Ecole d'Automne
UMLV
November 16-20
2009

Empirical methods
and
selection of characters

I) Reconstruction of signals with small support by random methods.

Let $y \in \mathbb{R}^N$ (or $\mathbb{C}^N$) be an unknown signal.

We receive Φy with Φ an $n \times N$ matrix

i.e. $\Phi = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$, $\Phi y = (\langle Y_i, y \rangle)_{1 \leq i \leq n}$.

with $n \ll N$.

We know that y has a small support in the canonical basis chosen at the beginning i.e. $|\mathrm{supp}\, y| \leq m$. We also say that $y \in \Sigma_m$ is m-sparse.

Problem: what are the conditions on Φ, m, n and N such that the solution of the problem

$$(P) \qquad \min_{t \in \mathbb{R}^N} \{ |t|_1, \ \Phi y = \Phi t \}$$

is unique and equal to y.

Proposition: For every $y \in \Sigma_m$, the solution of (P) is unique and equal to y

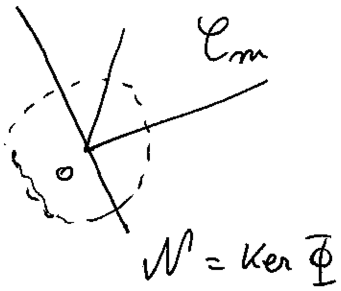
iff

$\forall h \in \mathrm{Ker}\, \Phi = \mathcal{N}$, $h \neq 0$
$\forall I \subset [N]$, $\# I \leq m$, $|h_I|_1 < |h_{I^c}|_1$.

Let $\mathcal{C}_m$ be the cone

$$\mathcal{C}_m = \{ h \in \mathbb{R}^N, \ \exists I \subset [N] \text{ with } \# I \leq m, \ |h_{I^c}|_1 \leq |h_I|_1 \}$$

This condition is equivalent to $\mathcal{N} \cap \mathcal{C}_m = \{0\}$.



Conclusion: when $x \in \Sigma_m$, the solution of (P) is unique and equal to x iff $\mathcal{N} \cap \mathcal{C}_m \cap S^{N-1} = \emptyset$.

Remark: if $t \in \mathcal{C}_m \cap S^{N-1}$

then $$|t|_1 = \sum_{i=1}^{N} |t_i| = \sum_{i \in I} |t_i| + \sum_{i \in I^c} |t_i|$$

$$\leq 2 \sum_{i \in I} |t_i| \leq 2\sqrt{m}$$

so $\mathcal{C}_m \cap S^{N-1} \subset 2\sqrt{m}\, B_1^N \cap S^{N-1}$

We will study the following sufficient condition:

If $\boxed{\mathcal{N} \cap 2\sqrt{m}\, B_1^N \cap S^{N-1} = \emptyset}$

then the solution of (P) is unique and equal to x.

This condition is equivalent to $\boxed{\operatorname{diam}(\mathcal{N} \cap B_1^N) < \frac{1}{2\sqrt{m}}}$

where the diameter is taken with respect to the Euclidean distance

1) Local theory of Banach spaces.

Gelfand numbers : $u : X \to Y$

$$c_k(u) = \inf \left\{ |||u_{|S}|||, \quad S \subset X, \ \text{codim } S < k \right\}$$

$$= \inf_{S} \ \sup_{\substack{x \in S \\ \|x\|_X \leq 1}} \|u(x)\|_Y$$

Let's take $u = id : \ell_1^N \to \ell_2^N$ then $c_k(u) = \inf_{\text{codim} S < k} \ \sup_{\substack{x \in S \\ |x|_1 \leq 1}} |x|_2$

$$= \inf_{\text{codim} S < k} \text{diam}(S \cap B_1^N)$$

Lot of work in the 80's.

Garnaev - Gluskin '84 : $c_k\left(id : \ell_1^N \to \ell_2^N\right) \approx \min \left\{ 1, \sqrt{\frac{\log \frac{N}{k}}{k}} \right\}$

and is "attained" for $S = \text{Ker } \Phi$

where $\Phi : \mathbb{R}^N \to \mathbb{R}^k$, $\Phi = (g_{ij})$ with $g_{ij} \sim \mathcal{N}(0,1)$

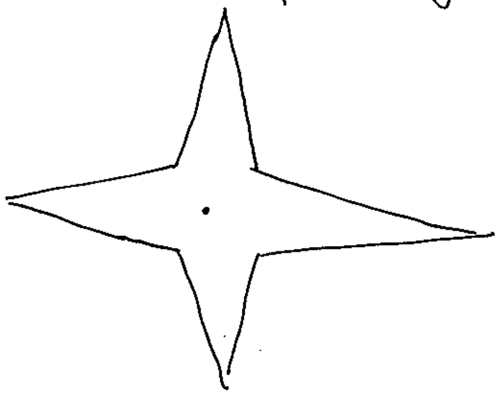
Immediate corollary : if $\Phi = (g_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq N}} : \mathbb{R}^N \to \mathbb{R}^n$

and if $m \approx \dfrac{n}{\log \frac{N}{m}}$ i.e. $n \approx m \log \frac{N}{m}$

then the solution of (P) is unique and equal to f.

