_p \leq CN^{1/p} \log N [\log(\log N) \log^2(\log(\log N))]^{1-1/p}$$

for some constant C , depending on p and the p th and $2p$ th moments of the random variable. We also get a similar upper bound when the entries of the random matrix are independent random variables with bounded $2p$ th moment and are not necessarily identically distributed.

Chapter 1

History and Motivation

The following Gaussian result, was proved by Yehoram Gordon [Gordon85], on embedding subsets of euclidean space into normed space. Let $(g_j)_{j=1}^N$, $(g_i)_{i=1}^n$ and $(g_{ij})_{(1 \leq i \leq n), (1 \leq j \leq N)}$ be independent standard Gaussian random variables and set $(e_j)_{j=1}^N$ to be the standard unit vector basis on $\mathbb{R}^N$.

Theorem 1.1. Gaussian Embedding theorem *There exists an absolute constant C , for which the following holds. Set $n, k \geq 1$, consider $T \subset \mathbb{S}^{n-1}$ and let $E = (\mathbb{R}^N, \|\cdot\|_p)$ be a normed space such that for every $x \in \mathbb{R}^N$, $\|\cdot\|_p \leq \|x\|_2$. Fix $\varepsilon > 0$ and assume that*

$$\mathbb{E} \sup_{t \in T} \left\| \sum_{i=1}^n g_i t_i \right\|_2 \leq C \varepsilon \mathbb{E} \left\| \sum_{j=1}^N g_j e_j \right\|_p$$

(i.e, for a universal constant c , $N \geq c^p p^{p/2} \varepsilon^{-p} n^{p/2}$).

If the random operator $\Gamma : \mathbb{R}^n \rightarrow \mathbb{R}^N$, is defined by

$$\Gamma t = \sum_{j=1}^N \sum_{i=1}^n g_{ij} t_i e_j$$

then, there is a realisation of Γ such that for every $t \in T$,

$$(1 - \varepsilon) \mathbb{E} \left\| \sum_{j=1}^N g_j e_j \right\|_p \leq \|\Gamma t\|_p \leq (1 + \varepsilon) \mathbb{E} \left\| \sum_{j=1}^N g_j e_j \right\|_p$$

The theorem was proved as an application of the Gaussian min-max theorem by Y. Gordon and is an easy modification of the method of the proof of Dvoretzky's theorem and other applications in [Gordon88] and [Gordon07]. It was rediscovered by G. Schechtman [Schechtman06], who gave it a new proof using the majorizing measures approach. Further generalizations of the above result, in the form of an embedding theorem for subgaussian and log concave cases, where the embedding is not isometric, but is rather isomorphic, was carried out in [MendelsonJaegerman08] and [GuedonRudelson07] respectively. Here, we give a flavor of the proof of the subgaussian embedding theorem, as presented in [Rudelson13].

Definition 1.2. Isotropic random vectors: A random vector in $\mathbb{R}^n$ is said to be isotropic, if

$$\mathbb{E}\langle X, x \rangle^2 = 1, \text{ for all } x \in \mathbb{S}^{n-1}$$

Theorem 1.3. *Let X be an isotropic subgaussian vector in $\mathbb{R}^n$. Let $X_1, \dots, X_N$ be independent copies of X . Then, for any $p > 2$ and any $N \geq n^{p/2}$, the inequality*

$$c \leq \left(\frac{1}{N} \sum_{j=1}^N |\langle x, X_j \rangle|^p \right)^{1/p} \leq C\sqrt{p}$$

holds with high probability for all $x \in \mathbb{S}^{n-1}$.

Proof. Denote by A , the $N \times n$, matrix with rows $X_1, \dots, X_N$. Assume that n is large enough, so that $N \geq n^{p/2} \geq \delta_0^{-1}n$, where δ_0 is an absolute constant that is as defined in [Rudelson06]. Combining this proposition with the inequality $\|y\|_2 \leq N^{1/2-1/p} \cdot \|y\|_p$, valid for all $y \in \mathbb{R}^N$, we obtain

$$\mathbb{P} \left(\min_{x \in \mathbb{S}^{n-1}} \|Ax\|_p \leq c_1 N^{1/p} \right) \leq e^{-c_2 N},$$

which establishes the left hand side inequality, with probability exponentially close to 1. To prove the second inequality, we use the method of majorizing measures, or generic chaining, developed by Talagrand. [Talagrand05]. Let $\{X_t\}_{t \in T}$ be a real valued random process, i.e, a collection of interdependent random variables, indexed by some set T . In the setup below, we can assume that T is finite or countable, eliminating the question of measurability of $\sup_{t \in T} X_t$. We shall call the process $\{X_t\}_{t \in T}$ centered, if $\mathbb{E}X_t = 0$ for all $t \in T$.

Definition 1.4. Let (T, d) be a metric space. A random process $\{X_t\}_{t \in T}$ is called subgaussian with respect to the metric d , if for any $t, s \in T, t \neq s$, the random variable $(X_t - X_s)/d(t, s)$ is subgaussian. A random process $\{G_t\}_{t \in T}$ is called Gaussian with respect to the metric d , if for any finite set $F \subset T$, the joint distribution of $\{G_t\}_{t \in F}$ is Gaussian, and for any $t, s \in T, t \neq s$ $(G_t - G_s)/d(t, s)$ is a $N(0, 1)$ random variable.

We use the following fundamental result due to Talagrand:

Theorem 1.5. *Majorizing measures theorem*

Let (T, d) be a metric space, and let $\{G_t\}_{t \in T}$, be a Gaussian random process with respect to the metric d . For any centered random process $\{X_t\}_{t \in T}$, which is subgaussian with respect to the same metric,

$$\mathbb{E} \sup_{t \in T} X_t \leq C \mathbb{E} \sup_{t \in T} G_t$$

For $(s, y) \in \mathbb{R}^N \times \mathbb{R}^n$, define the random variable $X_{s,y}$ by

$$X_{s,y} = \sum_{j=1}^N s_j \langle X_j, y \rangle.$$

Then, for any $T \subset B_2^N \times B_2^n$, the random process $\{X_{s,y}\}_{(s,y) \in T}$ is a subgaussian with

respect to the Euclidean metric. Indeed, for any $(s, y), (s', y') \in T$,

$$X_{s,y} - X_{s',y'} = \sum_{j=1}^N \left((s_j - s'_j) \langle X_j, y \rangle + s'_j \langle X_j, y - y' \rangle \right)$$

Let $\lambda \in \mathbb{R}$. Since the vector X is centered subgaussian, for any $z \in \mathbb{R}^N$, $\exp(\lambda \langle X, z \rangle) \leq \exp(C\lambda^2 \|z\|_2^2)$. Hence, using the independence of X_j and applying Cauchy-Schwartz inequality we get

$$\begin{aligned} & \mathbb{E} \exp(\lambda(X_{s,y} - X_{s',y'})) \\ &= \prod_{j=1}^N \mathbb{E} \left[\exp(\lambda(s_j - s'_j) \langle X_j, y \rangle + \lambda s'_j \langle X_j, y - y' \rangle) \right] \\ &\leq \prod_{j=1}^N \exp(2C\lambda^2((s_j - s'_j)^2 \|y\|_2^2)) \cdot \prod_{j=1}^N \exp(2C\lambda^2(s'_j)^2 \|y - y'\|_2^2) \\ &\leq \exp(2C\lambda^2(\|s - s'\|_2^2 + \|y - y'\|_2^2)) \end{aligned}$$

Using the Laplace transform property of subgaussian random variables, the r.v.

$$\frac{X_{s,y} - X_{s',y'}}{\|(s, y) - (s', y')\|_2}$$

is sub-gaussian.

Let Y and Z be independent standard Gaussian vectors in $\mathbb{R}^n$ and $\mathbb{R}^N$ respectively.

Set

$$G_{s,y} = \langle s, Z \rangle + \langle y, Y \rangle$$

Then, for any $T \subset \mathbb{R}^N \times \mathbb{R}^n$, $\{G_{s,y}\}_{(s,y) \in T}$ is a Gaussian process with respect to the Euclidean metric. let $1/p + 1/p^* = 1$, and set $T = B_{p^*}^N \times B_2^n \subset B_2^N \times B_2^n$. By the majorizing measures theorem,

$$\mathbb{E} \sup_{(s,y) \in T} X_{s,y} \leq C \mathbb{E} \sup_{(s,y) \in T} G_{s,y}$$

Therefore,

$$\begin{aligned}
\mathbb{E} \sup_{y \in B_2^n} \left(\frac{1}{N} \sum_{j=1}^N |\langle X_j, y \rangle|^p \right)^{1/p} &= \frac{1}{N^{1/p}} \mathbb{E} \sup_{s \in B_{p^*}^N} \sup_{y \in B_2^n} \sum_{j=1}^N s_j \langle X_j, y \rangle \\
&\leq \frac{C}{N^{1/p}} \mathbb{E} \sup_{s \in B_{p^*}^N} \sup_{y \in B_2^n} G_{s,y} = \frac{C}{N^{1/p}} \mathbb{E}(\|Z\|_p + \|Y\|_2) \\
&\leq C(\sqrt{p} + \frac{\sqrt{n}}{N^{1/p}})
\end{aligned}$$

Since, $N \geq n^{p/2}$, the last sum does not exceed $C' \sqrt{p}$. To finish the proof, we combine this estimate with the Chebyshev's inequality. ■

Conjecture 1.6. *Let X be a random vector in $\mathbb{R}^n$, whose entries are independent random variables with $\mathbb{E}|X|^{2p} < \infty$. Let $X_1, \dots, X_N$ be independent copies of X . Then, for any $p > 2$ and any $N \geq n^{p/2}$, the inequality*

$$c \leq \left(\frac{1}{N} \sum_{j=1}^N |\langle x, X_j \rangle|^p \right)^{1/p} \leq C \sqrt{p}$$

holds with high probability for all $x \in \mathbb{S}^{n-1}$.

In the case of this random matrix with heavy tails, the majorization technique does not work directly. The work in this thesis involves an attempt toward establishing the upper bound. We do that by generalizing the beautiful arguments of Latała as in [Latała05].

Chapter 2

Introduction

We derive some of the useful inequalities involving random variables, which will be used later in the proof of the main theorem. To this end, we start with the some simple inequalities involving random variables and proceed to derive Rosenthal's inequality. We then give the proof of the concentration of measure on the sphere, using Brunn Minkowski's inequality.

2.1 Preliminaries

If X is a random variable, letting X' , be an identically distributed independent copy of X , we get the symmetrized random variable by setting, $X^s = X - X'$.

Proposition 2.1. *Let $p \geq 1$, let X and Y , be independent random variables. Suppose that, $\mathbb{E}|X|^p < \infty$, $\mathbb{E}|Y|^p < \infty$, and that $\mathbb{E} Y = 0$. Then,*

$$\mathbb{E}|X|^p \leq \mathbb{E}|X + Y|^p.$$

In particular, if X has mean 0, and X^s is the symmetrized random variable, then

$$\mathbb{E}|X|^p \leq \mathbb{E}|X^s|^p.$$

Proof. Let x be a real number. Then, since Y has mean 0,

$$|x|^p = |x + \mathbb{E}Y|^p \leq \mathbb{E}|x + Y|^p = \int_{-\infty}^{\infty} |x + y|^p dF_Y(y),$$

so that

$$\mathbb{E}|X|^p \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x+y|^p dF_Y(y) dF_X(x) = \mathbb{E}|X+Y|^p$$

The particular case follows by letting, $Y = -X'$, the independent and identically distributed copy of X . ■

Proposition 2.2. *Lyapunov's inequality*

For $0 < r \leq p$, $\|X\|_r \leq \|X\|_p$.

Proof. We use, Holder's inequality, with X replaced by $|X|^p$, and p by r/p , to get,

$$\mathbb{E}|X|^p \leq \| |X|^p \|_{r/p} \cdot 1 = (\|X\|_r)^p$$

■

Proposition 2.3. *c_p -inequality:*

Let $p > 0$. Suppose that $\mathbb{E}|X|^p$ and $\mathbb{E}|Y|^p$ are both finite, then

$$\mathbb{E}|X+Y|^p \leq c_p(\mathbb{E}|X|^p + \mathbb{E}|Y|^p), \text{ where } c_p = 1, \text{ for } 0 < p \leq 1 \text{ and } c_p = 2^{p-1}, \text{ when } p \geq 1.$$

Proof. Set $x = X(\omega)$ and $y = Y(\omega)$ for $\omega \in \Omega$. By the triangle inequality and the fact that for $0 < p < 1$, $x, y > 0$, $(x+y)^p < x^p + y^p$,

$$\mathbb{E}|X+Y|^p \leq \mathbb{E}(|X| + |Y|)^p \leq \mathbb{E}(|X|^p + |Y|^p)$$

which establishes the inequality for $0 < p \leq 1$. For $p \geq 1$, we follow the same proof with the inequality:

$$(x+y)^p \leq 2^{p-1}(x^p + y^p)$$

■

Definition 2.4. Radamacher Functions: Radamacher functions, are functions which when probabilistically interpreted, are a kind of rescaled and iterated coin tossing random variables. Namely, for $t \in \mathbb{R}$, let $r(t)$ be the periodically continued functions defined by

$$r(t) = \begin{cases} 1, & \text{for } 0 \leq t < \frac{1}{2}, \\ -1 & \text{for } \frac{1}{2} \leq t < 1 \end{cases}$$

and set, for $0 \leq t \leq 1, r_n(t) = r(2^{n-1}t), n = 1, 2, \dots$. The sequence $r_n, n \geq 1$ thus defined is the sequence of Radamacher functions. One of the most easily checked properties of the Radamachers is:

$$\int_0^1 r_n(t)dt = 0 \text{ and that } \int_0^1 r_n(t)r_m(t)dt = \begin{cases} 1, & \text{for } m = n, \\ 0, & \text{for } m \neq n, \end{cases}$$

The radamachers form an orthogonal system. They constitute a sequence of independent random variables, assuming the two values +1 and -1 with probability 1/2 each, which implies that the mean is 0, and the variance equals 1.

