On the study of the operator norms of some random matrices.

Olivier Guédon

Université Paris-Est Marne-la-Vallée

May 18, 2017 - Będlewo

Random operators.

Local theory of Banach spaces : how to produce subspaces with nice properties in a Banach space ?

The Euclidean case

$g_j \sim \mathcal{N}(0,1)$, $g = (g_j)_{j \leq k}$

Take $x \in (\mathbb{R}^k, |\cdot|_2)$ then $\sum x_j g_j \sim (\sum x_j^2)^{1/2} g_1$. That is

$$G(x) = \sum x_j g_j \in L_1$$

$\ell_2^k \overset{1}{\hookrightarrow} L_1$ isometrically.

Random operators.

Local theory of Banach spaces : how to produce subspaces with nice properties in a Banach space ?

The Euclidean case

$g_j \sim \mathcal{N}(0,1)$, $g = (g_j)_{j \leq k}$

Take $x \in (\mathbb{R}^k, |\cdot|_2)$ then $\sum x_j g_j \sim (\sum x_j^2)^{1/2} g_1$. That is

$$G(x) = \sum x_j g_j \in L_1$$

$\ell_2^k \overset{1}{\hookrightarrow} L_1$ isometrically.

More generally, $g_{ij} \sim \mathcal{N}(0,1)$

$$G = (g_{ij}) : \ell_2^k \to (\mathbb{R}^n, \|\cdot\|) = X$$

Find condition on k such that $\ell_2^k \overset{1+\varepsilon}{\hookrightarrow} X$

Dvoretzky-Milman in the '70's, Gordon in the '80's

Random operators.

Local theory of Banach spaces : how to produce subspaces with nice properties in a Banach space ?

The non Euclidean case. $1 \leq p \leq 2$

θ_j p-stable random variables, $\theta = (\theta_j)_{j \leq k}$

Random operators.

Local theory of Banach spaces : how to produce subspaces with nice properties in a Banach space ?

The non Euclidean case. $1 \le p \le 2$

θ_j p-stable random variables, $\theta = (\theta_j)_{j \le k}$ Take $x \in (\mathbb{R}^k, |\cdot|_p)$ then $\sum x_j \theta_j \sim (\sum |x_j|^p)^{1/p} \theta_1$. That is

$$\Theta(x) = \sum x_j \theta_j \in L_1$$

$\ell_p^k \overset{1}{\hookrightarrow} L_1$ isometrically.

Random operators.

Local theory of Banach spaces : how to produce subspaces with nice properties in a Banach space ?

The non Euclidean case. $1 \leq p \leq 2$

θ_j p-stable random variables, $\theta = (\theta_j)_{j \leq k}$

More generally, θ_{ij} p-stable r.v.

$$\Theta = (\theta_{ij}) : \ell_p^k \to (\mathbb{R}^n, \|\cdot\|) = X$$

Find condition on k such that $\ell_p^k \overset{1+\varepsilon}{\hookrightarrow} X$

Random operators.

Local theory of Banach spaces : how to produce subspaces with nice properties in a Banach space ?
The non Euclidean case. $1 \leq p \leq 2$
θ_j p-stable random variables, $\theta = (\theta_j)_{j \leq k}$

More generally, θ_{ij} p-stable r.v.

$$\Theta = (\theta_{ij}) : \ell_p^k \to (\mathbb{R}^n, \| \cdot \|) = X$$

Find condition on k such that $\ell_p^k \overset{1+\varepsilon}{\hookrightarrow} X$

Hum, not so easy....
Not good concentration properties. You have to perturb the random operator,
Pisier '83,
Friedland-G '2013

Sparsity and compressed sensing

Reconstruction of a signal.

Let Φ be an $n \times N$ matrix. You receive $\Phi x := y$ where $x \in \mathbb{R}^N$ is the unknown signal.

How to recover x? Assumption on the signal : it is an m-sparse signal with $m \ll n$.

Restricted Isometry Property with constant $\delta \in (0, 1)$: an $n \times N$ matrix Φ such that for all m-sparse vectors $x \in \mathbb{R}^N$,

$$(1 - \delta)|x|_2 \leq |\Phi x|_2 \leq (1 + \delta)|x|_2.$$

Sparsity and compressed sensing

Reconstruction of a signal.
Let Φ be an $n \times N$ matrix. You receive $\Phi x := y$ where $x \in \mathbb{R}^N$ is the unknown signal.
How to recover x? Assumption on the signal : it is an m-sparse signal with $m \ll n$.