- How to study the diameter of a section by a subspace of a star shape body :

Let T be a star shape body with respect to the origin



diam $(\mathcal{N} \cap T)$?

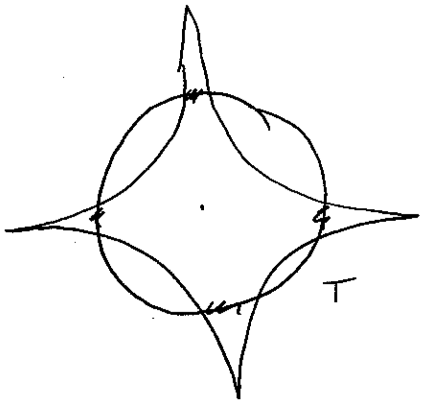
where $\mathcal{N} = \operatorname{Ker}\phi \subset \mathbb{R}^N$

Proposition: if $\inf_{y \in T \cap \rho S^{N-1}} \sum_{i=1}^{n} \langle Y_i, y \rangle^2 > 0$

then diam $(T \cap \operatorname{Ker} \Phi) \leq \rho$

where $\Phi = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} \in M_{n \times N}$.

Proof:



. If $y \in T \cap \rho S^{N-1}$ then $|\Phi y|_2^2 > 0$ so $y \notin \operatorname{Ker}\Phi$.

. Since T is star shaped, if $y \in T$ and $|y|_2 \geq \rho$ then $\frac{\rho y}{|y|_2} \in T \cap \rho S^{N-1}$

so $y \notin \operatorname{Ker}\Phi$.

↳ Pajor - Tomczak Jaegermann : Gelfand numbers and low M^*- estimate.

2) Random methods to study $\operatorname{Ker} \Phi \cap B_1^N$

How to find ρ such that $\inf_{y \in T \cap \rho S^{N-1}} \sum_{i=1}^{n} \langle Y_i, y \rangle^2 > 0$?

. Let $\varphi_1, \dots, \varphi_N$ be an orthonormal basis of ℓ_2^N such that $\forall i, |\varphi_i|_\infty \leq \frac{K}{\sqrt{N}}$.

Main examples : Discrete Fourier system , Walsh system

(i.e. for example $N = 2^p$

$$W_p = \frac{1}{\sqrt{2}} \begin{pmatrix} W_{p-1} & W_{p-1} \\ -W_{p-1} & W_{p-1} \end{pmatrix} , \quad W_0 = 1$$

$\uparrow$ matrix of size 2^p

The matrix that we get is a matrix with entries $\pm \frac{1}{\sqrt{N}}$ and the column vectors of the matrix form an orthonormal basis of ℓ_2^N. } (exercise).

(Fourier : ~~Φ~~ $\varphi_{ij} = \frac{1}{\sqrt{N}} \exp\left(-i \frac{2\pi ij}{N}\right), 1 \leq i,j \leq N$).

. First definition of the random vector : $Y = \varphi_i$ with proba $\frac{1}{N}$ and let $Y_1, \dots, Y_n$ be independent copies of Y

↳ properties : 1) $\mathbb{E} \langle Y, y \rangle^2 = \frac{1}{N} \sum_{i=1}^{N} \langle \varphi_i, y \rangle^2 = \frac{1}{N} |y|_2^2$

2) Let $\Phi = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$ then $\mathbb{E} |\Phi y|_2^2 = \frac{n}{N} |y|_2^2$.

We will study

$$(*)\quad \mathbb{E} \sup_{y\in T\cap \rho S^{N-1}} \left| \sum_{i=1}^{n} \langle Y_i, y\rangle^2 - \frac{n\rho^2}{N}\right| \overset{?}{\leq} \frac{2}{3}\,\frac{n\rho^2}{N}$$

If $(*)$ is ~~true~~ valid then $\exists$ a choice of $(Y_i)_{1\leq i\leq n}$ such that

$$\sup_{y\in T\cap \rho S^{N-1}} \left| \sum_{i=1}^{n} \langle Y_i, y\rangle^2 - \frac{n\rho^2}{N}\right| \leq \frac{2}{3}\,\frac{n\rho^2}{N}$$

hence $\forall y\in T\cap\rho S^{N-1}$, $\displaystyle\sum_{i=1}^{n} \langle Y_i, y\rangle^2 \geq \frac{1}{3}\,\frac{n\rho^2}{N} > 0$

and we have solved our problem.