Lemma 2.5. Khintchine's inequality: Let $p > 0$. let $c_1, c_2, \dots, c_n$ be arbitrary reals, and set

$$f_n(t) = \sum_{k=1}^n c_k r_k(t)$$

Then, there exists constants A_p and B_p depending only on p , such that:

$$A_p \left(\sum_{k=1}^n c_k^2 \right)^{p/2} \leq \int_0^1 |f_n(t)|^p dt \leq B_p \left(\sum_{k=1}^n c_k^2 \right)^{p/2}$$

Proof. Via rescaling it is no loss of generality to assume that $\sum_{k=1}^n c_k^2 = 1$. Moreover, it suffices to prove the lemma for integral values of p . We thus wish to prove that,

for $p = 1, 2, \dots$,

$$A_p \leq \int_0^1 |f_n(t)|^p dt \leq B_p.$$

We first consider f_n :

$$\begin{aligned} \int_0^1 (f_n(t))^p dt &\leq p! \int_0^1 \exp f_n(t) dt \\ &= p! \prod_{k=1}^n \int_0^1 \exp\{c_k r_k(t)\} dt = p! \prod_{k=1}^n \frac{1}{2} (e^{c_k} + e^{-c_k}) \\ &\leq p! \prod_{k=1}^n \exp c_k^2 = p! \exp \sum_{k=1}^n c_k^2 = p! e \end{aligned}$$

In this argument, we have used the fact that one term in the Taylor expansion of the exponential function is smaller than the function itself in the first inequality, independence in the second and finally, the fact that $\frac{1}{2}(e^x + e^{-x}) \leq \exp x^2$ (use Taylor expansion; the odd terms cancel and the even ones double).

By symmetry,

$$\int_0^1 |f_n(t)|^p dt = 2 \int_0^1 f_n(t) I\{f_n(t) > 0\} dt \leq 2p! \int_0^1 e^{f_n(t)} dt \leq 2p! e.$$

This establishes the upper bound.

We now establish the lower bound: For $p \geq 2$, we note that, by Lyapounov's inequality,

$$\left(\int_0^1 |f_n(t)|^p dt \right)^{1/p} \geq \left(\int_0^1 |f_n(t)|^2 dt \right)^{1/2} = \left(\sum_{k=1}^n c_k^2 \right) = 1. \quad (2.1)$$

Now, let $0 < p < 2$, and set $\gamma(p) = p \log(\int_0^1 |f_n(t)|^{1/p} dt)$. This is a convex function.

Let $\alpha = \frac{p}{4-p}$, so that

$$\alpha \frac{1}{p} + (1 - \alpha) \frac{1}{4} = \frac{1}{2}.$$

The convexity of γ yields,

$$\frac{1}{2} \log \left(\int_0^1 |f_n(t)|^2 dt \right) \leq \frac{1}{4-p} \log \left(\int_0^1 |f_n(t)|^p dt \right) + \frac{2-p}{8-2p} \log \left(\int_0^1 |f_n(t)|^4 dt \right).$$

By (2.1) and exponentiating on both sides, it follows that

$$1 \leq \left(\int_0^1 |f_n(t)|^p dt \right)^{\frac{1}{4-p}} \cdot \left(\int_0^1 |f_n(t)|^4 dt \right)^{(2-p)/(8-2p)}.$$

which shows that

$$\int_0^1 |f_n(t)|^p dt \geq A_p \text{ with } A_p = \left(\int_0^1 |f_n(t)|^4 dt \right)$$

. ■

2.2 Marcinkiewicz-Zygmund inequalities

Let $X_1, X_2, \dots, X_n$ be independent random variables with mean 0, and suppose that $\mathbb{E}|X_k|^p < \infty$, for $k = 1, 2, \dots, n$ and some $p \geq 1$. Set $S_n = \sum_{k=1}^n X_k$ and let, random variables with superscript s , denote symmetrized versions. Further, set,

$$T_n(t) = \sum_{k=1}^n X_k r_k(t), \text{ for } 0 \leq t \leq 1, n \geq 1,$$

where $r_k(t)$ is the sequence of radamacher functions.

Theorem 2.6. *Marcinkiewicz-Zygmund inequalities:*

Let $p \geq 1$. Suppose that $X_1, X_2, \dots, X_n$ are independent random variables with mean 0, such that $\mathbb{E}|X_k|^p < \infty$, for all k , and let $\{S_n, n \geq 1\}$, denote the partial sums. Then there exists constants, A_p^ and B_p^* , depending only on p , such that,*

$$A_p^* \mathbb{E} \left(\sum_{k=1}^n X_k^2 \right)^{p/2} \leq \mathbb{E}|S_n|^p \leq B_p^* \mathbb{E} \left(\sum_{k=1}^n X_k^2 \right)^{p/2}$$

Proof. Consider $S_n^s = \sum_{k=1}^n X_k^s$ and $T_n^s(t) = \sum_{k=1}^n X_k r_k(t)$. By the coin tossing property of the Radamacher functions and symmetry, it follows that $S_n^s \stackrel{d}{=} T_n^s(t)$. Invoking prop.(2.1) and the c_p -inequality, we obtain

$$\mathbb{E}|S_n|^p \leq \mathbb{E}|S_n^s|^p = \mathbb{E}|T_n^s(t)|^p \leq 2^p \mathbb{E}|T_n(t)|^p. \quad (2.2)$$

Integrating the extreme members, changing the order of integration and applying the Khintchine's lemma (for fixed ω) yields

$$\mathbb{E}|S_n|^p \leq 2^p \mathbb{E} \int_0^1 |T_n(t)|^p dt \leq 2^p B_p \left(\mathbb{E} \sum_{k=1}^n X_k^2 \right)^{p/2},$$

which proves the upper bound, with $B_p^* = 2^p B_p$, where B_p is the constant in Khintchine's inequality. By running (2.2) backward, with S_n and $T_n(t)$ playing the reversed roles, we obtain

$$\mathbb{E}|T_n(t)|^p \leq 2^p \mathbb{E}|S_n|^p$$

which, after the same intergration procedure and the lower bound of Khintchine's inequality, proves the lower bound with $A_p^* = 2^{-p} A_p$, where A_p is the constant in Khintchine's inequality. ■

If, in addition, $X_1, X_2, \dots, X_n$ are identically distributed, the upper bound can be further elaborated:

Corollary 2.7. *Let $p \geq 1$. Suppose that $X_1, X_2, \dots, X_n$ are independent, identically distributed (i.i.d) random variables, with mean 0 and $\mathbb{E}|X|^p < \infty$. Set $S_n = \sum_{k=1}^n X_k, n \geq 1$. Then, there exists a constant b_p depending only on p , such that*

$$\mathbb{E}|S_n|^p \leq \begin{cases} B_p n \mathbb{E}|X|^p & \text{when } 1 \leq p \leq 2, \\ B_p n^{p/2} \mathbb{E}|X|^p & \text{when } p \geq 2 \end{cases}$$

Proof. Let $1 \leq p \leq 2$. Since, $p/2 < 1$, we apply the c_p -inequality to the Marcinkiewicz-Zygmund inequalities:

$$\mathbb{E}|S_n|^p \leq B_p^* \mathbb{E} \left(\sum_{k=1}^n X_k^2 \right)^{p/2} \leq B_p^* \mathbb{E} \left(\sum_{k=1}^n (X_k^2)^{p/2} \right) = B_p^* n \mathbb{E}|X|^p.$$

For $p \geq 2$, we use the convexity of $|x|^{p/2}$:

$$\begin{aligned} \mathbb{E}|S_n|^p &\leq B_p^* \mathbb{E} \left(\sum_{k=1}^n X_k^2 \right)^{p/2} = B_p^* n^{p/2} \mathbb{E} \left(\frac{1}{n} \sum_{k=1}^n X_k^2 \right)^{p/2} \\ &\leq B_p^* n^{p/2} \frac{1}{n} \left(\sum_{k=1}^n \mathbb{E}(X_k^2)^{p/2} \right) = B_p^* n^{p/2} \mathbb{E}|X|^p \end{aligned}$$

■

2.3 Rosenthal's Inequality

Lemma 2.8. *Let $p \geq 1$. Suppose that $X_1, X_2, \dots, X_n$ are independent random variables such that $\mathbb{E}|X_k|^p < \infty$, for all k . Then,*

$$\mathbb{E}|S_n|^p \leq \max \left\{ 2^p \sum_{k=1}^n \mathbb{E}|X_k|^p, 2^{p^2} \left(\sum_{k=1}^n \mathbb{E}|X_k| \right)^p \right\}.$$

Proof. Since, $S_n \leq \sum_{k=1}^n |X_k|$, it is no restriction to assume that all summands are non-negative. Set $S_n^{(j)} = \sum_{k=1, k \neq j}^n X_k$. Using the c_p -inequality, independence and non-negativity and Lyapounov's inequality, we obtain

$$\begin{aligned} \mathbb{E}(S_n)^p &= \sum_{j=1}^n \mathbb{E}(S_n)^{p-1} X_j \leq 2^{p-1} \sum_{j=1}^n \mathbb{E}((X_j^{p-1} + (S_n^{(j)})^{p-1}) X_j) \\ &= 2^{p-1} \sum_{j=1}^n (\mathbb{E} X_j^p + \mathbb{E}(S_n^{(j)})^{p-1} \cdot \mathbb{E} X_j) \\ &\leq 2^{p-1} \sum_{j=1}^n (\mathbb{E} X_j^p + \mathbb{E}(S_n)^{p-1} \cdot \mathbb{E} X_j) \end{aligned}$$

$$\begin{aligned}
&= 2^{p-1} \left(\sum_{j=1}^n \mathbb{E} X_j^p + \mathbb{E}(S_n)^{p-1} \sum_{j=1}^n \mathbb{E} X_j \right) \\
&\leq 2^{p-1} \left(\sum_{j=1}^n \mathbb{E} X_j^p + (\mathbb{E}(S_n)^p)^{(p-1)/p} \sum_{j=1}^n \mathbb{E} X_j \right) \\
&\leq 2^p \max \left\{ \sum_{j=1}^n \mathbb{E} X_j^p + (\mathbb{E}(S_n)^p)^{(p-1)/p} \sum_{j=1}^n \mathbb{E} X_j \right\}
\end{aligned}$$

Thus,

$$\mathbb{E}(S_n)^p \leq 2^p \sum_{j=1}^n \mathbb{E} X_j^p \text{ and } \mathbb{E}(S_n)^p \leq 2^p (\mathbb{E}(S_n)^p)^{(p-1)/p} \sum_{j=1}^n \mathbb{E} X_j.$$

The conclusion follows. ■

Theorem 2.9. Rosenthal's inequality For $p > 2$, let $X_1, X_2, \dots, X_n$ be independent random variables with mean 0, and suppose that $\mathbb{E}|X_k|^p < \infty$, for all k . Then,

$$\mathbb{E}|S_n|^p \begin{cases} \leq D_p \max \left\{ \sum_{k=1}^n \mathbb{E}|X_k|^p, \left(\sum_{k=1}^n \mathbb{E} X_k^2 \right)^{p/2} \right\} \\ \geq 2^{-p} \max \left\{ \sum_{k=1}^n \mathbb{E}|X_k|^p, \left(\sum_{k=1}^n \mathbb{E} X_k^2 \right)^{p/2} \right\} \end{cases}$$

Proof. Applying the Marcinkiewicz-Zygmund inequalities and lemma (2.8) yields

$$\begin{aligned}
\mathbb{E}|S_n|^p &\leq B_p^* \mathbb{E} \left(\sum_{k=1}^n X_k^2 \right)^{p/2} \\
&\leq B_p^* \max \left\{ 2^{p/2} \sum_{k=1}^n \mathbb{E}((X_k^2)^{p/2}), 2^{(p/2)^2} \left(\sum_{k=1}^n \mathbb{E} X_k^2 \right)^{p/2} \right\} \\
&= D_p \max \left\{ \sum_{k=1}^n \mathbb{E}((X_k^2)^p), \left(\sum_{k=1}^n \mathbb{E} X_k^2 \right)^{p/2} \right\},
\end{aligned}$$

with $D_p = B_p^* 2^{(p/2)^2}$. As for the lower bound, the Marcinkiewicz-Zygmund inequalities tell us that

$$\mathbb{E}|S_n|^p \geq 2^{-p} \left(\sum_{k=1}^n \mathbb{E} X_k^2 \right)^{\frac{p}{2}},$$

and the following result, if $X_1, X_2, \dots, X_n$ are independent random variables with mean 0 and finite moments of order r , then $\sum_{k=1}^n \mathbb{E}|X_k|^r \leq 2^r \mathbb{E}|\sum_{k=1}^n X_k|^r$, tells us

$$\mathbb{E}|S_n|^p \geq 2^{-p} \sum_{k=1}^n \mathbb{E}|X_k|^p,$$

which together establish the lower bound. ■

2.4 Brunn Minkowski Inequality

Theorem 2.10. Prekopa-Leindler Inequality *Let $f, g, m : \mathbb{R}^n \rightarrow [0, \infty)$ be integrable functions and take $\lambda \in (0, 1)$. Assume that for all $x, y \in \mathbb{R}^n$, we have the following inequality:*

$$m(\lambda x + (1 - \lambda)y) \geq f(x)^\lambda g(y)^{(1-\lambda)}$$

Then,

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

Proof. Let μ_n denote the n -dimensional Lebesgue measure. We will prove the result by induction on n . For the case $n = 1$, we use the fact that if $A, B \subset \mathbb{R}$, are nonempty and measurable, then