Restricted Isometry Property with constant $\delta \in (0, 1)$: an $n \times N$ matrix Φ such that for all m-sparse vectors $x \in \mathbb{R}^N$,

$$(1 - \delta)|x|_2 \leq |\Phi x|_2 \leq (1 + \delta)|x|_2.$$

Candès, Romberg, Tao ['06]
For such matrices, if δ is small enough, any m-sparse vectors is uniquely defined by

$$\min\{|t|_1 \text{ subject to } \Phi t = y\}$$

Sparsity and compressed sensing

Reconstruction of a signal.
Let Φ be an $n \times N$ matrix. You receive $\Phi x := y$ where $x \in \mathbb{R}^N$ is the unknown signal.
How to recover x ? Assumption on the signal : it is an m-sparse signal with $m \ll n$.

Restricted Isometry Property with constant $\delta \in (0,1)$: an $n \times N$ matrix Φ such that for all m-sparse vectors $x \in \mathbb{R}^N$,

$$(1-\delta)|x|_2 \leq |\Phi x|_2 \leq (1+\delta)|x|_2.$$

Random matrices are good !
$\Phi = (g_{ij})/\sqrt{N}$ satisfies RIP with $m \approx c(\delta)\frac{n}{\log(eN/n)}$.

Question : give an explicit construction of such a matrix ?

Graph theory

Random graph

Classical Erdös-Rény model : n vertices, edges between i and j are taken with probability p

Inhomogeneous Erdös-Rény model : n vertices, edges between i and j with probability p_{ij}

Adjacency matrix

$$Adj = (x_{ij}) \quad \text{and} \quad \mathbb{E}Adj = (p_{ij})$$

Graph theory

Random graph

Classical Erdös-Rény model : n vertices, edges between i and j are taken with probability p

Inhomogeneous Erdös-Rény model : n vertices, edges between i and j with probability p_{ij}

Adjacency matrix

$$Adj = (x_{ij}) \quad \text{and} \quad \mathbb{E}Adj = (p_{ij})$$

How does Adj concentrates around $\mathbb{E}Adj$? Good information to detect communities in the graph.

Graph theory

Random graph

Classical Erdös-Rény model : n vertices, edges between i and j are taken with probability p

Inhomogeneous Erdös-Rény model : n vertices, edges between i and j with probability p_{ij}

Adjacency matrix

$$Adj = (x_{ij}) \quad \text{and} \quad \mathbb{E}Adj = (p_{ij})$$

How does Adj concentrates around $\mathbb{E}Adj$? Good information to detect communities in the graph.

Want to evaluate $\mathbb{E}\|Adj - \mathbb{E}Adj\|_{2\to 2}$? You start with a classical trick

$$\mathbb{E}\|(x_{ij} - \mathbb{E}x_{ij})\| = \mathbb{E}\|(x_{ij} - \mathbb{E}'x'_{ij})\| \leq \mathbb{E}\mathbb{E}'\|(x_{ij} - x'_{ij})\|$$
$$= \mathbb{E}\mathbb{E}'\mathbb{E}_\varepsilon\|(\varepsilon_{ij}(x_{ij} - x'_{ij}))\| \leq 2\mathbb{E}\mathbb{E}_\varepsilon\|(\varepsilon_{ij}x_{ij})\|$$

Operator norms of random matrix

Take a deterministic matrix

$$X = (x_{ij})$$

Let R_i be its rows and C_j be its columns. Then

$$\|X\|_{2\to 2} \geq \max\left(\max_i |R_i|_2, \max_j |C_j|_2\right).$$

Operator norms of random matrix

Take a deterministic matrix

$$X = (x_{ij})$$

Let R_i be its rows and C_j be its columns. Then

$$\|X\|_{2\to 2} \geq \max\left(\max_i |R_i|_2, \max_j |C_j|_2\right).$$

Theorem [Seginer 2000]
If x_{ij} are independent identically distributed random variables then

$$\mathbb{E}\|X\|_{2\to 2} \approx \mathbb{E}\max_i |R_i|_2 + \mathbb{E}\max_j |C_j|_2.$$

Operator norms of random matrix

Take a deterministic matrix

$$X = (x_{ij})$$

Let R_i be its rows and C_j be its columns. Then

$$\|X\|_{2\to 2} \geq \max\left(\max_i |R_i|_2, \max_j |C_j|_2\right).$$

Theorem [Seginer 2000]
If x_{ij} are independent identically distributed random variables then

$$\mathbb{E}\|X\|_{2\to 2} \approx \mathbb{E}\max_i |R_i|_2 + \mathbb{E}\max_j |C_j|_2.$$