<u>rk</u>: $\frac{2}{3}$ can be replaced by any number < 1.

. 2nd <u>definition of randomness</u>:

Let δ_i be i.i.d random variables with $\delta_i = 1$ with proba δ and $\delta_i = 0$ with proba $(1-\delta)$.

We start from the orthogonal matrix $\begin{pmatrix}\varphi_1\\ \vdots\\ \varphi_N\end{pmatrix}$ and we "select" randomly ~~some~~ rows

i.e. $\Phi(\omega) = \begin{pmatrix}\delta_1\varphi_1\\ \vdots\\ \delta_N\varphi_N\end{pmatrix}$ ← and you "delete" the zero lines.

Then you study

$$(**)\quad \mathbb{E} \sup_{y\in T\cap \rho S^{N-1}} \left| \sum_{i=1}^{N} \delta_i \langle \varphi_i, y\rangle^2 - \delta\rho^2\right| \overset{?}{\leq} \frac{2}{3}\delta\rho^2$$

↳ and you are done.

$\boxed{\text{cf page 10 : THM}}$

3) Empirical processes.

Let $Y_1, \dots, Y_n$ be independent copies of a random vector Y

Let ~~f~~ $\mathcal{F}$ be a class of functionals on these vectors

Theorem 1: symmetrization principle

Let $\varepsilon_1, \dots, \varepsilon_n$ be ~~an~~ iid r.v. with $\mathbb{P}(\varepsilon_i = +1) = \frac{1}{2}$ and $\mathbb{P}(\varepsilon_i = -1) = 1/2$.

Then for a countable class of functions $\mathcal{F}$ ~~f~~,

(1) $$\mathbb{E} \sup_{f \in \mathcal{F}} \left| \sum_{i=1}^n f(Y_i) - \mathbb{E} f(Y_i) \right| \leq 2 \, \mathbb{E} \mathbb{E}_\varepsilon \sup_{f \in \mathcal{F}} \left| \sum_{i=1}^n \varepsilon_i f(Y_i) \right|$$

(2) $$\mathbb{E} \sup_{f \in \mathcal{F}} \sum_{i=1}^n |f(Y_i)| \leq \sup_{f \in \mathcal{F}} \mathbb{E} |f(Y_i)| + 2 \, \mathbb{E} \mathbb{E}_\varepsilon \sup_{f \in \mathcal{F}} \left| \sum_{i=1}^n \varepsilon_i f(Y_i) \right|$$

(3) If $\mathbb{E} f(Y_i) = 0$ then
$$\mathbb{E} \mathbb{E}_\varepsilon \sup_{f \in \mathcal{F}} \left| \sum_{i=1}^n \varepsilon_i f(Y_i) \right| \leq 2 \, \mathbb{E} \sup_{f \in \mathcal{F}} \left| \sum_{i=1}^n f(Y_i) \right|$$

Proof. (1) ok.

(2) Apply (1) for $|f|$, triangle inequality and contraction principle to conclude that
$$\mathbb{E}_\varepsilon \sup_{f \in \mathcal{F}} \left| \sum \varepsilon_i |f(Y_i)| \right| \leq \mathbb{E}_\varepsilon \sup_{f \in \mathcal{F}} \left| \sum \varepsilon_i f(Y_i) \right|.$$
($\varphi(t) = |t|$ is 1-lipschitzienne).

(3) To prove (3), work conditionally on $(\varepsilon_i)_{i=1}^n$.

Let $I = \{ i \, ; \, \varepsilon_i = 1 \}$

Then we have

$$\mathbb{E}\mathbb{E}_{\varepsilon} \sup_{f\in\mathcal{F}} \left|\sum_{i=1}^{m} \varepsilon_i f(X_i)\right| \leq \mathbb{E}\mathbb{E}_{\varepsilon} \sup_{f\in\mathcal{F}} \left|\sum_{i\in I} f(X_i)\right| + \mathbb{E}\mathbb{E}_{\varepsilon} \sup_{f\in\mathcal{F}} \left|\sum_{i\in I^c} f(X_i)\right|$$

But $\mathbb{E} f(X_i) = 0$ so by Jensen,

$$\mathbb{E}_{x} \sup_{f\in\mathcal{F}} \left|\underbrace{\sum_{i\in I} f(X_i)}_{\left|\sum_{i\in I} f(X_i) + \mathbb{E}\sum_{i\in I^c} f(X_i)\right|}\right| \leq \mathbb{E} \sup_{f\in\mathcal{F}} \left|\sum_{i=1}^{m} f(X_i)\right|$$

And (3) is proved $\square$.