$$\mu_1(A + B) \geq \mu_1(A) + \mu_1(B)$$

To prove this, it suffices to prove this for compact sets by set approximation results. Also, we are allowed to translate A, B , as we please, since μ_1 is translation invariant. Thus, move A , so that the max value is 0 and move B , so that the min. value is 0, then $A \cap B = \{0\}$. Now, note that since $0 \in A \cap B$, $A + B \supset A \cup B$, and this gives

the result since

$$\mu_1(A + B) \geq \mu_1(A \cup B) = m_1(A) + \mu_1(B)$$

as A, B , overlap on a set of measure zero 0. Continuing the proof by approximation, it suffices to prove the result for bounded f, g . By scaling, we may assume that $\|f\|_\infty = \|g\|_\infty = 1$. Set some $t \in (0, 1)$. Let $A = \{x : f(x) \geq t\}, B = \{x : g(x) \geq t\}$, and $C = \{x : m(x) \geq t\}$. Now if $x \in A, y \in B$, then

$$m(\lambda x + (1 - \lambda)y) \geq f(x)^\lambda g(y)^{1-\lambda} \geq t^\lambda t^{1-\lambda} = t$$

so that $\lambda A + (1 - \lambda)B \subset C$. Note that by assumption and since $t < 1, A, B$ are nonempty. Thus applying the earlier result, we have that

$$\begin{aligned} \mu_1(C) &\geq \mu_1(\lambda A + (1 - \lambda)B) \\ &\geq \mu_1(\lambda A) + \mu_1((1 - \lambda)B) \\ &= \lambda m_1(A) + (1 - \lambda)\mu_1(B) \end{aligned}$$

$$\mu_1(m \geq t) \geq \lambda \mu_1(f \geq t) + (1 - \lambda) \mu_1(g \geq t)$$

Integrating over t and using Fubini, we get that

$$\begin{aligned} \int_0^\infty \mu_1(m \geq t) dt &\geq \lambda \int_0^\infty \mu_1(f \geq t) dt + (1 - \lambda) \int_0^\infty \mu_1(g \geq t) dt \\ \int_{\mathbb{R}} m(x) dx &\geq \lambda \int_{\mathbb{R}} f(x) dx + (1 - \lambda) \int_{\mathbb{R}} g(x) dx \\ &\geq \left(\int_{\mathbb{R}} f(x) dx \right)^\lambda \left(\int_{\mathbb{R}} g(x) dx \right)^{1-\lambda} \end{aligned}$$

by the AM-GM inequality.

Note that for $n = 1$ stronger result holds, that we can bound by the arithmetic

mean. This does not hold for larger dimensions. For the induction step, assume the result holds for $n - 1$. We will use $\mathbb{R}^n = \mathbb{R} \times \mathbb{R}^{n-1}$. For $(t, s) \in \mathbb{R} \times \mathbb{R}^{n-1}$, define $f_{(t,s)}, g_t(s) = g(t, s)$ and $m_t(s) = m(t, s)$. Take $t_0, t_1 \in \mathbb{R}$ and define $t = \lambda t_0 + (1 - \lambda)t_1$. If $x, y \in \mathbb{R}^{n-1}$, then

$$\begin{aligned} m_t(\lambda x + (1 - \lambda)y) &= m(\lambda(t, x) + (1 - \lambda)(t, y)) \\ &\geq f(t_0, x)^\lambda g(t, y)^{1-\lambda} \\ &\geq f_{t_0}(x)^\lambda g_{t_1}(y)^{1-\lambda} \end{aligned}$$

and thus the triple f_{t_0}, g_{t_1} and m_t satisfy the assumptions for $n - 1$. Induction gives

$$\int_{\mathbb{R}^{n-1}} m_t(x) d\mu_{n-1}(x) \geq \left(\int_{\mathbb{R}^{n-1}} f_{t_0}(x) d\mu_{n-1}(x) \right)^\lambda \left(\int_{\mathbb{R}^{n-1}} g_{t_1}(x) d\mu_{n-1}(x) \right)^{1-\lambda}$$

If we denote

$$\begin{aligned} \tilde{m}(t) &= \int_{\mathbb{R}^{n-1}} m_t(x) d\mu_{n-1}(x) \\ \tilde{f}(t) &= \int_{\mathbb{R}^{n-1}} f_t(x) d\mu_{n-1}(x) \\ \tilde{g}(t) &= \int_{\mathbb{R}^{n-1}} g_t(x) d\mu_{n-1}(x) \end{aligned}$$

Then, the above says that

$$\tilde{m}(\lambda t_0 + (1 - \lambda)t_1) \geq \tilde{f}(t_0)^\lambda \tilde{g}(t_1)^{1-\lambda}$$

Thus, the triple $\tilde{m}, \tilde{f}$ and $\tilde{g}$, satisfy the hypothesis for the case $n = 1$ and we have that

$$\int_{\mathbb{R}} \tilde{m}(t) d\mu_1(t) \geq \left(\int_{\mathbb{R}} \tilde{f}(t) d\mu_1(t) \right)^\lambda \left(\int_{\mathbb{R}} \tilde{g}(t) d\mu_1(t) \right)^{1-\lambda}$$

and this is exactly the result we want. ■

Corollary 2.11. Brunn Minkowski inequality

Let $A, B \subset \mathbb{R}^n$ be measurable and non empty. Then,

1. For all $\lambda \in [0, 1]$, $\text{vol}(\lambda A + (1 - \lambda)B) \geq \text{vol}(A)^\lambda \text{vol}(B)^{1-\lambda}$
2. $\text{vol}(A + B)^{1/n} \geq \text{vol}(A)^{1/n} + \text{vol}(B)^{1/n}$

Proof. For (1), we apply the Prekopa-Leindler for $f = \mathbf{1}_A, g = \mathbf{1}_B, m = \mathbf{1}_{\lambda A + (1-\lambda)B}$, so that

$$m(\lambda x + (1 - \lambda)y) \geq f(x)^\lambda g(y)^{1-\lambda}$$

since, if $x \in A$ and $y \in B$, then $\lambda x + (1 - \lambda)y \in \lambda A + (1 - \lambda)B$, and otherwise, the inequality holds trivially. Then, the inequality gives

$$\text{vol}(\lambda A + (1 - \lambda)B) = \int_{\mathbb{R}} \mathbf{1}_{\lambda A + (1 - \lambda)B} dx \geq \left(\int_{\mathbb{R}} \mathbf{1}_A dx \right)^\lambda \left(\int_{\mathbb{R}} \mathbf{1}_B dx \right)^{1-\lambda} = \text{vol}(A)^\lambda \text{vol}(B)^{1-\lambda}$$

(2) is a consequence of (1). First, we normalize the sets so that $A = \frac{\tilde{A}}{(\text{vol}(A))^{1/n}}, B = \frac{\tilde{B}}{(\text{vol}(B))^{1/n}}$. Then, we see that

$$\text{vol} \left(\frac{\text{vol}(A)^{\frac{1}{n}}}{\text{vol}(A)^{1/n} + \text{vol}(B)^{1/n}} \tilde{A} + \frac{\text{vol}(B)^{\frac{1}{n}}}{\text{vol}(A)^{1/n} + \text{vol}(B)^{1/n}} \tilde{B} \right) \geq 1$$

(Note that $\text{vol}(\tilde{A}) = 1, \text{vol}(\tilde{B}) = 1$). But then,

$$\frac{\text{vol}(A)^{\frac{1}{n}}}{\text{vol}(A)^{1/n} + \text{vol}(B)^{1/n}} \tilde{A} + \frac{\text{vol}(B)^{\frac{1}{n}}}{\text{vol}(A)^{1/n} + \text{vol}(B)^{1/n}} \tilde{B} = \frac{A + B}{\text{vol}(A)^{1/n} + \text{vol}(B)^{1/n}}$$

And, thus, we have

$$\frac{\text{vol}(A + B)}{(\text{vol}(A)^{1/n} + \text{vol}(B)^{1/n})^n} \geq 1$$

Rearranging gives the desired result.

■

2.5 Concentration of measure on the sphere

Let $(X, \|\cdot\|)$ be a normed space, and $K = \{x \in X : \|x\| \leq 1\}$.

Definition 2.12. The modulus of uniform convexity for the normed space $(X, \|\cdot\|)$ is

$$\delta_{\|\cdot\|}(\varepsilon) = \inf \left\{ 1 - \left\| \frac{x+y}{2} \right\| : \|x\|, \|y\| \leq 1, \|x-y\| \geq \varepsilon \right\}$$

Example 2.13. If X is a Hilbert space, then we have the parallelogram identity:

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

and if $\|x\| = \|y\| = 1$, and, $\|x-y\| \geq \varepsilon$, then

$$\|x+y\|^2 + \varepsilon^2 \leq 4$$

so that $\left\| \frac{x+y}{2} \right\| \leq \sqrt{1 - \frac{\varepsilon^2}{4}} \approx 1 - \frac{\varepsilon^2}{8}$

Remark 2.14. It is non trivial result to show that if for a normed space, $\delta(\varepsilon) > \frac{\varepsilon^2}{8}$, then it must be a Hilbert space. In other words, Hilbert spaces have the best modulus of uniform convexity.

Now, let K be a centrally symmetric convex body, and let $\delta K = S$, where δK is the boundary of K .

Define

$$\nu(A) = \frac{\text{vol}(A \cap K)}{\text{vol}(K)}, A \subset K$$

$$\mu(A) = \frac{\text{vol}(\{ta : a \in A, t \in (0, 1)\})}{\text{vol}(K)}, A \subset S$$

If $K = B_2^n$, then μ is the normalised surface measure.

Theorem 2.15. *Let $K \subset \mathbb{R}^n$ be centrally symmetric convex body, and $S = \delta K$. Let*

$\delta = \delta_{\|\cdot\|}$. Then,

$$1. \text{ For all } A \subset K, \nu(A_\varepsilon) \geq 1 - \frac{1}{\nu(A)} e^{-2n\delta(\varepsilon)}$$

$$2. \text{ For all } A \subset S, \mu(A_\varepsilon) \geq 1 - \frac{2}{\mu(A)} e^{-2n\delta(\frac{\varepsilon}{4})}$$

Here, $A_\varepsilon = A + \varepsilon K = \{x \in \mathbb{R}^n : d(x, A) \leq \varepsilon \text{ and } d(x, y) = \|x - y\|\}$.

Proof. (1) Let $A \subset K$. Define $B = K \setminus A_\varepsilon$. If $a \in A$ and $b \in B$, then $\|a - b\| \geq \varepsilon$, and

$$\left\| \frac{a+b}{2} \right\| \leq 1 - \delta(\varepsilon)$$

This shows that

$$\frac{A+B}{2} \subset (1 - \delta(\varepsilon))K$$

and by Brunn-Minkowski,

$$\nu(A)^{1/2} \nu(B)^{1/2} \leq \nu\left(\frac{A+B}{2}\right) \leq \nu(1 - \delta(\varepsilon)K) = (1 - \delta(\varepsilon))^n \leq e^{-n\delta(\varepsilon)}$$

This implies that

$$1 - \nu(A_\varepsilon) = \nu(B) \leq \frac{1}{\nu(A)} e^{-2n\delta(\varepsilon)}$$

(2) Now let's take $A \subset S$, and again, use $B = S \setminus A_\varepsilon$. Define the partial cones

$$\hat{A} = \left\{ ta : a \in A, \frac{1}{2} \leq t \leq 1 \right\}, \hat{B} = \left\{ tb : b \in B, \frac{1}{2} \leq t \leq 1 \right\}$$

Claim 2.16. If $\hat{a} \in \hat{A}$, and $\hat{b} \in \hat{B}$, then $\|\hat{a} - \hat{b}\| \geq \frac{\varepsilon}{4}$.

Now, let $\hat{a} = \alpha a$ and $\hat{b} = \beta b$, so that $\alpha, \beta \in [1/2, 1]$ and $a \in A, b \in B$. We split this into two cases:

- If $|\alpha - \beta| \geq \frac{3}{4}$, then we are done by triangle inequality

$$\|\hat{a} - \hat{b}\| \geq |\alpha - \beta| \geq \frac{\varepsilon}{4}$$

- Otherwise, $|\alpha - \beta| \leq \frac{\varepsilon}{4}$, and we have

$$\|\hat{a} - \hat{b}\| = \|\alpha(a - b) + (\alpha - \beta)b\| \geq \|\alpha(a - b)\| - \|(\alpha - \beta)b\| \geq \frac{1}{2}\varepsilon - \frac{\varepsilon}{4} = \frac{\varepsilon}{4}$$

This implies that $\frac{\hat{A} + \hat{B}}{2} \subset (1 - \delta(\frac{\varepsilon}{4}))K$, and with the same computation as before, we get

$$\nu(\hat{B}) \leq \frac{1}{\nu(\hat{A})} e^{-2n\delta(\varepsilon/4)}$$

To finish the result, we note that

$$\mu(A) = \nu([0, 1]A)$$

$$\begin{aligned} \nu(\hat{A}) &= \nu([0, 1]A \setminus [0, 1/2]A) = \mu(A) - \nu(\frac{1}{2}[0, 1]A) \\ &= (1 - \frac{1}{2^n})\mu(A) \end{aligned}$$

and, this means that

$$(1 - \frac{1}{2^n})^2 \mu(B) \leq \frac{1}{\mu(A)} e^{-2n\delta(\varepsilon/4)}$$

Since, $B = S \setminus A_\varepsilon$, we have that

$$\mu(A_\varepsilon) \geq 1 - \frac{1}{\mu(A)} e^{-2n\delta(\varepsilon/4)} (1 - \frac{1}{2^n})^{-2} \approx 1 - \frac{1}{\mu(A)} e^{-2n\delta(\varepsilon/4)}$$

■

Now, with μ being the normalised Haar measure on $\mathbb{S}^{n-1}$, we have the following corollary:

Corollary 2.17. *Let $f : \mathbb{S}^{n-1} \rightarrow \mathbb{R}$, be an L -Lipschitz map, i.e. for all $x, y \in \mathbb{S}^{n-1}$, we have that*

$$|f(x) - f(y)| \leq L\|x - y\|_2$$

Let M be a median of f , i.e. $\mu(f \geq M), \mu(f \leq M) \geq \frac{1}{2}$ (which always exists for any probability measure). Then, for all $\varepsilon > 0$,

$$\mu(x \in \mathbb{S}^{n-1} : |f(x) - M| \geq \varepsilon) \leq Ce^{-c' n \varepsilon^2 / L^2}$$

This says that a Lipschitz function is essentially a constant on $\mathbb{S}^{n-1}$ for large n .