More is true : for any $p \leq \log n$,

$$\mathbb{E}\|X\|_{2\to 2}^p \approx \mathbb{E}\max_i |R_i|_2^p + \mathbb{E}\max_j |C_j|_2^p.$$

Operator norms of random matrix

The moment method

$$\|X\|_{2\to 2} = \max_{1\le i\le n} |s_i(X)| \le \left(\sum_{i=1}^{n} |s_i(X)|^p \right)^{1/p} \le n^{1/p} \max_{1\le i\le n} |s_i(X)|$$

Hence for $p \approx \log n$,

$$\|X\|_{2\to 2} \approx \left(\operatorname{tr} (X^* X)^{p/2} \right)^{1/p}$$

Operator norms of random matrix

The moment method

$$\|X\|_{2\to 2} = \max_{1\le i\le n} |s_i(X)| \le \left(\sum_{i=1}^{n} |s_i(X)|^p\right)^{1/p} \le n^{1/p} \max_{1\le i\le n} |s_i(X)|$$

Hence for $p \approx \log n$,

$$\|X\|_{2\to 2} \approx \left(\mathrm{tr}\,(X^*X)^{p/2}\right)^{1/p}$$

We have proven that for $p \approx \log n$,

$$\mathbb{E}\|X\|_{2\to 2}^{p} \approx \mathbb{E}\mathrm{tr}\,(X^*X)^{p/2}$$

And you use combinatorics.

Operator norms of random matrix

The moment method

$$\|X\|_{2\to 2} = \max_{1\le i\le n} |s_i(X)| \le \left(\sum_{i=1}^{n} |s_i(X)|^p\right)^{1/p} \le n^{1/p} \max_{1\le i\le n} |s_i(X)|$$

Hence for $p \approx \log n$,

$$\|X\|_{2\to 2} \approx \left(\mathrm{tr}\,(X^*X)^{p/2}\right)^{1/p}$$

We have proven that for $p \approx \log n$,

$$\mathbb{E}\|X\|_{2\to 2}^p \approx \mathbb{E}\mathrm{tr}\,(X^*X)^{p/2}$$

And you use combinatorics.
Not enough. You need to use the assumption identically distributed. Hence you add an additional randomness by allowing random permutation on each coordinates.

Operator norms of random matrix

A simple case of independent not identically distributed r.v. :

Theorem [Seginer 2000]

If $x_{ij} = \varepsilon_{ij} a_{ij}$ where ε_{ij} are independent random variables then

$$\mathbb{E}\|X\|_{2\to 2} \leq (\log n)^{1/4} \left(\max_i \sqrt{\sum_j a_{ij}^2} + \max_j \sqrt{\sum_i a_{ij}^2} \right).$$

Operator norms of random matrix

A simple case of independent not identically distributed r.v. :

Theorem [Seginer 2000]

If $x_{ij} = \varepsilon_{ij} a_{ij}$ where ε_{ij} are independent random variables then

$$\mathbb{E}\|X\|_{2\to 2} \le (\log n)^{1/4} \left(\max_i \sqrt{\sum_j a_{ij}^2} + \max_j \sqrt{\sum_i a_{ij}^2} \right).$$

He has an example where $(\log n)^{1/4}$ is necessary.

Operator norms of random matrix

A simple case of independent not identically distributed r.v. :

Theorem [Seginer 2000]

If $x_{ij} = \varepsilon_{ij} a_{ij}$ where ε_{ij} are independent random variables then

$$\mathbb{E}\|X\|_{2\to 2} \leq (\log n)^{1/4} \left(\max_i \sqrt{\sum_j a_{ij}^2} + \max_j \sqrt{\sum_i a_{ij}^2} \right).$$

He has an example where $(\log n)^{1/4}$ is necessary.

Conjecture

$$\mathbb{E}\|X\|_{2\to 2} \approx \mathbb{E}\max_i |R_i|_2 + \mathbb{E}\max_j |C_j|_2.$$

is not valid in general.