~~[illegible]~~ Proposition 2: for any countable set T,

$$\mathbb{E} \sup_{t\in T} \left|\sum_{i=1}^{m} \varepsilon_i t_i\right| \leq \sqrt{\frac{\pi}{2}}\, \mathbb{E} \sup_{t\in T} \left|\sum_{i=1}^{m} g_i t_i\right|$$

where g_i are iid random $\mathcal{N}(0,1)$ variables.

proof: $g_i \sim \varepsilon_i |g_i|$

So $$\mathbb{E}_{\varepsilon}\mathbb{E}_{g} \sup_{t\in T} \left|\sum \varepsilon_i |g_i| t_i\right| \geq \mathbb{E}_{\varepsilon} \sup_{t\in T} \left|\mathbb{E}_{g} \sum \varepsilon_i |g_i| t_i\right|$$

$$= \sqrt{\frac{2}{\pi}}\, \mathbb{E}_{\varepsilon} \sup_{t\in T} \left|\sum \varepsilon_i t_i\right| \square$$

Conclusion:

$$\mathbb{E} \sup_{y\in T\cap\rho S^{N-1}} \left|\sum_{i=1}^{m} \langle Y_i, y\rangle^2 - \mathbb{E}\langle Y_i, y\rangle^2\right| \leq 2\,\mathbb{E}\mathbb{E}_{\varepsilon} \sup_{y\in T\cap\rho S^{N-1}} \left|\sum_{i=1}^{m} \varepsilon_i \langle Y_i, y\rangle^2\right|$$

$$\leq \sqrt{2\pi}\, \mathbb{E}\mathbb{E}_{g} \sup_{y\in T\cap\rho S^{N-1}} \left|\sum_{i=1}^{m} g_i \langle Y_i, y\rangle^2\right|.$$

Theorem (Rudelson '97) :

$$\mathbb{E}_{\varepsilon} \sup_{y \in S^{N-1}} \left| \sum_{i=1}^{n} \underset{\substack{\text{or}\\ g_i}}{\varepsilon_i} \langle Y_i , y \rangle^2 \right| \lesssim \sqrt{\log n} \cdot \max_{1 \le i \le n} |Y_i|_2 \cdot \sup_{y \in S^{N-1}} \left(\sum_{i=1}^{n} \langle Y_i , y \rangle^2 \right)^{1/2} .$$

for any fixed vectors $Y_1, \dots, Y_n$.

proof: $\sup_{y \in S^{N-1}} \sum_{i=1}^{n} \varepsilon_i \langle Y_i , y \rangle^2 = \sup_{y \in S^{N-1}} \langle \sum_{i=1}^{n} \varepsilon_i \langle Y_i , y \rangle Y_i \ , y \rangle$

Let $T_i : \mathbb{R}^N \to \mathbb{R}^N$

$y \mapsto \langle Y_i , y \rangle Y_i$

hence

$$\sup_{y \in S^{N-1}} \left| \sum_{i=1}^{n} \varepsilon_i \langle Y_i , y \rangle^2 \right| = \sup_{y \in S^{N-1}} |\langle \sum \varepsilon_i T_i \ y , y \rangle|$$

$$= \left|\left|\left| \sum_{i=1}^{n} \varepsilon_i T_i \right|\right|\right|_{\ell_2^N \to \ell_2^N}$$

$$= \left\| \sum \varepsilon_i T_i \right\|_{S_\infty^N} = \sup_{1 \le i \le N} |\lambda_i (\ )|$$

But

$$\sup_{1 \le i \le N} |\lambda_i| \le \Big(\overset{\substack{\text{because rk} \le n\\ (n)\\ \downarrow}}{\sum_{i=1}^{N}} |\lambda_i|^q \Big)^{1/q} := \left\| \sum \varepsilon_i T_i \right\|_{S_q^N} \le N^{1/q} \sup \lambda_i$$

$= e \sup \lambda_i$ for $q = \ln N$

Khinchine inequality for S_q^N (Lust Piquard - Pisier '86)

$$\mathbb{E} \left\| \sum \varepsilon_i T_i \right\|_{S_q^N} \lesssim \sqrt{q} \cdot \max \left\{ \left\| \left(\sum T_i^* T_i \right)^{1/2} \right\|_{S_q^N} , \left\| \left(\sum T_i T_i^* \right)^{1/2} \right\|_{S_q^N} \right\} .$$

But $T_i T_i^* = T_i^* T_i : y \mapsto |Y_i|^2 \langle Y_i , y \rangle Y_i$

$= |Y_i|^2 T_i$

And

$$\left\| \left(\sum |Y_i|^2 T_i \right)^{1/2} \right\|_{S_q^N} \le e \left\| \sum |Y_i|^2 T_i \right\|_{S_\infty^N}^{1/2}$$

$$\le e \max_{1 \le i \le n} |Y_i| \ \left\| \sum T_i \right\|_{S_\infty^N}^{1/2} \qquad \square$$