Proof. Let $A = \{x \in \mathbb{S}^{n-1} : f(x) \leq M\}$. Then, $x \in A_{\varepsilon/L}$ if and only if there exists $y \in A$, with $\|x - y\| \leq \varepsilon/L$, in which case $|f(x) - f(y)| \leq \varepsilon$ and $f(x) \leq M + \varepsilon$. This means that

$$\{x \in \mathbb{S}^{n-1} : f(x) > M + \varepsilon\} \subset \mathbb{S}^{n-1} \setminus A_{\varepsilon/L}$$

Since, $\mu(A) \geq \frac{1}{2}$, we have

$$\mu(x \in \mathbb{S}^{n-1} : f(x) - M > \varepsilon) \leq Ce^{-c' n \varepsilon^2 / L^2}$$

By symmetry, examining $B = \{x \in \mathbb{S}^{n-1} : f(x) \geq M\}$, we have the opposite inequality, that

$$\mu(x \in \mathbb{S}^{n-1} : f(x) - M < -\varepsilon) \leq Ce^{-c' n \varepsilon^2 / L^2}$$

and since the union of these sets is $\mu(x \in \mathbb{S}^{n-1} : \|f(x) - M\| > \varepsilon)$, the result follows by subadditivity. ■

Corollary 2.18. Gaussian Concentration:

Let $f : \mathbb{S}^{n-1} \rightarrow \mathbb{R}$, be an L -Lipschitz map, i.e. for all $x, y \in \mathbb{S}^{n-1}$, we have that

$$\mu(x \in \mathbb{S}^{n-1} : |f(x) - \int f d\mu| \geq \varepsilon) \leq Ce^{-c' n \varepsilon^2 / L^2}$$

Proof. Without loss of generality, $L = 1$ (by scaling). Denote

$$\mathbb{E}(f) := \int_{\mathbb{S}^{n-1}} f d\mu$$

Take the product measure $\mu \times \mu$ on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$. Note

$$\begin{aligned} (\mu \times \mu)((x, y) \in \mathbb{S}^{n-1} \times \mathbb{S}^{n-1} : |f(x) - f(y)| \geq \varepsilon) &\leq (\mu \times \mu) \left(\left\{ |f(x) - M| \geq \frac{\varepsilon}{2} \right\} \cup \left\{ |f(y) - M| \geq \frac{\varepsilon}{2} \right\} \right) \\ &\leq 2\mu \left(\left\{ |f(x) - M| \geq \frac{\varepsilon}{2} \right\} \right) \\ &\leq 2ae^{bn\varepsilon^2/4} \end{aligned}$$

This means, for most pair of points, $f(x)$ and $f(y)$ are close. Now, fix $\lambda > 0$, we will be using a Chernoff bound:

$$\mu(|f(x) - \mathbb{E}(f)| > \varepsilon) = \mu(e^{\lambda^2|f(x) - \mathbb{E}(f)|^2 \geq e^{\lambda^2\varepsilon^2}}) \leq e^{-\lambda^2|f(x) - \mathbb{E}(f)|^2} d\mu(x)$$

Note that since $e^{\lambda^2 t^2}$ is convex, Jensen's inequality gives

$$e^{\lambda^2|f(x) - \int f(y) d\mu(y)|^2} \leq \int_{\mathbb{S}^{n-1}} e^{\lambda^2|f(x) - f(y)|^2} d\mu(y)$$

and, hence,

$$\int_{\mathbb{S}^{n-1}} e^{\lambda^2|f(x) - \mathbb{E}|^2} d\mu(x) \leq \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} e^{\lambda^2|f(x) - f(y)|^2} d\mu(x) d\mu(y)$$

Now, using Fubini, note that if ψ is an increasing function from $\mathbb{R}^+ \rightarrow \mathbb{R}^+$, then

$$\mathbb{E}[\psi] = \int_0^\infty \psi'(t) \mathbb{P}(X \geq t) dt$$

Thus,

$$\begin{aligned} \int_{\mathbb{S}^{n-1}} \int_{\mathbb{S}^{n-1}} e^{\lambda^2|f(x) - f(y)|^2} d\mu(x) d\mu(y) &\leq \int_0^\infty 2\lambda^2 t e^{\lambda^2 t} (\mu \times \mu)(\{(x, y) : |f(x) - f(y)| \geq t\}) dt \\ &\leq \int_0^\infty 2\lambda^2 t e^{\lambda^2 t} 2ae^{-bnt^2/4} dt \end{aligned}$$

$$\leq \int_0^\infty \frac{bn}{8} 2te^{-bnt^2/8} dt = 2a, \text{ setting } \left(\lambda^2 = \frac{bn}{8} \right)$$

Putting everything together, we conclude that

$$\mu(|f(x) - \mathbb{E}(f)| > \varepsilon) \leq 2ae^{-bn\varepsilon^2/8}$$

■

Chapter 3

Gaussian Concentration and its consequences

3.1 Preliminaries

Lemma 3.1. (*Gaussian Concentration*) Suppose that g_n are iid $\mathcal{N}(0, 1)$ random vari-

ables and, $G = \sup_i (l_i + \sum_n k_{i,n} g_n)$,

Then,

$$\mathbb{P}(G \geq (EG^2)^{1/2} + wt) \leq \mathbb{P}(G \geq EG + wt) \leq e^{-t^2/2},$$

where $w = \sup_i \sqrt{\sum_n k_{i,n}^2}$.

Proof. We may treat G as a w -Lipschitz function of the variables g_n and the statement is an immediate consequence of concentration of Gaussian measures, corollary (2.18).

■

Lemma 3.2. Suppose that random variables Y_x, A_x and numbers $b_x \geq 0$, satisfy the inequality

$$\mathbb{P}(Y_x \geq A_x + tb_x) \leq e^{-t}, \forall t > 0, x \in U.$$

Then, for all $u \geq -\log \#U$,

$$\mathbb{E} \max_{x \in U} [Y_x - A_x - b_x(u + \log \#U)] \leq e^{-u} \max_{x \in U} b_x. \quad (3.1)$$

In particular,

$$\mathbb{E} \max_{x \in U} [Y_x - A_x - b_x \log \#U] \leq \max_{x \in U} b_x. \quad (3.2)$$

Proof. Let $Z_x = Y_x - A_x - b_x(u + \log \#U)$, then,

$$\begin{aligned} \mathbb{E} \max_{x \in U} Z_x &\leq \int_0^\infty \mathbb{P}(\max_{x \in U} Z_x \geq s) ds \leq \sum_{x \in U} \int_0^\infty \mathbb{P}(Y_x \geq A_x + b_x(u + \log \#U + \frac{s}{b_x})) ds \\ &\leq \sum_{x \in U} e^{-U} \frac{1}{\#U} \int_0^\infty e^{-\frac{s}{b_x}} ds = e^{-u} \frac{1}{\#U} \sum_{x \in U} b_x \leq e^{-u} \max_{x \in U} b_x \end{aligned}$$

and (3.2) follows. ■

Lemma 3.3. Suppose that $g_{i,j}$ are iid $\mathcal{N}(0, 1)$ r.v., U is a finite subset of l_2 and for each $x \in U$, there are given sets $I(x)$ and r.v.s, $(L_j(x))_j$ independent of $(g_{ij})_{i \in I(x), j}$. Then, for any $c > 0$,

$$\begin{aligned} &\mathbb{E} \max_{x \in U} [\sum_j (L_j(x) + \sum_{i \in I(x)} a_{ij} g_{ij} x_i)^2 - (1+c) \sum_j (L_j(x))^2 + \sum_{i \in I(x)} a_{ij}^2 x_i^2) \\ &\quad - 2(1+c^{-1}) \max_j \sum_{i \in I(x)} a_{ij}^2 x_i^2 \log \#U] \\ &\leq 2(1+c^{-1}) \max_{x \in U} \max_j \sum_{i \in I(x)} a_{ij}^2 x_i^2. \end{aligned}$$

Proof. Let

$$G_x = \sqrt{\sum_j (L_j(x) + \sum_{i \in I(x)} a_{ij} g_{ij} x_i)^2} = \sup_{\|y\|_2 \leq 1} \sum_j (L_j(x) + \sum_{i \in I(x)} a_{ij} g_{ij} x_i) y_j$$

Then, $\mathbb{E}_g G_x^2 = \sum_j (L_j(x))^2 + \sum_{i \in I(x)} a_{ij}^2 x_i^2$, where, $\mathbb{E}_g$ denotes the expectation with respect to the gaussian random variables, $(g_{ij})_{i \in I(x), j}$.

If we put,

$$w(x) = \sup_{\|y\|_2 \leq 1} \sqrt{\sum_j \sum_{i \in I(x)} a_{ij}^2 x_i^2 y_j^2} = \max_j \sqrt{\sum_{i \in I(x)} a_{ij}^2 x_i^2}$$

Then since,

$$(\sqrt{\mathbb{E}_g G_x^2} + tw(x))^2 \leq (1+c)\mathbb{E}_g G_x^2 + (1+c^{-1})t^2 w^2(x)$$

we get, by Lemma (3.1)

$$\begin{aligned} \mathbb{P}(G_x^2 \geq (1+c)\mathbb{E}_g G_x^2 + (1+c^{-1})t^2 w^2(x)) \\ \leq \mathbb{P}(G_x \geq \sqrt{\mathbb{E}_g G_x^2} + tw(x)) \leq e^{-t^2/2} \end{aligned}$$

Therefore, by Lemma (3.3) follows by Lemma(3.2) applied to

$$Y_x = G_x^2, A_x = (1+c) \sum_j (L_j(x)^2 + \sum_{i \in I(x)} a_{ij}^2 x_i^2), b_x = 2(1+c^{-1})w^2(x)$$

■

3.2 Computing the norm on a subset of the unit ball of l_2

We introduce a subset of l_2 , on which the p -norm can be computed easily and is proportional to the p -norm computed on the unit sphere of l_2 . We do this so that we have explicit control over the cardinality of this set and the infinity-norm on sections of this subset.

Consider the subset of l_2 , D_2 , defined by,

$$D_2 = \{x \in l_2 : \|x\|_2 \leq 1, \forall i, \exists k \in \mathbb{Z}, \text{ such that } x_i^2 = 2^{-k} I_{x_i \neq 0}\},$$

and its finite dimensional counterpart:

$$D_2^n = \{x \in l_2^{2^n} : \|x\|_2 \leq 1, \forall i, \exists k \in \mathbb{Z}, \text{ such that } x_i^2 = 2^{-k} I_{x_i \neq 0}\},$$

Remark 3.4. With a little abuse of notation, we will treat $l_2^{2^n}$, as a subset of l_2 and D_2^n , as subset of D_2 .

For $x \in D_2, k \in \mathbb{Z}_+$, define,

$$A_k(x) = \{i : i \geq 2^k, x_i^2 \geq 2^{-k} \text{ or } 2^l \leq i < 2^{l+1}, x_i^2 \geq 2^{l-2k}, l = 0, 1, \dots, k-1\}$$

$A_k(x)$ captures those coordinates of x , whose content are in a specific range. We use these coordinates to get the values residing in these coordinates, by defining the set, Π_k^n .

Let $\pi_k(x) = (\pi_{k,i}(x))_i$, where $\pi_{k,i}(x) = x_i I_{A_k(x)}(i)$.

Now, set $\Pi_k^n = \{\pi_k(x) : x \in D_2^n\}$. Having defined this set, we obtain an upper bound on its cardinality.

Lemma 3.5. For any $k, n \in \mathbb{Z}_+$, we

$$\log \# \Pi_k^n \leq \begin{cases} C 2^k (1 + n - k) & \text{for } k \leq n \\ C 2^n (1 + k - n) & \text{for } k \geq n \end{cases}$$

Proof. Let

$$U_l^n = \{x \in l_2^{2^n}, \|x\|_2 \leq 1 : \forall i, x_i^2 = 2^{-l} I_{x_i \neq 0}\}$$

and

$$V_l^n = \{x \in l_2^{2^n}, \|x\|_2 \leq 1 : \forall i, \exists m = 0, 1, \dots, l, x_i^2 = 2^{-m} I_{x_i \neq 0}\}.$$

We shall now get an upperbound on the cardinality of U_l^n and use that to bound the cardinality of V_l^n .