The Gaussian case

We have $$\mathbb{E}\|(\varepsilon_{ij}a_{ij})\|_{2\to 2} \le \sqrt{\frac{2}{\pi}}\mathbb{E}\|(g_{ij}a_{ij})\|_{2\to 2}$$

The Gaussian case

We have
$$\mathbb{E}\|(\varepsilon_{ij}a_{ij})\|_{2\to 2} \le \sqrt{\frac{2}{\pi}}\mathbb{E}\|(g_{ij}a_{ij})\|_{2\to 2}$$

Theorem [Latała 2005]

$$\mathbb{E}\|(g_{ij}a_{ij})\|_{2\to 2} \lesssim \max_i \sqrt{\sum_j a_{ij}^2} + \max_j \sqrt{\sum_i a_{ij}^2} + \left(\sum_{ij} a_{ij}^4\right)^{1/4}$$

The Gaussian case

We have
$$\mathbb{E}\|(\varepsilon_{ij}a_{ij})\|_{2\to 2} \leq \sqrt{\frac{2}{\pi}}\mathbb{E}\|(g_{ij}a_{ij})\|_{2\to 2}$$

Theorem [Latała 2005]

$$\mathbb{E}\|(g_{ij}a_{ij})\|_{2\to 2} \lesssim \max_i \sqrt{\sum_j a_{ij}^2} + \max_j \sqrt{\sum_i a_{ij}^2} + \left(\sum_{ij} a_{ij}^4\right)^{1/4}$$

Theorem [Bandeira - Van Handel 2016]

$$\mathbb{E}\|(g_{ij}a_{ij})\|_{2\to 2} \leq \max_i \sqrt{\sum_j a_{ij}^2} + \max_j \sqrt{\sum_i a_{ij}^2} + C\sqrt{\log n}\max|a_{ij}|$$

where C is a universal constant

The Gaussian case

Conjecture. In the Gaussian case, $X = (g_{ij}a_{ij})$, we have

$$\mathbb{E}\|X\|_{2\to 2} \approx \mathbb{E}\max_i |R_i|_2 + \mathbb{E}\max_j |C_j|_2.$$

The Gaussian case

Conjecture. In the Gaussian case, $X = (g_{ij}a_{ij})$, we have

$$\mathbb{E}\|X\|_{2\to 2} \approx \mathbb{E}\max_i |R_i|_2 + \mathbb{E}\max_j |C_j|_2.$$

Lemma [Van Handel 2017]

$$\mathbb{E}\max_i \sqrt{\sum_j a_{ij}^2 g_{ij}^2} + \mathbb{E}\max_j \sqrt{\sum_i a_{ij}^2 g_{ij}^2}$$

$$\leq \max_i \sqrt{\sum_j a_{ij}^2} + \max_j \sqrt{\sum_i a_{ij}^2} + C\,\mathbb{E}\max |a_{ij}g_{ij}|$$

The Gaussian case

Conjecture. In the Gaussian case, $X = (g_{ij}a_{ij})$, we have

$$\mathbb{E}\|X\|_{2\to 2} \approx \mathbb{E}\max_i |R_i|_2 + \mathbb{E}\max_j |C_j|_2.$$

Lemma [Van Handel 2017]

$$\mathbb{E}\max_i \sqrt{\sum_j a_{ij}^2 g_{ij}^2} + \mathbb{E}\max_j \sqrt{\sum_i a_{ij}^2 g_{ij}^2}$$

$$\leq \max_i \sqrt{\sum_j a_{ij}^2} + \max_j \sqrt{\sum_i a_{ij}^2} + C\,\mathbb{E}\max |a_{ij}g_{ij}|$$

This would fit with Seginer's example.

The Gaussian case

Theorem [Bandeira - Van Handel 2016]

$$\mathbb{E}\|(g_{ij}a_{ij})\|_{2\to 2} \leq \max_i \sqrt{\sum_j a_{ij}^2} + \max_j \sqrt{\sum_i a_{ij}^2} + C\sqrt{\log n} \max |a_{ij}|$$

where C is a universal constant
They use the moment method.

The Gaussian case

Maybe not the good strategy.

$$\mathbb{E}\|X\|_{2\to 2} \le \mathbb{E}\left(\operatorname{tr}(X^*X)^{p/2}\right)^{1/p} \le \left(\mathbb{E}\operatorname{tr}(X^*X)^{p/2}\right)^{1/p}$$

and you take $p \approx \log n$.