Let's come back to our Gaussian process:

$$X_y = \sum_{i=1}^{n} g_i \langle Y_i, y\rangle^2 = \sum_{i=1}^{n} g_i f^2(Y_i)$$

$$d(y,\bar y)^2 = \mathbb{E}|X_y - X_{\bar y}|^2 = \sum_{i=1}^{n} \overset{\cancel{f(Y_i) - \bar f^2(Y_i)}}{\left(\langle Y_i, y\rangle^2 - \langle Y_i,\bar y\rangle^2\right)^2}$$

$$= \sum_{i=1}^{n} \overset{(f-\bar f)(Y_i)^2}{\langle Y_i, y-\bar y\rangle^2} \overset{(f(Y_i)+\bar f(Y_i))^2}{\left(\langle Y_i, y\rangle + \langle Y_i, \bar y\rangle\right)^2}$$

$$\leq 2 \sum_{i=1}^{n} \overset{(f-\bar f)(Y_i)^2}{\langle Y_i, y-\bar y\rangle^2} \left(\overset{f^2(Y_i)}{\langle Y_i, y\rangle^2} + \overset{\bar f(Y_i)^2}{\langle Y_i,\bar y\rangle^2}\right)$$

Inequality (1): $d(y,\bar y) \leq 2 \sup_y \left(\sum_{i=1}^{n} \overset{\cancel{f^2(Y_i)}}{\langle Y_i, y\rangle^2}\right)^{1/2} \cdot \max_{1\leq i\leq n} \left|\overset{\cancel{(f-\bar f)(Y_i)}}{\langle Y_i, y-\bar y\rangle}\right|$

Therefore: $\mathbb{E}_g \sup_{\cancel{f}\in \cancel{F}} \sum g_i \cancel{\langle Y_i, y\rangle} \cancel{f^2(Y_i)} \langle Y_i, y\rangle^2$

$$\leq \sup_{\cancel{f}\in B} \left(\sum_{i=1}^{n} \overset{\cancel{f^2(Y_i)}}{\langle Y_i, y\rangle^2}\right)^{1/2} \gamma_2(\cancel{F}, d_{\infty,n})$$

where $d_{\infty,n}(z) = \max_{1\leq i\leq n} \overset{\cancel{|f(Y_i)|, \forall f\in F}}{|\langle Y_i, z\rangle|}, \forall z \in \mathbb{R}^N$

Main Theorem: Let $Y = \varphi_i$ with proba $\frac{1}{N}$ where $(\varphi_1, \dots, \varphi_N)$ o.n.b. of ℓ_2^N

such that $\forall i, |\varphi_i|_\infty \leq \frac{K}{\sqrt{N}}$

Let $Y_1, \dots, Y_n$ be ind. copies of Y

If $m \leq \frac{n}{\log N (\log n)^3}$ (i.e. $n \geq m (\log N)(\log(m \log N))^3$)

then with proba $\geq$

$\Phi = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$ is such that the problem (P) has a unique solution equal to U

Remarks:

Remarks: ① Candès - Tao proved $n \gtrsim m (\log N)^6$
IEEE 2006

② Rudelson - Vershynin proved $n \gtrsim m \log N \cdot \log(m \log N) (\log m)^2$
Communications on Pure and Applied Math. '2008.

Key theorem: Let Y be a random vector in $\mathbb{R}^N$, $Y_1, \dots, Y_n$ be ind. copies of Y

$$d_{\infty,n}(z, \bar{z}) = \max_{1 \le i \le n} |\langle Y_i, z - \bar{z} \rangle| .$$

For any set B,

$$\mathbb{E} \sup_{y \in B} \left| \sum_{i=1}^{n} \langle Y_i, y \rangle^2 - \mathbb{E} \langle Y_i, y \rangle^2 \right| \lesssim \max \left(\sqrt{n}\, \sigma_B U_n, U_n^2 \right)$$

where $U_n = \left(\mathbb{E}\, \gamma_2^2 (B, d_{\infty,n}) \right)^{1/2}$ and $\sigma_B = \sup_{y \in B} \left(\mathbb{E} \langle Y, y \rangle^2 \right)^{1/2}$

proof: we start with symmetrization

Let $A := \mathbb{E} \sup_{y \in B} \left| \sum_{i=1}^{n} \langle Y_i, y \rangle^2 - \mathbb{E} \langle Y_i, y \rangle^2 \right|$

then

$$A \lesssim 2\, \mathbb{E}\, \mathbb{E}_\varepsilon \sup_{y \in B} \sum \varepsilon_i \langle Y_i, y \rangle^2$$

$$\lesssim c\, \mathbb{E} \sup_{y \in B} \left(\sum \langle Y_i, y \rangle^2 \right)^{1/2} \gamma_2 (B, d_{\infty,n}) \quad \text{by inequality (1)}$$