Notice that if $x \in U_l^n$, then, $\#\{i : x_i \neq 0\} \leq 2^l$, so, for $l \leq n$,

$$\#U_l^n \leq \sum_{k=0}^{2^l} \binom{2^n}{k} 2^k \leq 1 + \sum_{k=1}^{2^l} \left(\frac{2e2^n}{k}\right)^k \leq 1 + 2^l (2e2^{n-l})^{2^l} \leq \exp(C 2^l (1 + n - l))$$

where the second inequality follows by the classical estimate $\binom{n}{k} \leq \frac{n}{k!} \leq \left(\frac{en}{k}\right)^k$ and the third one by the monotonicity of the function $x \rightarrow a^x x^{-x}$ on $(0, \frac{a}{e}]$. ■

For $l \geq n$, we get,

$$\#U_l^n \leq \sum_{k=0}^{2^n} \binom{2^n}{k} 2^k = 3^{2^n}$$

Now notice that for any $a, b > 0$, there exist finite constants, $C_{a,b}$ and $C'_{a,b}$ depending only on a and b such that $\forall p \geq 2$,

$$\sum_{l=-\infty}^m 2^{al} (1 + b(n-l))^{\frac{p}{2}} \leq C_{a,b} 2^{am} (1 + b(n-m))^{\frac{p}{2}} \text{ for } m \leq n \quad (3.3)$$

and

$$\sum_{l=m}^{\infty} 2^{-al} (1 + b(l-n))^{\frac{p}{2}} \leq C'_{a,b} 2^{-am} (1 + b(m-n))^{\frac{p}{2}} \text{ for } m \geq n \quad (3.4)$$

Indeed, observe that for $l \leq m \leq n$, $1 + b(n-l) \leq (1 + b(m-l))(1 + b(n-m))$, so

$$\begin{aligned} \sum_{l=-\infty}^m 2^{al} (1 + b(n-l))^{\frac{p}{2}} &\leq 2^{am} (1 + b(n-m))^{\frac{p}{2}} \sum_{l=-\infty}^m 2^{a(l-m)} (1 + b(m-l))^{\frac{p}{2}} \\ &= 2^{am} (1 + b(n-m))^{\frac{p}{2}} \sum_{k=0}^{\infty} 2^{-ak} (1 + bk)^{\frac{p}{2}} \end{aligned}$$

(3.3) follows. The proof of (3.4) is similar.

Thus, if we use (3.3) with $p = 2, a = b = 1$, and the previously obtained estimates of $\#U_l^n$, we get that for any n, k ,

$$\log \#V_k^n \leq \sum_{l=0}^k \log \#U_l^n \leq \begin{cases} C 2^k (1 + n - k) & \text{for } k \leq n \\ C 2^n (1 + k - n) & \text{for } k \geq n \end{cases}$$

Now, for $k \leq n$, we have

$$\#\Pi_k^n \leq \#V_k^n \prod_{l=0}^{k-1} \#V_{2k-l}^l,$$

so, by (3.3) with $a = 1, b = 2$, we obtain,

$$\log \#\Pi_k^n \leq C[2^k(1+n-k) + \sum_{l=0}^{k-1} 2^l(1+2k-2l)] \leq C2^k(1+n-k).$$

Similarly, for $k \geq n$,

$$\#\Pi_k^n \leq \prod_{l=0}^n \#V_{2k-l}^l$$

and hence,

$$\log \#\Pi_k^n \leq C \sum_{l=0}^n 2^l(1+2k-2l) \leq C2^n(1+k-n).$$

Lemma 3.6. *For any $x \in l_2$ with $\|x\|_2 \leq 1$, we can find $y \in D_2$, such that $\|x-y\|_2 \leq \frac{3}{10}$. Therefore, for any linear operator T on l_2 , and $p \geq 2$,*

$$\sup_{\|x\|_2 \leq 1} \|Tx\|_p \leq \frac{10}{7} \sup_{x \in D_2} \|Tx\|_p$$

Proof. To prove the first part of the statement, it is enough to show that

$$\forall a \in [-1, 1], \exists b, \text{ such that } b^2 = 2^{-k}I_{b \neq 0}, b^2 \leq a^2, |a-b| \leq \frac{3}{10}|a|$$

If we take $a = 0$, we take $b = 0$, we may assume that $2^{-k} \leq a^2 \leq 2^{-k+1}$ and put $b = \text{sgn}(a)2^{\frac{-k}{2}}$, then,

$$\frac{|a-b|}{|a|} = \left(1 - \frac{|b|}{|a|}\right) \leq \left(1 - \frac{1}{\sqrt{2}}\right) \leq \frac{3}{10}$$

Therefore, any $x \in l_2$, with $\|x\|_2 \leq 1$, can be represented as

$$x = \sum_{i \leq 0} \left(\frac{3}{10}\right)^i y_i, \text{ where } y_0, y_1, \dots \in D_2$$

and the last part of the lemma follows immediately from this. ■

Lemma 3.7. Let $\Delta = \max_j (\sum_{i < 2^n} a_{ij}^4)^{\frac{1}{4}}$. Then,

$$\mathbb{E} \sup_{x \in D_2} \left[\sum_j \left(\sum_{i < 2^n} a_{ij} g_{ij} x_i \right)^2 - C_1 \sum_{ij} a_{ij}^2 x_i^2 - C_1 \Delta^2 2^{\frac{n}{2}} \sum_{l=1}^n 2^{\frac{l-n}{8}} \sqrt{\sum_{2^{l-1} \leq i < 2^l} x_i^2} \right] \leq C_1 \Delta^2 2^{\frac{n}{4}}$$

Proof. For $x \in D_2^n$, $l, n = 1, 2, \dots$, let us define

$$A_l^n(x) = \{i < 2^n : i \in A_l(x)\},$$

$$B_0^n(x) = A_0^n(x) = \{i < 2^n : x_i^2 = 1\}, b_0(x) = 2^{\frac{n}{4}} \log \# \Pi_0^n \max_j \sum_{i \in B_0^n(x)} a_{ij}^2 x_i^2,$$

$$B_l^n(x) = A_l^n(x) \setminus A_{l-1}^n(x) = \{i < 2^n : i \geq 2^{l-1}, x_i^2 = 2^{-l}\}$$

$$\text{or } 2^{m-1} \leq i < 2^m, x_i^2 \in \{2^{m-2l}, 2^{m-1-2l}\}, m = 1, 2, \dots, l\},$$

$$c_l(x) = \sqrt{\sum_{2^{l-1} \leq i < 2^l} x_i^2}, b_l(x) = 2^{\frac{|n-l|}{4}} \log \# \Pi_l^n \max_j \sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2$$

and

$$d_k = \prod_{l=0}^k \left(1 + 2^{\frac{-|n-l|}{4}} \right)$$

We prove by induction on $k = 0, 1, \dots$, that

$$\begin{aligned} \mathbb{E} \sup_{x \in D_2^n} \left[\sum_j \left(\sum_{i \in A_k^n(x)} a_{ij} g_{ij} x_i \right)^2 - d_k \sum_{i \in A_k^n(x), j} a_{ij}^2 x_i^2 - 4d_k \sum_{l=0}^k b_l(x) \right] \\ \leq 4d_k \sum_{l=0}^k 2^{\frac{|n-l|}{4}} \sup_{x \in D_2^n} \max_j \sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2. \end{aligned} \quad (3.5)$$

For $k = 0$, inequality (3.5) follows by Lemma(3.3), applied to $U = \Pi_0^n, L_j(x) = 0, I(x) = B_0^n(x)$ and $c = 2^{-\frac{n}{4}}$.

To show the inductive step, let us denote

$$S_k(x) = \sum_j \left(\sum_{i \in A_k^n(x)} a_{ij} g_{ij} x_i \right)^2 - d_k \sum_{i \in A_k^n(x), j} a_{ij}^2 x_i^2 - 4d_k \sum_{l=0}^k b_l(x).$$

Then,

$$\begin{aligned}
& S_k(x) - \left(1 + 2^{\frac{-|n-l|}{4}}\right) S_{k-1}(x) \\
& \leq \sum_j \left(\sum_{i \in A_k^n(x)} a_{ij} g_{ij} x_i \right)^2 - \left(1 + 2^{\frac{-|n-k|}{4}}\right) \sum_j \left(\sum_{i \in A_{k-1}^n(x)} a_{ij} g_{ij} x_i \right)^2 \\
& - \left(1 + 2^{\frac{-|n-k|}{4}}\right) \sum_j \sum_{i \in B_k^n(x)} a_{ij}^2 x_i^2 - 2 \left(1 + 2^{\frac{-|n-k|}{4}}\right) \log \# \Pi_k^n \max_j \sum_{i \in B_k^n(x)} a_{ij}^2 x_i^2.
\end{aligned}$$

Lemma (3.3) applied to

$$U = \Pi_k^n, L_j(x) = \sum_{i \in A_{k-1}^n(x)} a_{ij} g_{ij} x_i, I(x) = B_k^n(x), c = 2^{\frac{-|n-k|}{4}}$$

implies

$$\mathbb{E} \sup_{x \in D_2^n} \left[S_k(x) - \left(1 + 2^{\frac{-|n-k|}{4}}\right) S_{k-1}(x) \right] \leq 2 \left(1 + 2^{\frac{-|n-k|}{4}}\right) \sup_{x \in D_2^n} \max_j \sum_{i \in B_k^n(x)} a_{ij}^2 x_i^2.$$

The induction step from $k-1$ to k easily follows, since

$$\mathbb{E} \sup_{x \in D_2^n} S_k(x) \leq \mathbb{E} \sup_{x \in D_2^n} \left[S_k(x) - \left(1 + 2^{\frac{-|n-k|}{4}}\right) S_{k-1}(x) \right] + \left(1 + 2^{\frac{-|n-k|}{4}}\right) \mathbb{E} \sup_{x \in D_2^n} S_{k-1}(x).$$

When we take $k \rightarrow \infty$, in (3.5), we get

$$\begin{aligned}
& \mathbb{E} \sup_{x \in D_2^n} \left[\sum_j \left(\sum_{i < 2^n} a_{ij} g_{ij} x_i \right)^2 - d_\infty \sum_{i,j} a_{ij}^2 x_i^2 - 4d_\infty \sum_{l=0}^k b_l(x) \right] \\
& \leq 4d_\infty \sum_{l \geq 0} 2^{\frac{|n-l|}{4}} \sup_{x \in D_2^n} \max_j \sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2.
\end{aligned} \tag{3.6}$$

where

$$d_\infty = \lim_{k \rightarrow \infty} d_k = \prod_{l \geq 0} \left(1 + 2^{\frac{-|n-l|}{4}}\right) \leq C.$$

Now notice that

$$\max_j \sum_{i < 2^n} a_{ij}^2 y_i^2 \leq \max_j \sqrt{\sum_{i < 2^n} a_{ij}^4} \sqrt{\sum_i y_i^4} \leq \Delta^2 \|y\|_2 \|y\|_\infty. \tag{3.7}$$

Thus

$$4 \sum_{l \geq 0} 2^{\frac{|n-l|}{4}} \sup_{x \in D_2^n} \max_j \sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \leq 4\Delta^2 \left(\sum_{l=0}^n 2^{\frac{|n-l|}{4}} 2^{\frac{-l}{2}} + \sum_{l \geq n+1} 2^{\frac{|l-n|}{4}} 2^{\frac{n-2l}{2}} \right) \leq C\Delta^2 2^{\frac{n}{4}}$$

Using inequality (3.7), we also get

$$\max_j \sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \leq \begin{cases} \Delta^2 \left(\sum_{m=1}^l 2^{\frac{m-2l}{2}} c_m(x) + 2^{\frac{-l}{2}} \sum_{m=l+1}^n c_m(x) \right) & \text{for } l \leq n \\ \Delta^2 \left(\sum_{m=1}^n 2^{\frac{m-2l}{2}} c_m(x) \right) & \text{for } l > n \end{cases}$$

Therefore, by Lemma (3.5),

$$\sum_{l \geq 0} b_l(x) = \sum_{l \geq 0} 2^{\frac{|n-l|}{4}} \log \# \Pi_l^n \max_j \sum_{i \in B_l(x)} a_{ij}^2 x_i^2 \leq C\Delta^2 \sum_{m=1}^n \alpha_m c_m(x),$$

where

$$\alpha_m = \sum_{l=0}^{m-1} 2^{\frac{n-l}{4}} 2^l (1+n-l) 2^{\frac{-l}{2}} + \sum_{l=m}^n 2^{\frac{n-l}{4}} 2^l (1+n-l) 2^{\frac{m-2l}{2}} + \sum_{l \geq n+1} 2^{\frac{l-n}{4}} 2^n (1+l-n) 2^{\frac{m-2l}{2}}.$$

To estimate the first and the third term, we use respectively, (3.3) with $a = \frac{1}{4}$ and $b = 1$ and (3.4) with $a = \frac{3}{4}$ and $b = 1$, when $p = 2$. To bound the second term, we notice that $1 + n - l \leq 1 + n - m$, for $m \leq l \leq n$. Hence, we get

$$\alpha_m \leq C 2^{\frac{n+m}{4}} (1 + n - m) \leq C 2^{\frac{n}{2}} 2^{\frac{m-n}{8}} \sup_{x \geq 0} ((1+x) 2^{-\frac{x}{8}}) \leq C 2^{\frac{n}{2}} 2^{\frac{m-n}{8}}.$$

Therefore,

$$\sum_{l \geq 0} b_l(x) \leq C\Delta^2 2^{\frac{n}{2}} \sum_{l=1}^n 2^{\frac{l-n}{8}} c_l(x)$$

and the lemma follows by the inequality (3.6). ■

Theorem 3.8.

$$\mathbb{E} \sup_{\|x\|_2 \leq 1} \sum_j \left(\sum_{i \leq j} a_{ij} g_{ij} x_i \right)^2 \leq C \left(\max_i \sum_j a_{ij}^2 + \sqrt{\sum_{ij} a_{ij}^4} \right)$$

Proof. By Lemma (3.6), it is enough to show that

$$\mathbb{E} \sup_{x \in D_2} \sum_j \left(\sum_{i \leq j} a_{ij} g_{ij} x_i \right)^2 \leq C \left(\max_i \sum_j a_{ij}^2 + \sqrt{\sum_{ij} a_{ij}^4} \right)$$

Since the norm is invariant to column and row exchanges, we will assume that

$$\sum_{i=1}^n a_{i,j_1}^4 \geq \sum_{i=1}^n a_{i,j_2}^4, \text{ for } j_1 < j_2$$