The Gaussian case

Maybe not the good strategy.

$$\mathbb{E}\|X\|_{2\to 2} \le \mathbb{E}\left(\mathrm{tr}\,(X^*X)^{p/2}\right)^{1/p} \le \left(\mathbb{E}\,\mathrm{tr}\,(X^*X)^{p/2}\right)^{1/p}$$

and you take $p \approx \log n$. Example :

$$A = (a_{ij}) = \mathrm{diag}(a_1, \dots, a_n)$$

Hence

$$\left(\mathbb{E}\mathrm{tr}\,(X^*X)^{p/2}\right)^{1/p} = \left(\mathbb{E}\sum_{i=1}^{n} |a_i|^p |g_i|^p\right)^{1/p} \approx \sqrt{p}\left(\sum_{i=1}^{n} |a_i|^p\right)^{1/p}$$

And then, you take $p \approx \log n$, which gives $\sqrt{\log n}\max|a_{ij}|$
While obviously

$$\mathbb{E}\|\mathrm{diag}(a_1 g_1, \dots, a_n g_n)\| = \mathbb{E}\max|a_i g_i|.$$

Other strategy.

[Van Handel 2017] A strategy based on some comparison of Gaussian processes.

Conjecture is proved when (a_{ij}^2) is a symmetric positive matrix !

But not in general.

Other strategy.

[Van Handel 2017] A strategy based on some comparison of Gaussian processes.
Conjecture is proved when (a_{ij}^2) is a symmetric positive matrix !
But not in general.

Other operator norms.
$X = (a_{ij}g_{ij}) : \ell_{p^*}^n \to \ell_q^n$ where $1 < p^* \leq 2 \leq q$.

$$\|X\|_{p^* \to q} \geq \max\left(\max_i |R_i|_p, \max_j |C_j|_q\right)$$

Other strategy.

[Van Handel 2017] A strategy based on some comparison of Gaussian processes.

Conjecture is proved when (a_{ij}^2) is a symmetric positive matrix !

But not in general.

Other operator norms.

$X = (a_{ij} g_{ij}) : \ell_{p^*}^n \to \ell_q^n$ where $1 < p^* \leq 2 \leq q$.

$$\|X\|_{p^* \to q} \geq \max \left(\max_i |R_i|_p, \max_j |C_j|_q \right)$$

Conjecture. Let p be the conjugate of p^* then

$$\mathbb{E}\|X\|_{p^* \to q} \approx_{p,q} \mathbb{E} \max_i |R_i|_p + \mathbb{E} \max_j |C_j|_q$$

$$\approx_{p,q} \max_i \left(\sum_j |a_{ij}|^p \right)^{1/p} + \max_j \left(\sum_i |a_{ij}|^q \right)^{1/q} + \mathbb{E} \max |a_{ij} g_{ij}|$$

Other strategy - Other operator norms

Theorem [G-Hinrichs-Litvak-Prochno]. We have

$$\mathbb{E}\|X\|_{p^*\to q} \lesssim_{p,q} (\log n)^{1/q} \left(\max_i \left(\sum_j |a_{ij}|^p \right)^{1/p} + \mathbb{E} \max |a_{ij} g_{ij}| \right.$$
$$\left. + \max_j \left(\sum_i |a_{ij}|^q \right)^{1/q} \right).$$

Other strategy - Other operator norms

Theorem [G-Hinrichs-Litvak-Prochno]. We have

$$\mathbb{E}\|X\|_{p^*\to q} \lesssim_{p,q} (\log n)^{1/q} \left(\max_i \left(\sum_j |a_{ij}|^p \right)^{1/p} + \mathbb{E} \max |a_{ij} g_{ij}| \right) + \max_j \left(\sum_i |a_{ij}|^q \right)^{1/q} .$$

Observe that

$$\mathbb{E}\|X\|_{p^*\to q} = \mathbb{E} \sup_{y \in B^n_{p^*}} \left(\sum_i |\langle R_i, y\rangle|^q \right)^{1/q} \leq \left(\mathbb{E} \sup_{y \in B^n_{p^*}} \sum_i |\langle R_i, y\rangle|^q \right)^{1/q}$$

$$\leq \left(\mathbb{E} \sup_{y \in B^n_{p^*}} \sum_i |\langle R_i, y\rangle|^q - \mathbb{E}|\langle R_i, y\rangle|^q \right)^{1/q} + \sup_{y \in B^n_{p^*}} \left(\mathbb{E}|\langle R_i, y\rangle|^q \right)^{1/q} .$$