$$\le c \left(\mathbb{E}\, \gamma_2^2 (B, d_{\infty,n}) \right)^{1/2} \left(\mathbb{E} \sup_{y \in B} \sum_{i=1}^{n} \langle Y_i, y \rangle^2 - \mathbb{E}\langle Y_i, y \rangle^2 + \mathbb{E} \langle Y_i, y \rangle \right)^{1/2}$$

$$\le c\, U_n \left(A + n \sigma_B^2 \right)^{1/2}$$

Therefore

$$A^2 \le c U_n^2 A + c\, U_n^2\, n \sigma_B^2$$

and

$$\left(A - c U_n^2 \right)^2 \le c \cdot n \sigma_B^2 U_n^2 + c U_n^4$$

$$\lesssim \max \left(\sqrt{n}\, \sigma_B U_n, U_n^2 \right)^2$$

so $A \lesssim \max \left(\sqrt{n}\, \sigma_B U_n, U_n^2 \right)$. □

Proof of the Main Thm.

Our goal is to prove $\operatorname{diam}(B_1^N \cap \operatorname{Ker} \Phi) \le \frac{1}{2\sqrt{m}}$.

It's enough to prove that

$$\mathbb{E} \sup_{y\in B_1^N \cap \rho S^{N-1}} \left| \sum_{i=1}^{n} \langle Y_i, y\rangle^2 - \frac{n\rho^2}{N} \right| \le \frac{2}{3} \frac{n\rho^2}{N} \quad ?$$

$$B = B_1^N \cap \rho S^{N-1}$$

so $\sigma_B = \sup_{y\in B} \left(\mathbb{E} \langle Y, y\rangle^2 \right)^{1/2} = \frac{\rho}{\sqrt{N}}$.

and $\gamma_2(B, d_{\infty,n}) \le \gamma_2(B_1^N, d_{\infty,n})$

We just use Dudley's estimate:

$$\gamma_2(B_1^N, d_{\infty,n}) \le \int_0^{+\infty} \sqrt{\log N(B_1^N, \varepsilon d_{\infty,n})}\, d\varepsilon.$$

Let $S: \mathbb{R}^n \to \mathbb{R}^N$ $\qquad S: \ell_1^n \to \ell_\infty^N$

$e_i \mapsto Y_i$ $\qquad S^*: \ell_1^N \to \ell_\infty^n$

$$\begin{aligned} d_{\infty,n}(z, \bar{z}) &= \max_{1\le i\le n} \langle Y_i, z-\bar{z}\rangle \\ &= \max_{1\le i\le n} \langle Se_i, z-\bar{z}\rangle \\ &= \max_{1\le i\le n} \langle e_i, S^*z - S^*\bar{z}\rangle \\ &= \| S^*(z-\bar{z})\|_\infty \end{aligned}$$

- $\sqrt{\log N(B_1^N, \varepsilon d_{\infty,n})} \leq \frac{c}{\sqrt{N}} \sqrt{\log n} \cdot \sqrt{\log N} \cdot \frac{1}{\varepsilon}$

  $\forall \varepsilon > 0$

∘ $\leq$

$$\sqrt{\log N(B_1^N, \varepsilon d_{\infty,n})} \leq \sqrt{n \log\left(1 + \frac{3}{\varepsilon\sqrt{N}}\right)}$$

$u = \varepsilon\sqrt{N}.$

$$\frac{1}{\sqrt{N}} \int_0^{+\infty} \sqrt{\log N\left(B_1^N, \frac{u}{\sqrt{N}} B_{\infty,n}\right)}\, du$$

~~$\frac{1}{\sqrt{N}} \int_0$~~

$$\sqrt{\log N\left(B_1^N, \frac{u}{\sqrt{N}} B_{\infty,n}\right)} \leq \left\{ \begin{array}{l} c \cdot \frac{\sqrt{\log n}\ \sqrt{\log N}}{u} \\ \\ \sqrt{n \log\left(1 + \frac{3}{u}\right)}. \end{array} \right.$$

$$u \sim \frac{\sqrt{\log N}}{\sqrt{n}}$$

$$\cdot \int_0^{\frac{c}{\sqrt{n}}} \sqrt{n \log\left(1+\frac{3}{u}\right)}\, du = \sqrt{n} \cdot \int_0^{\frac{c}{\sqrt{n}}} \sqrt{\log\left(1+\frac{3}{u}\right)}\, du$$

$$\sqrt{n}\, u = v.$$

$$= \int_0^c \sqrt{\log\left(1+\frac{3\sqrt{n}}{v}\right)}\, dv$$

$$\le \int_0^c c\sqrt{\log n + \underset{\substack{\uparrow \\ \text{integrable at} \\ \text{the origin}}}{\log\left(\frac{1}{v}\right)}}\, dv$$