Set for $n = 1, 2, \dots$, and $x \in D_2$,

$$\Delta_n^4 = \sum_i a_{i,2^{n-1}}^4 = \max_{j \geq 2^{n-1}} \sum_i a_{ij}^4 \geq \max_{j \geq 2^{n-1}} \sum_{i \leq j} a_{ij}^4$$

and

$$f_n(x) = \sum_{2^{n-1} \leq j < 2^n} \sum_{i \leq j} a_{ij}^2 x_i^2 + \Delta_n^2 2^{\frac{n}{2}} \sum_{l \leq n} 2^{\frac{l-n}{8}} c_l(x),$$

where as in the proof of lemma (3.7),

$$c_n(x) = \sqrt{\sum_{2^{n-1} \leq i < 2^n} x_i^2}.$$

Let C_1 be the same constant as in lemma (3.7), then

$$\begin{aligned} & \mathbb{E} \sup_{x \in D_2} \left[\sum_j \left(\sum_{i \leq j} a_{ij} g_{ij} x_i \right)^2 - C_1 \sum_{n=1}^{\infty} f_n(x) \right] \\ & \leq \sum_{n=1}^{\infty} \mathbb{E} \sup_{x \in D_2} \left[\sum_{2^{n-1} \leq j < 2^n} \left(\sum_{i \leq j} a_{ij} g_{ij} x_i \right)^2 - C_1 f_n(x) \right] \leq \sum_{n=1}^{\infty} C_1 \Delta_n^2 2^{\frac{n}{4}}. \end{aligned} \quad (3.8)$$

Notice that $2^{n-1} \Delta_n^4 \leq \sum_{ij} a_{ij}^4$, so,

$$\sum_{n \geq 1} \Delta_n^2 2^{\frac{n}{4}} \leq C \sqrt{\sum_{ij} a_{ij}^4}. \quad (3.9)$$

Obviously, we have,

$$\sup_{x \in D_2} \sum_{i \leq j} a_{ij}^2 x_i^2 \leq \max_i \sum_j a_{ij}^2. \quad (3.10)$$

For any $u, v \in l_2$, we have

$$\sum_{l \leq n} u_l v_n 2^{\frac{l-n}{8}} = \sum_{i \geq 0} 2^{-\frac{i}{8}} \sum_l u_l v_{l+i} \leq \sum_{i \geq 0} 2^{-\frac{i}{8}} \|u\|_2 \|v\|_2 \leq C \|u\|_2 \|v\|_2$$

Since $\sum_l c_l(x)^2 = \|x\|_2^2$ and $\sum_n \Delta_n^4 2^n \leq C \sum_{ij} a_{ij}^4$, we get

$$\sum_{l \leq j} c_l(x) \Delta_j^2 2^{\frac{j}{2}} 2^{\frac{l-j}{8}} \leq C \|x\|_2 \sqrt{\sum_{ij} a_{ij}^4}. \quad (3.11)$$

By (3.10) and (3.11), we get

$$\sum_{n=1}^{\infty} f_n(x) \leq \max_i \sum_j a_{ij}^2 + C \sqrt{\sum_{ij} a_{ij}^4} \text{ for any } x \in D_2. \quad (3.12)$$

Inequalities, (3.8), (3.9) and (3.12) imply the theorem. ■

Chapter 4

Extending the result for p -norm

We extend the main result in the previous chapter for the p -norm, $p > 2$.

Lemma 4.1. (*Gaussian Concentration*) Suppose that g_n are iid $\mathcal{N}(0, 1)$ random vari-

ables and, $G = \sup_i (l_i + \sum_n k_{i,n} g_n)$,

Then,

$$\mathbb{P}(G \geq (EG^p)^{1/p} + wt) \leq \mathbb{P}(G \geq EG + wt) \leq e^{-t^2/2},$$

$$\text{where } w = \sup_i \sqrt{\sum_n k_{i,n}^2}.$$

Proof. The proof is identical to the proof of (3.1). ■

Lemma 4.2. Suppose that random variables Y_x, A_x and numbers $b_x \geq 0$, satisfy the inequality

$$\mathbb{P}(Y_x \geq A_x + tb_x) \leq e^{-t^{\frac{2}{p}}}, \forall t > 0, x \in U.$$

Then,

$$\mathbb{E} \max_{x \in U} \left[Y_x - A_x - b_x \left(2^{\frac{p-2}{p}} \log \#U \right)^{\frac{p}{2}} \right] \leq C \max_{x \in U} b_x. \quad (4.1)$$

Proof. Let $Z_x = Y_x - A_x - b_x(2^{\frac{p-2}{p}} \log \#U)^{\frac{p}{2}}$. Then,

$$\begin{aligned} \mathbb{E} \max_{x \in U} Z_x &\leq \int_0^\infty \mathbb{P}(\max_{x \in U} Z_x \geq s) ds \leq \sum_{x \in U} \int_0^\infty \mathbb{P}\left(Y_x \geq A_x + b_x \left[\left(2^{\frac{p-2}{p}} \log \#U\right)^{\frac{p}{2}} + \frac{s}{b_x} \right]\right) ds \\ &\leq \sum_{x \in U} \frac{1}{\#U} \int_0^\infty \exp\left[-\left(\frac{s}{2^{\frac{p-2}{p}} b_x}\right)^{\frac{2}{p}}\right] ds = \frac{C}{\#U} \sum_{x \in U} b_x \leq C \max_{x \in U} b_x \end{aligned}$$

where we used the fact that $(|a|^{\frac{p}{2}} + |b|^{\frac{p}{2}})^{\frac{2}{p}} \geq \frac{|a|}{2^{\frac{p-2}{p}}} + \frac{|b|}{2^{\frac{p-2}{p}}}$

■

Lemma 4.3. *Suppose that $g_{i,j}$ are iid $\mathcal{N}(0, 1)$ r.v., U is a finite subset of l_2 and for each $x \in U$, there are given sets $I(x)$ and r.v.s, $(L_j(x))_j$ independent of $(g_{ij})_{i \in I(x), j}$. Then, for any $c_1, c_2 > 0$,*

$$\begin{aligned} \mathbb{E} \max_{x \in U} &\left[\sum_j (|L_j(x) + \sum_{i \in I(x)} a_{ij} g_{ij} x_i|)^p - (1 + c_2)^{p-1} (1 + c_1)^{p-1} \sum_j |L_j(x)|^p \right. \\ &- C_1 (1 + c_2)^{p-1} (1 + c_1)^{p-1} \left(\frac{1}{c_1}\right)^{p-1} \sum_j \left(\sum_{i \in I(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \\ &- C_2 (1 + c_2)^{p-1} \left(\frac{1}{c_2}\right)^{p-1} \max_j \left(\sum_{i \in I(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} (\log \#U)^{\frac{p}{2}} \Big] \\ &\leq C_2 (1 + c_2)^{p-1} \left(\frac{1}{c_2}\right)^{p-1} \max_{x \in U} \max_j \left(\sum_{i \in I(x)} a_{ij}^2 x_i^2 \right)^{p/2}. \end{aligned}$$

where C_1, C_2 are some constants depending on p .

Proof. Let

$$G_x = \sqrt[p]{\sum_j |L_j(x) + \sum_{i \in I(x)} a_{ij} g_{ij} x_i|^p} = \sup_{\|y\|_{p^*} \leq 1} \sum_j (L_j(x) + \sum_{i \in I(x)} a_{ij} g_{ij} x_i) y_j$$

Then, $\forall c_1 > 0$, $\mathbb{E}_g G_x^p \leq (1 + c_1)^{p-1} \sum_j |L_j(x)|^p + C(1 + \frac{1}{c_1})^{p-1} \sum_j \left(\sum_{i \in I(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}}$,

where, $\mathbb{E}_g$ denotes the expectation with respect to the gaussian random variables, $(g_{ij})_{i \in I(x)}$.

Here, we used Rosenthal's inequality, theorem (2.9).

If we put,

$$w(x) = \sup_{\|y\|_2 \leq 1} \sqrt{\sum_j \sum_{i \in I(x)} a_{ij}^2 x_i^2 y_j^2} = \max_j \sqrt{\sum_{i \in I(x)} a_{ij}^2 x_i^2}$$

Then since, $\forall c_2 > 0$,

$$\begin{aligned} & \left(\sqrt[p]{\mathbb{E}_g G_x^p} + tw(x) \right)^p \leq (1 + c_2)^{p-1} \mathbb{E}_g G_x^p + \left(1 + \frac{1}{c_2}\right)^{p-1} t^p w^p(x) \\ & \leq (1 + c_2)^{p-1} (1 + c_1)^{p-1} \sum_j |L_j(x)|^p + C(1 + c_2)^{p-1} \left(1 + \frac{1}{c_1}\right)^{p-1} \sum_j \left(\sum_{i \in I(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} + \left(1 + \frac{1}{c_2}\right)^{p-1} t^p w^p(x) \end{aligned}$$

we get, by Lemma (4.1)

$$\begin{aligned} & \mathbb{P}(G_x^p \geq (1 + c_2)^{p-1} \mathbb{E}_g G_x^p + \left(1 + \frac{1}{c_2}\right)^{p-1} t^p w^p(x)) \\ & \leq \mathbb{P}(G_x \geq \sqrt[p]{\mathbb{E}_g G_x^p} + tw(x)) \leq e^{-t^2/2} \end{aligned}$$

Therefore, Lemma (4.3) follows by Lemma(4.2) applied to

$$\begin{aligned} Y_x &= G_x^p, A_x = (1 + c_2)^{p-1} (1 + c_1)^{p-1} \sum_j |L_j(x)|^p + C(1 + c_2)^{p-1} \left(1 + \frac{1}{c_1}\right)^{p-1} \left(\sum_j \left(\sum_{i \in I(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \right) \\ b_x &= 2^{\frac{p}{2}} \left(1 + \frac{1}{c_2}\right)^{p-1} \max_j \left(\sum_{i \in I(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \end{aligned}$$

■

Lemma 4.4. *Let $\Delta = \max_{2^{n-1} \leq j < 2^n} \left(\sum_{i < 2^n} a_{ij}^{\frac{2p}{p-1}} \right)^{\frac{p-1}{2p}}$. Then,*

$$\begin{aligned} & \mathbb{E} \sup_{x \in D_2} \left[\sum_{2^{n-1} \leq j < 2^n} \left(\left| \sum_{i < 2^n} a_{ij} g_{ij} x_i \right| \right)^p - C'_1 n [n \ln(n) \ln^2(\ln(n))]^{p-1} \sum_{2^{n-1} \leq j < 2^n} \left(\sum_{i < 2^n} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \right. \\ & \quad \left. - C'_2 \Delta^p 2^{\frac{n}{2}} \sum_{l=1}^n 2^{\frac{l-n}{8}} \sqrt{\sum_{2^{i-1} \leq i < 2^l} x_i^2} \right] \leq C'_2 \Delta^p 2^{\frac{n}{4}} \end{aligned}$$

Proof. As in the proof of Lemma(3.7),

For $x \in D_2^n, l, n = 1, 2, \dots$, let us define

$$A_l^n(x) = \{i < 2^n : i \in A_l(x)\},$$

$$B_0^n(x) = A_0^n(x) = \{i < 2^n : x_i^2 = 1\}, b_0(x) = 2^{\frac{n}{4}} (\log \# \Pi_0^n \max_j \sum_{i \in B_0^n(x)} a_{ij}^2 x_i^2)^{\frac{p}{2}},$$

$$B_l^n(x) = A_l^n(x) \setminus A_{l-1}^n(x) = \{i < 2^n : i \geq 2^{l-1}, x_i^2 = 2^{-l}$$

$$\text{or } 2^{m-1} \leq i < 2^m, x_i^2 \in \{2^{m-2l}, 2^{m-1-2l}\}, m = 1, 2, \dots, l\},$$

$$c_l(x) = \sqrt{\sum_{2^{l-1} \leq i < 2^l} x_i^2}, b_l(x) = 2^{\frac{|n-l|}{4}} (\log \# \Pi_l^n \max_j \sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2)^{\frac{p}{2}}$$

$$d_k = \prod_{l=0}^k \left[\left(1 + 2^{\frac{-|n-l|}{4(p-1)}} \right) \right]^{p-1}$$

and

$$e_k = \prod_{l=0}^k \left[\left(1 + \frac{1}{(|n-l|+3)(\ln(|n-l|+3))(\ln^2(\ln(|n-l|+3)))} \right) \right]^{p-1}$$

We prove by induction on $k = 0, 1, \dots$, that

$$\begin{aligned} & \mathbb{E} \sup_{x \in D_2^n} \left[\sum_j \left(\left| \sum_{i \in A_k^n(x)} a_{ij} g_{ij} x_i \right| \right)^p \right. \\ & - C_1 d_k e_k \sum_j \sum_{l=0}^k \left[(|n-l|+3)(\ln(|n-l|+3))(\ln^2(\ln(|n-l|+3))) \right]^{p-1} \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \\ & \left. - C_2 d_k e_k \sum_{l=0}^k b_l(x) \right] \\ & \leq C_2 d_k e_k \sum_{l=0}^k 2^{\frac{|n-l|}{4}} \sup_{x \in D_2^n} \max_j \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}}, \text{ where } 2^{n-1} \leq j < 2^n. \end{aligned} \quad (4.2)$$

For $k = 0$, inequality (4.2) follows by Lemma(4.3), applied to $U = \Pi_0^n, L_j(x) =$

$0, I(x) = B_0^n(x), c_1 = [(n+3)(\ln(n+3))(\ln^2(\ln(n+3)))]^{-1}$ and $c_2 = 2^{-\frac{n}{4}}$.