Other strategy - Other operator norms

$$\mathbb{E}\|X\|_{p^*\to q} \leq$$

$$\leq \left(\mathbb{E} \sup_{y\in B^n_{p^*}} \sum_i |\langle R_i, y\rangle|^q - \mathbb{E}'|\langle R_i', y\rangle|^q\right)^{1/q} + \sup_{y\in B^n_{p^*}} \left(\mathbb{E}|\langle R_i, y\rangle|^q\right)^{1/q}.$$

$$\leq \left(\mathbb{E}\mathbb{E}' \sup_{y\in B^n_{p^*}} \sum_i |\langle R_i, y\rangle|^q - |\langle R_i', y\rangle|^q\right)^{1/q} + \sup_{y\in B^n_{p^*}} \left(\mathbb{E}|\langle R_i, y\rangle|^q\right)^{1/q}.$$

$$\leq \left(2\mathbb{E}\mathbb{E}_\varepsilon \sup_{y\in B^n_{p^*}} \sum_i \varepsilon_i |\langle R_i, y\rangle|^q\right)^{1/q} + \sup_{y\in B^n_{p^*}} \left(\mathbb{E}|\langle R_i, y\rangle|^q\right)^{1/q}.$$

How to evaluate this first term ?

Non commutative Khinchine inequality

The case p=q=2.

$$\mathbb{E}_\varepsilon \sup_{y \in B_2^n} \sum_{i=1}^n \varepsilon_i |\langle R_i, y\rangle|^2 = \mathbb{E}_\varepsilon \left\| \sum_{i=1}^n \varepsilon_i R_i \otimes R_i \right\|_{2\to 2}$$

$$\lesssim \sqrt{\log n} \left(\left\| \sum_{i=1}^n |R_i|_2^2 \, R_i \otimes R_i \right\|_{2\to 2} \right)^{1/2}$$

$$\lesssim \sqrt{\log n} \max_i |R_i|_2 \left(\left\| \sum_{i=1}^n R_i \otimes R_i \right\|_{2\to 2} \right)^{1/2}$$

Non commutative Khinchine inequality

The case p=q=2.

$$\mathbb{E}_\varepsilon \sup_{y\in B_2^n} \sum_{i=1}^n \varepsilon_i |\langle R_i, y\rangle|^2 = \mathbb{E}_\varepsilon \left\| \sum_{i=1}^n \varepsilon_i R_i \otimes R_i \right\|_{2\to 2}$$

$$\lesssim \sqrt{\log n} \left(\left\| \sum_{i=1}^n |R_i|_2^2 \, R_i \otimes R_i \right\|_{2\to 2} \right)^{1/2}$$

$$\lesssim \sqrt{\log n} \max_i |R_i|_2 \left(\left\| \sum_{i=1}^n R_i \otimes R_i \right\|_{2\to 2} \right)^{1/2}$$

You have bounded a quantity by its square root !

Some generalization.

Theorem [G-Rudelson '07, G-Mendelson-Pajor-Tomczak '08]

$$\mathbb{E} \sup_{y \in B^n_{p^*}} \left| \sum_{i=1}^{n} \varepsilon_i |\langle R_i, y\rangle|^q \right| \lesssim_p$$

$$\sqrt{\log n} \max_{1 \le i \le n} |R_i|_p \sup_{y \in B^n_{p^*}} \left(\sum_{i=1}^{n} |\langle R_i, y\rangle|^{2(q-1)} \right)^{1/2} .$$

Some generalization.

Theorem [G-Rudelson '07, G-Mendelson-Pajor-Tomczak '08]

$$\mathbb{E} \sup_{y \in B_{p^*}^n} \left| \sum_{i=1}^n \varepsilon_i |\langle R_i, y \rangle|^q \right| \lesssim_p$$

$$\sqrt{\log n} \max_{1 \le i \le n} |R_i|_p \sup_{y \in B_{p^*}^n} \left(\sum_{i=1}^n |\langle R_i, y \rangle|^{2(q-1)} \right)^{1/2}.$$

Valid because $q \geq 2$, $1 < p^* \leq 2$ and $\ell_{p^*}^n$ has modulus of convexity of power type 2.

Some generalization.

Theorem [G-Rudelson '07, G-Mendelson-Pajor-Tomczak '08]

$$\mathbb{E} \sup_{y \in B_{p^*}^n} \left| \sum_{i=1}^{n} \varepsilon_i |\langle R_i, y\rangle|^q \right| \lesssim_p$$

$$\sqrt{\log n} \max_{1 \le i \le n} |R_i|_p \sup_{y \in B_{p^*}^n} \left(\sum_{i=1}^{n} |\langle R_i, y\rangle|^{2(q-1)} \right)^{1/2} .$$

Valid because $q \geq 2$, $1 < p^* \leq 2$ and $\ell_{p^*}^n$ has modulus of convexity of power type 2.

Use concentration of Lipschitz function of a Gaussian vector, and compute all parameters