$$\lesssim c\sqrt{\log n}$$

$$\cdot \int_{\frac{c}{\sqrt{n}}}^1 \frac{c\sqrt{\log n}\sqrt{\log N}}{u}\, du = c\sqrt{\log N}\sqrt{\log n}\,(\log n)$$

$$= c\sqrt{\log N}\,(\log n)^{3/2}.$$

So $U_n \le \dfrac{\sqrt{\log N}\,(\log n)^{3/2}}{\sqrt{N}}$

$$\mathbb{E}\sup_{y\in B_1^N\cap\rho S^{N-1}} \left|\sum_{i=1}^m \langle Y_i, y\rangle^2 - \frac{m\rho^2}{N}\right| \lesssim \max\left(\frac{\rho\sqrt{m}}{\sqrt{N}} \cdot \frac{\sqrt{\log N}\,(\log n)^{3/2}}{\sqrt{N}},\ \frac{\log N\cdot(\log n)^3}{N}\right).$$

Let's ~~take~~ ~~ε small enough~~ define ρ such that

such that $\dfrac{\sqrt{\log N}\,(\log n)^{3/2}}{\sqrt{N}} = \underset{\substack{\uparrow \\ \varepsilon \text{ smaller constant}}}{\frac{2}{3}}\rho\sqrt{\dfrac{m}{N}}$ then $\lesssim \dfrac{2}{3}\dfrac{m\rho^2}{N}$

$\square$

II) Harmonic analysis.

Let $\varphi_1, \dots, \varphi_N$ an o.n.b. of ℓ_2^N such that $\forall i, |\varphi_i|_\infty \leq \frac{K}{\sqrt{N}}$.

Let $Y = \varphi_i$ with proba $\frac{1}{N}$ and $Y_1, \dots, Y_n$ be ind. copies of Y

We have proved that

$$\mathbb{E} \sup_{y \in B_1^N \cap \rho S^{N-1}} \left| \sum_{i=1}^n \langle Y_i, y\rangle^2 - \frac{n\rho^2}{N} \right| \leq \frac{1}{3} \frac{n\rho^2}{N}.$$

when $\rho \simeq \frac{1}{\sqrt{N}} \cdot \sqrt{\frac{N}{n}} \sqrt{\log N} (\log n)^{3/2}$ $(\ast)$

Recall that $\Phi = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$

Properties of Φ: 1) $\operatorname{Ker} \Phi = \operatorname{span}\left\{\{\varphi_1, \dots, \varphi_N\} \setminus \{Y_i\}_{i=1}^n\right\} := \operatorname{span}(\varphi_i)_{i\in I}$

2) $(\operatorname{Ker} \Phi)^\perp = \operatorname{span}(\varphi_i)\ i \notin I$

3) If $(\ast)$ is satisfied ~~then~~ and $n < \frac{3N}{4}$ then $\operatorname{diam}(\operatorname{Ker}\Phi \cap B_1^N) \leq \rho$ and $\operatorname{diam}((\operatorname{Ker}\Phi)^\perp \cap B_1^N) \leq \rho$

proof: 1) and 2).

3) We have already seen that $\operatorname{diam}(\operatorname{Ker}\Phi \cap B_1^N) \leq \rho$.

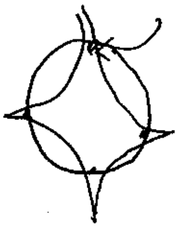
Moreover from the upper bound, $\forall y \in B_1^N \cap \rho S^{N-1}$

$$\sum_{i \in I} \langle \varphi_i, y\rangle^2 = \sum_{i=1}^N \langle \varphi_i, y\rangle^2 - \sum_{i \notin I} \langle \varphi_i, y\rangle^2$$

$$= |y|_2^2 - \sum_{i=1}^n \langle Y_i, y\rangle^2 \geq \rho^2 - \frac{4}{3} \frac{n\rho^2}{N}$$

$$= \rho^2\left(1 - \frac{4n}{3N}\right) > 0$$

$\psi = \begin{pmatrix} \varphi_i \\ \vdots \end{pmatrix}_{i \in I}$ hence



$\forall y \in B_1^N \cap \rho S^{N-1}$, $|\psi y|_2 > 0$

and $y \notin \operatorname{Ker}\psi$

As before, $\operatorname{diam}(\operatorname{Ker}\psi \cap B_1^N) \leq \rho$. But $\operatorname{Ker}\psi = (\operatorname{Ker}\Phi)^\perp$.