To show the inductive step, let us denote

$$\begin{aligned}
S_k(x) = & \left[\sum_j \left(\left| \sum_{i \in A_k^n(x)} a_{ij} g_{ij} x_i \right| \right)^p \right. \\
& - C_1 d_k e_k \sum_j \sum_{l=0}^k \left[(|n-l|+3)(\ln(|n-l|+3))(\ln^2(\ln(|n-l|+3))) \right]^{p-1} \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \\
& \left. - C_2 d_k e_k \sum_{l=0}^k b_l(x) \right]
\end{aligned}$$

Then,

$$\begin{aligned}
S_k(x) - & \left(1 + 2^{\frac{-|n-k|}{4}} \right)^{p-1} \left(1 + \left[(|n-k|+3)(\ln(|n-k|+3))(\ln^2(\ln(|n-k|+3))) \right]^{-1} \right)^{p-1} S_{k-1}(x) \\
= & \sum_j \left(\left| \sum_{i \in A_{k-1}^n(x)} a_{ij} g_{ij} x_i \right| + \sum_{i \in B_k^n(x)} a_{ij} g_{ij} x_i \right)^p - d_k e_k \sum_j \left| \sum_{i \in A_{k-1}^n(x)} a_{ij} g_{ij} x_i \right|^p \\
& - C_1 d_k e_k \sum_j \left[(|n-k|+3)(\ln(|n-k|+3))(\ln^2(\ln(|n-k|+3))) \right]^{p-1} \left(\sum_{i \in B_k^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \\
& - C_2 d_k e_k 2^{\frac{|n-k|}{4}} (\log \# \Pi_k^n)^{\frac{p}{2}} \max_j \left(\sum_{i \in B_k^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}}
\end{aligned}$$

Lemma (4.3) applied to

$$U = \Pi_k^n, L_j(x) = \sum_{i \in A_{k-1}^n(x)} a_{ij} g_{ij} x_i, I(x) = B_k^n(x),$$

$$c_1 = \left[(|n-k|+3)(\ln(|n-k|+3))(\ln^2(\ln(|n-k|+3))) \right]^{-1}, c_2 = 2^{\frac{-|n-k|}{4(p-1)}}$$

implies

$$\begin{aligned}
& \mathbb{E} \sup_{x \in D_2^n} \left[S_k(x) - (1+c_1)^{p-1} (1+c_2)^{p-1} S_{k-1}(x) \right] \\
& \leq C_2 (1+c_1)^{p-1} (1+c_2)^{p-1} 2^{\frac{|n-k|}{4}} \sup_{x \in D_2^n} \max_j \left(\sum_{i \in B_k^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}}.
\end{aligned}$$

The induction step from $k-1$ to k easily follows, since

$$\mathbb{E} \sup_{x \in D_2^n} S_k(x) \leq \mathbb{E} \sup_{x \in D_2^n} \left[S_k(x) - (1+c_1)^{p-1} (1+c_2)^{p-1} S_{k-1}(x) \right]$$

$$+(1+c_1)^{p-1}(1+c_2)^{p-1}\mathbb{E}\sup_{x\in D_2^n} S_{k-1}(x).$$

When we take $k \rightarrow \infty$, in (4.2), we get

$$\begin{aligned} & \mathbb{E}\sup_{x\in D_2^n} \left[\sum_j \left(\left| \sum_{i<2^n} a_{ij} g_{ij} x_i \right| \right)^p \right. \\ & - C_1(de)_\infty \sum_j \sum_{l \geq 0} [(|n-l|+3)(\ln(|n-l|+3))(\ln^2(\ln(|n-l|+3)))]^{p-1} \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \\ & \left. - C_2(de)_\infty \sum_{l=0}^k b_l(x) \right] \\ & \leq C_2(de)_\infty \sum_{l \geq 0} 2^{\frac{|n-l|}{4}} \sup_{x \in D_2^n} \max_j \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}}. \end{aligned} \quad (4.3)$$

where

$$(de)_\infty = \prod_{l=0}^{\infty} \left(\left(1 + 2^{\frac{-|n-l|}{4(p-1)}} \right)^{p-1} \left(1 + \frac{1}{(|n-l|+3)(\ln(|n-l|+3))(\ln^2(\ln(|n-l|+3)))} \right)^{p-1} \right) \leq C$$

Now notice that

$$\max_j \left(\sum_{i<2^n} a_{ij}^2 y_i^2 \right)^{\frac{p}{2}} \leq \max_j \left(\sum_{i<2^n} a_{ij}^{\frac{2p}{p-1}} \right)^{\frac{p-1}{2}} \sqrt{\sum_i y_i^{2p}} \leq \Delta^p \|y\|_2 \|y\|_\infty^{p-1}. \quad (4.4)$$

Thus,

$$\sum_{l \geq 0} 2^{\frac{|n-l|}{4}} \sup_{x \in D_2^n} \max_j \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \leq \Delta^p \left(\sum_{l=0}^n 2^{\frac{|n-l|}{4}} 2^{\frac{-l(p-1)}{2}} + \sum_{l \geq n+1} 2^{\frac{|l-n|}{4}} 2^{\frac{n-pl}{2}} \right) \leq C \Delta^p 2^{\frac{n}{4}}$$

Using inequality (4.4), we also get

$$\max_j \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \leq \begin{cases} \Delta^p \left(\sum_{m=1}^l 2^{\frac{m-pl}{2}} c_m(x) + 2^{\frac{-l(p-1)}{2}} \sum_{m=l+1}^n c_m(x) \right) & \text{for } l \leq n \\ \Delta^p \left(\sum_{m=1}^n 2^{\frac{m-pl}{2}} c_m(x) \right) & \text{for } l > n \end{cases}$$

Therefore, by Lemma (3.5),

$$\sum_{l \geq 0} b_l(x) = \sum_{l \geq 0} 2^{\frac{|n-l|}{4}} \left(\log \# \Pi_l^n \max_j \sum_{i \in B_l(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \leq C \Delta^p \sum_{m=1}^n \alpha_m c_m(x),$$

where

$$\alpha_m = \sum_{l=0}^{m-1} 2^{\frac{n-l}{4}} 2^{\frac{lp}{2}} (1+n-l)^{\frac{p}{2}} 2^{\frac{-l(p-1)}{2}} + \sum_{l=m}^n 2^{\frac{n-l}{4}} 2^{\frac{lp}{2}} (1+n-l)^{\frac{p}{2}} 2^{\frac{m-pl}{2}} + \sum_{l \geq n+1} 2^{\frac{l-n}{4}} 2^{\frac{np}{2}} (1+l-n)^{\frac{p}{2}} 2^{\frac{m-pl}{2}}.$$

To estimate the first and the third term, we use respectively, (3.3) with $a = \frac{1}{4}$ and $b = 1$ and (3.4) with $a = \frac{3}{4}$ and $b = 1$. To bound the second term, we notice that $1 + n - l \leq 1 + n - m$, for $m \leq l \leq n$. Hence, we get

$$\alpha_m \leq C 2^{\frac{n+m}{4}} (1+n-m)^{\frac{p}{2}} \leq C 2^{\frac{n}{2}} 2^{\frac{m-n}{8}} \sup_{x \geq 0} ((1+x)^{\frac{p}{2}} 2^{-\frac{x}{8}}) \leq C 2^{\frac{n}{2}} 2^{\frac{m-n}{8}}.$$

Therefore,

$$\begin{aligned} \sum_{l \geq 0} b_l(x) &\leq C \Delta^p 2^{\frac{n}{2}} \sum_{l=1}^n 2^{\frac{l-n}{8}} c_l(x) \\ &\quad \sum_{2^{n-1} \leq j < 2^n} \sum_{l \geq 0} [(|n-l|+3)(\ln(|n-l|+3))(\ln^2(\ln(|n-l|+3)))]^{p-1} \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \\ &= \sum_{2^{n-1} \leq j < 2^n} \sum_{l=0}^n [(|n-l|+3)(\ln(|n-l|+3))(\ln^2(\ln(|n-l|+3)))]^{p-1} \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \\ &\quad + \sum_{2^{n-1} \leq j < 2^n} \sum_{l=n+1}^{\infty} [(|n-l|+3)(\ln(|n-l|+3))(\ln^2(\ln(|n-l|+3)))]^{p-1} \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \end{aligned}$$

Since $2^{n-1} \leq j \leq 2^n$,

$$\begin{aligned} &\sum_j \sum_{l=n+1}^{\infty} [(|n-l|+3)(\ln(|n-l|+3))(\ln^2(\ln(|n-l|+3)))]^{p-1} \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \\ &\leq C_3 \sum_{l=0}^{\infty} 2^{\frac{|n-l|}{4}} \left(\log \# \Pi_l^n \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right) \right)^{\frac{p}{2}} \end{aligned}$$

and

$$\begin{aligned} &\sum_j \sum_{l=0}^n [(|n-l|+3)(\ln(|n-l|+3))(\ln^2(\ln(|n-l|+3)))]^{p-1} \left(\sum_{i \in B_l^n(x)} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \\ &\leq C_4 n [(n+3)(\ln(n+3))(\ln^2(\ln(n+3)))]^{p-1} \sum_j \left(\sum_{i < 2^n} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \end{aligned}$$

and the lemma follows by simplifying inequality (4.3) with the above relations. ■

Theorem 4.5.

$$\mathbb{E} \sup_{\|x\|_2 \leq 1} \sum_{j=1}^N \left(\left| \sum_{i \leq j} a_{ij} g_{ij} x_i \right| \right)^p \leq C_N \left[\max_i \sum_{j=1}^N |a_{ij}|^p + \sqrt{\sum_{j=1}^N \left(\sum_{i=1}^m a_{ij}^{\frac{2p}{p-1}} \right)^{p-1}} \right]$$

where $C_N = C (\ln N)^p [\ln(\ln N) \ln^2(\ln(\ln(N)))]^{p-1}$

Proof. By Lemma (3.6), it is enough to show that

$$\mathbb{E} \sup_{x \in D_2} \sum_{j=1}^N \left(\left| \sum_{i \leq j} a_{ij} g_{ij} x_i \right| \right)^p \leq C_N \left[\max_i \sum_{j=1}^N |a_{ij}|^p + \sqrt{\sum_{j=1}^N \left(\sum_{i=1}^m a_{ij}^{\frac{2p}{p-1}} \right)^{p-1}} \right]$$

Since, the norm is invariant to column and row exchanges, we may assume that

$$\sum_{i=1}^m a_{i,j_1}^{\frac{2p}{p-1}} \geq \sum_{i=1}^m a_{i,j_2}^{\frac{2p}{p-1}} \text{ for } j_1 < j_2$$

Set for $n = 1, 2, \dots$, and $x \in D_2$,

$$\Delta_n^{\frac{2p}{p-1}} = \sum_i a_{i,2^{n-1}}^{\frac{2p}{p-1}} = \max_{j \geq 2^{n-1}} \sum_i a_{ij}^{\frac{2p}{p-1}} \geq \max_{j \geq 2^{n-1}} \sum_{i \leq j} a_{ij}^{\frac{2p}{p-1}}$$

and

$$f_n(x) = n^p [\ln(n) \ln^2(\ln(n))]^{p-1} \left(\sum_{2^{n-1} \leq j < 2^n} \left(\sum_{i \leq j} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \right) + \Delta_n^p 2^{\frac{n}{2}} \sum_{l \leq n} 2^{\frac{l-n}{8}} c_l(x),$$

where as in the proof of lemma (4.4),

$$c_n(x) = \sqrt{\sum_{2^{n-1} \leq i < 2^n} x_i^2}.$$

Let $C_1 = \max(C'_1, C'_2)$, the constants in lemma (4.4). Then,

$$\begin{aligned} & \mathbb{E} \sup_{x \in D_2} \left[\sum_{j=1}^N \left(\left| \sum_{i \leq j} a_{ij} g_{ij} x_i \right| \right)^p - C_1 \sum_{n=1}^{\lceil \log N + 1 \rceil} f_n(x) \right] \\ & \leq \sum_{n=1}^{\lceil \log N + 1 \rceil} \mathbb{E} \sup_{x \in D_2} \left[\sum_{2^{n-1} \leq j < 2^n} \left(\sum_{i \leq j} a_{ij} g_{ij} x_i \right)^p - C_1 f_n(x) \right] \leq \sum_{n=1}^{\lceil \log N + 1 \rceil} C_1 \Delta_n^p 2^{\frac{n}{4}}. \end{aligned} \tag{4.5}$$

Notice that $2^{n-1}\Delta_n^{2p} \leq \sum_j \left(\sum_i a_{ij}^{\frac{2p}{p-1}} \right)^{p-1}$, so,

$$\sum_{n \geq 1} \Delta_n^p 2^{\frac{n}{4}} \leq C \sqrt{\sum_j \left(\sum_i a_{ij}^{\frac{2p}{p-1}} \right)^{p-1}}. \quad (4.6)$$

We have,

$$\sup_{x \in D_2} \sum_{j=1}^N \left(\sum_{i \leq j} a_{ij}^2 x_i^2 \right)^{\frac{p}{2}} \leq \max_i \sum_{j=1}^N |a_{ij}|^p. \quad (4.7)$$

For any $u, v \in l_2$, we have

$$\sum_{l \leq n} u_l v_n 2^{\frac{l-n}{8}} = \sum_{i \geq 0} 2^{-\frac{i}{8}} \sum_l u_l v_{l+i} \leq \sum_{i \geq 0} 2^{-\frac{i}{8}} \|u\|_2 \|v\|_2 \leq C \|u\|_2 \|v\|_2$$

Since $\sum_l c_l(x)^2 = \|x\|_2^2$ and $\sum_n \Delta_n^{2p} 2^n \leq C \sum_j \left(\sum_i a_{ij}^{\frac{2p}{p-1}} \right)^{p-1}$, we get

$$\sum_{j=1}^{\lceil \log N+1 \rceil} \sum_{l \leq j} c_l(x) \Delta_j^p 2^{\frac{j}{2}} 2^{\frac{l-j}{8}} \leq C \|x\|_2 \sqrt{\sum_{j=1}^N \left(\sum_{i=1}^m a_{ij}^{\frac{2p}{p-1}} \right)^{p-1}}. \quad (4.8)$$

By (4.7) and (4.8), we get

$$\sum_{n=1}^{\lceil \log N+1 \rceil} f_n(x) \leq C_N \left(\max_i \sum_{j=1}^N |a_{ij}|^p + \sqrt{\sum_{j=1}^N \left(\sum_{i=1}^m a_{ij}^{\frac{2p}{p-1}} \right)^{p-1}} \right) \text{ for any } x \in D_2. \quad (4.9)$$

Inequalities, (4.5), (4.6) and (4.9) imply the theorem. ■

Chapter 5

Putting it all together

Theorem 5.1. *For any $p > 2$ and $N \times m$ matrix (a_{ij}) , the following estimate holds:*

$$\mathbb{E} \|(a_{ij}g_{ij})\|_p \leq (C_N)^{\frac{1}{p}} \left(\max_{1 \leq i \leq m} \sqrt[p]{\sum_{j=1}^N a_{ij}^p} + \sqrt[p]{\sum_{j=1}^N \left(\sum_{i=1}^m a_{ij}^{\frac{2p}{p-1}} \right)^{p-1}} + \max_{1 \leq j \leq m} \sqrt[p]{\sum_{i=1}^N a_{ij}^2} + \sqrt[p]{\sum_{j=1}^m \sum_{i=1}^m a_{ij}^4} \right)$$

Proof. Notice that,

$$\begin{aligned} \sup_{\|x\|_2, \|y\|_{p^*} \leq 1} \sum_{ij} a_{ij}g_{ij}x_iy_j &\leq \sup_{\|x\|_2, \|y\|_{p^*}} \sum_{i \leq j} a_{ij}g_{ij}x_iy_j + \sup_{\|x\|_2, \|y\|_{p^*}} \sum_{j < i} a_{ij}g_{ij}x_iy_j \\ &= \sup_{\|x\|_2 \leq 1} \sqrt[p]{\sum_{j=1}^N \left(\sum_{i \leq j} a_{ij}g_{ij}x_i \right)^p} + \sup_{\|y\|_{p^*} \leq 1} \sqrt[p]{\sum_{i=1}^m \left(\sum_{j < i} a_{ij}g_{ij}y_j \right)^2} \\ &\leq \sup_{\|x\|_2 \leq 1} \sqrt[p]{\sum_{j=1}^N \left(\sum_{i \leq j} a_{ij}g_{ij}x_i \right)^p} + \sup_{\|y\|_2 \leq 1} \sqrt[p]{\sum_{i=1}^m \left(\sum_{j < i} a_{ij}g_{ij}y_j \right)^2} \\ \left(\sup_{\|x\|_2, \|y\|_{p^*} \leq 1} \sum_{ij} a_{ij}g_{ij}x_iy_j \right)^2 &\leq C \left[\sup_{\|x\|_2 \leq 1} \sqrt[p]{\sum_{j=1}^N \left(\sum_{i \leq j} a_{ij}g_{ij}x_i \right)^p} + \sup_{\|y\|_2 \leq 1} \sum_{i=1}^m \left(\sum_{j < i} a_{ij}g_{ij}y_j \right)^2 \right] \end{aligned}$$

Hence, to prove theorem (5.1), it is enough to show

$$\mathbb{E} \sup_{\|x\|_2 \leq 1} \sum_{j=1}^N \left(\sum_{i \leq j} a_{ij}g_{ij}x_i \right)^p \leq C_N \left(\max_i \sum_{j=1}^N |a_{ij}|^p + \sqrt[p]{\sum_{j=1}^N \left(\sum_{i=1}^m a_{ij}^{\frac{2p}{p-1}} \right)^{p-1}} \right)$$

and

$$\mathbb{E} \sup_{\|y\|_2 \leq 1} \sum_{i=1}^m \left(\sum_{j < i} a_{ij}g_{ij}y_j \right)^2 \leq C \left(\max_{1 \leq j \leq m} \sum_{i=1}^m a_{ij}^2 + \sqrt[p]{\sum_{j=1}^m \sum_{i=1}^m a_{ij}^4} \right)$$

These follow from (4.5) and (3.8), using the assumption that

$$\sum_{j=1}^N a_{i_1,j}^4 \geq \sum_{j=1}^N a_{i_2,j}^4 \text{ for } i_1 < i_2.$$

. ■

Theorem 5.2. *For any $p > 2$ and $N \times m$ random matrix, whose entries (X_{ij}) , satisfy*

$\mathbb{E}X_{ij}^{2p} < \infty$, the following estimate holds:

$$\begin{aligned} \mathbb{E}\|(X_{ij})\|_p &\leq (C_N)^{\frac{1}{p}} \left[\max_{1 \leq i \leq m} \sqrt[p]{\sum_{j=1}^N \mathbb{E}|X_{ij}|^p} + \sqrt[p]{m^{p-2} \sum_{i=1}^m \sum_{j=1}^N \mathbb{E}X_{ij}^{2p}} \right. \\ &\quad \left. + \max_{1 \leq j \leq N} \sqrt{\sum_{i=1}^m \mathbb{E}X_{ij}^2} + \sqrt[4]{\sum_{i=1}^m \sum_{j=1}^m \mathbb{E}X_{ij}^4} \right] \end{aligned}$$

Proof. Let $\tilde{X}_{ij}$ be an independent copy of random variables X_{ij} and ε_{ij} be a sequence of independent symmetric Bernoulli random variables; i.e., $\mathbb{P}(\varepsilon_{ij} = \pm 1) = 1/2$, independent of all other random variables. Then, by Jensen's inequality,

$$\mathbb{E}\|(X_{ij})\|_p \leq \mathbb{E}\|(X_{ij} - \tilde{X}_{ij})\|_p = \mathbb{E}\|(\varepsilon_{ij}(X_{ij} - \tilde{X}_{ij}))\|_p \leq 2\mathbb{E}\|(\varepsilon_{ij}X_{ij})\|_p.$$

Thus, it is enough to show theorem (5.2), in the case of symmetric random variables.

We have,

$$\mathbb{E}\|(a_{ij}g_{ij})\|_p = \mathbb{E}\|(a_{ij}\varepsilon_{ij}|g_{ij}|)\|_p \geq \mathbb{E}_\varepsilon\|(a_{ij}\varepsilon_{ij}\mathbb{E}|g_{ij}|)\|_p = \sqrt{\frac{2}{\pi}}\mathbb{E}\|(a_{ij}\varepsilon_{ij})\|_p.$$

Now, by the symmetry of X_{ij} s,

$$\begin{aligned} \mathbb{E}\|(X_{ij})\|_p^2 &= \mathbb{E}_X \mathbb{E}_\varepsilon \|(\varepsilon_{ij}X_{ij})\|_p^2 \leq \sqrt[p]{C_N} \mathbb{E} \left(\max_i \sum_{j=1}^N |X_{ij}|^p + \sqrt{\sum_{j=1}^N \left(\sum_{i=1}^m X_{ij}^{\frac{2p}{p-1}} \right)^{p-1}} \right) \\ &\quad + C \mathbb{E} \left(\max_{1 \leq j \leq m} \sqrt{\sum_{i=1}^m X_{ij}^2} + \sqrt[4]{\sum_{j=1}^m \sum_{i=1}^m X_{ij}^4} \right) \end{aligned}$$

$$\begin{aligned}
\mathbb{E} \max_i \sum_{j=1}^N (|X_{ij}|^p - \mathbb{E}|X_{ij}|^p) &\leq \mathbb{E} \sqrt{\sum_{i=1}^m \left(\sum_{j=1}^N (|X_{ij}|^p - \mathbb{E}|X_{ij}|^p) \right)^2} \\
&\leq \sqrt{\sum_{i=1}^m \mathbb{E} \left(\sum_{j=1}^N |X_{ij}|^p - \mathbb{E}|X_{ij}|^p \right)^2} \leq \sqrt{\sum_{i=1}^m \sum_{j=1}^N \mathbb{E}(X_{ij})^{2p}}
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E} \max_i \sum_{j=1}^N |X_{ij}|^p &\leq \mathbb{E} \max_i \sum_{j=1}^N (|X_{ij}|^p - \mathbb{E}|X_{ij}|^p) + \max_{1 \leq i \leq m} \sum_{j=1}^N \mathbb{E}|X_{ij}|^p \\
\mathbb{E} \max_i \sum_{j=1}^N |X_{ij}|^p &\leq \sqrt{\sum_{i=1}^m \sum_{j=1}^N \mathbb{E}(X_{ij})^{2p}} + \max_{1 \leq i \leq m} \sum_{j=1}^N \mathbb{E}|X_{ij}|^p
\end{aligned}$$

By Jensen's inequality,

$$\begin{aligned}
\mathbb{E} \sqrt{\sum_{j=1}^N \left(\sum_{i=1}^m X_{ij}^{\frac{2p}{p-1}} \right)^{p-1}} &\leq \sqrt{\sum_{j=1}^N \mathbb{E} \left(\sum_{i=1}^m X_{ij}^{\frac{2p}{p-1}} \right)^{p-1}} \\
\mathbb{E} \left(\sum_{i=1}^m X_{ij}^{\frac{2p}{p-1}} \right)^{p-1} &\leq m^{p-2} \sum_{i=1}^m \mathbb{E} X_{ij}^{2p}
\end{aligned}$$

Hence,

$$\mathbb{E} \sqrt{\sum_{j=1}^N \left(\sum_{i=1}^m X_{ij}^{\frac{2p}{p-1}} \right)^{p-1}} \leq \sqrt{m^{p-2} \sum_{j=1}^N \left(\sum_{i=1}^m \mathbb{E} X_{ij}^{2p} \right)}$$

■

Corollary 5.3. *For any $p > 2$ and $N \times m$ random matrix, whose entries (X_{ij}) are iid random variables, that satisfy $\mathbb{E} X_{ij}^{2p} = \mathbb{E} X^{2p} < \infty$, and $N \geq m^{\frac{p}{2}}$, the following estimate holds:*

$$\mathbb{E} \|(X_{ij})\|_p \leq (C_N)^{\frac{1}{p}} C \sqrt[p]{N}$$

Follows from theorem (5.2), where C is dependent on $\mathbb{E} X^p$ and $\mathbb{E} X^{2p}$.

Bibliography

- [Gordon85] Y. Gordon, Some inequalities for Gaussian processes and applications, *Israel Journal of Mathematics*, **50** (1985), 265-289.
- [Gordon88] Y. Gordon, On Milmans inequality and random spaces which escape through a mesh, in *Geometric Aspects of Functional Analysis (1986/87) Lecture Notes in Mathematics*, (1988), 84-106
- [Gordon07] Y. Gordon, A note on an observation of G. Schechtman, in *Geometric Aspects of Functional Analysis Lectures Notes in Mathematics*, **1910** (2007) 127 – 132
- [GiannopoulosMilman00] A. A. Giannopoulos, V. D. Milman, Concentration property on probability spaces, *Adv. Math.*, **156** (2000), 77-106.
- [GuedonRudelson07] O. Guedon, M. Rudelson, L_p moments of random vectors via majorizing measures, *Adv. Math.* **208** (2007), 798-823.
- [Latala05] R. Latała, Some estimates of norms of random matrices, *Proc. Amer. Math. Soc.*, **133** (2005), 1273-1282
- [LedouxTalagrand91] M. Ledoux and M. Talagrand, *Probability in Banach spaces*, Springer, 1991
- [MendelsonJaegermann08] S. Mendelson and N. Tomczak-Jaegermann, A subgaussian embedding theorem, *Israel Journal of Mathematics* **164** (2008) 349 – 364
- [Rudelson13] M. Rudelson, Lecture notes on non-asymptotic theory of random matrices, **2013**
- [Rudelson06] M. Rudelson, Lower estimates for the singular values of random matrices, *Compt. Rendus Math. de L'Academie des Sciences*, **342** (2006), 247-252
- [Schechtman06] G. Schechtman, Two observations regarding embedding subsets of Euclidean spaces in normed spaces, *Advances in Mathematics*, **200** (2006), 125 – 135
- [Talagrand96] M. Talagrand, Majorizing measures: the generic chaining, *Ann. Probab.* **24** (1996), 1049 - 1103.

- [Vershynin12] R. Vershynin, Introduction to the non-asymptotic analysis of random matrices. *Compressed Sensing, Theory and Applications*, ed. Y. Eldar and G. Kutyniok, Cambridge University Press, (2012), 210 – 268.
- [KlartagMendelson04] B. Klartag and S. Mendelson, Empirical processes and random projections, *Journal of Functional Analysis*, **225** (2005), 229 – 245.
- [Talagrand05] M. Talagrand, The Generic Chaining, *Springer*, (2005)

VITA

Chandra Vaidyanathan, was born in Karaikudi, Tamilnad, India, on September 21, 1975. After his schooling at Kendriya Vidyalaya, CECRI campus and Muthiah Alagappa Matriculation Higher Secondary School, he obtained his Bachelor's degree in Electronics and Communications Engineering, from Sri Venkateswara College of Engineering in 1996. After a brief stint as research associate at Madras Institute of Technology, Chennai and pursuing his Masters in Power and Controls from Northeastern University, he worked as software development engineer at Analog Devices Inc., Norwood from 2000 to 2002. He received his M.S.E.E. from the University of Notre Dame, with specialisation in Communications and Signal Processing in 2004. He also obtained an M.S. in Applied Mathematics from the University of Notre Dame, in 2006. He received his PhD in Mathematics from the University of Missouri, Columbia, in 2014.