[illegible]

Conclusion :

If $n < \frac{3N}{4}$ then there exists a subset I of cardinality greater than $N-n$ such that :

$$\begin{cases} \left|\sum_{i\in I} a_i \varphi_i\right|_2 \le \frac{1}{\sqrt{N}} \sqrt{\frac{N}{n}} \sqrt{\log N} (\log n)^{3/2} \left|\sum_{i\in I} a_i\varphi_i\right|_1 \\ \text{and} \quad \left|\sum_{i\notin I} a_i \varphi_i\right|_2 \le \frac{1}{\sqrt{N}} \cdot \sqrt{\frac{N}{n}} \sqrt{\log N} (\log n)^{3/2} \left|\sum_{i\in I} a_i\varphi_i\right|_1 \end{cases}$$

In fact $\mathbb{P}(\exists I$ defined by ... satisfying the conclusion$) \ge 1/2$.

And if $n = \lfloor \log 2\, N \rfloor$ $\cancel{\le \frac{3}{4} cN}$ then $n < \frac{3}{4} N$

and $\mathbb{P}\left(\frac{N}{2} - c\sqrt{N} \le \# I \le \frac{N}{2} + c\sqrt{N} \right) \ge 1/2$

We have proved :

Theorem : $\exists$ a subset I such that $\frac{N}{2} - c\sqrt{N} \le \# I \le \frac{N}{2} + c\sqrt{N}$

G-M-P -Tomczak Jaegermann '06

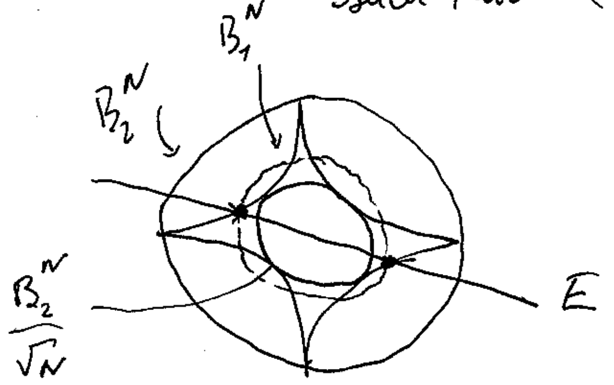
and $\left|\sum_{i\in I} a_i \varphi_i\right|_2 \le \frac{1}{\sqrt{N}} \cdot \cancel{\sqrt{\log N}} (\log N)^{\cancel{3/2}\, 2} \left|\sum_{i\in I} a_i\varphi_i\right|_1$

and $\left|\sum_{i\notin I} a_i \varphi_i\right|_2 \le \frac{1}{\sqrt{N}} \cdot (\log N)^2 \left|\sum_{i\notin I} a_i\varphi_i\right|_1$

Historical comments.

Milman '71 : $\forall \varepsilon \in (0,1)$, $\exists E \subset \mathbb{R}^N$, $\dim E = n \sim c\frac{\varepsilon^2}{\log(1+\frac{2}{\varepsilon})} N$

such that $(1-\varepsilon)\, r \frac{B_2^N}{\sqrt{N}} \subset B_1^N \cap E \subset (1+\varepsilon)\, r \frac{B_2^N}{\sqrt{N}}$

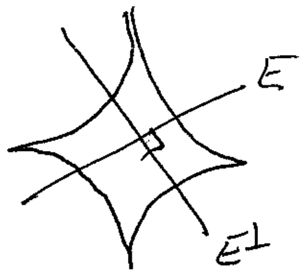


Kashin '77 : If $N = 2n$ then $\exists E \subset \mathbb{R}^N$ of dim n such that

$$\forall x \in E, \quad \frac{|x|_1}{\sqrt{N}} \le |x|_2 \le C \cdot \frac{|x|_1}{\sqrt{N}}$$

and $\forall x \in E^\perp$, $$\frac{|x|_1}{\sqrt{N}} \le |x|_2 \le c \, \frac{|x|_1}{\sqrt{N}} .$$

where C is a universal constant



Szarek
Szarek - Tomczak Jaegermann
"volume ratios".

↳ Algorithmic construction of such subspace ? Indyk

- In our problem, the basis $(\varphi_1, \dots, \varphi_N)$ is given and we want to find a coordinate subspace that satisfies good properties.

Remark: we allways have

$$\Big|\sum_{i\in I} a_i \varphi_i\Big|_1 \geqslant \max_{\cancel{k} i \in I} |a_i| \frac{\sqrt{N}}{K} \geqslant \frac{1}{\sqrt{\# I}} \left(\sum |a_i|^2\right)^{1/2} \frac{\sqrt{N}}{K}$$

hence $$\frac{K\sqrt{\# I}}{\sqrt{N}} \Big|\sum_{i\in I} a_i \varphi_i\Big|_1 \geqslant \Big|\sum_{i\in I} a_i \varphi_i\Big|_2$$

Talagrand '98
Bourgain

//: $\exists \, \delta_0 > 0$ small constant, $\exists I$, $\# I \geqslant \delta_0 N$ such that $$\Big|\sum_{i\in I} a_i \varphi_i\Big|_2 \lesssim \frac{1}{\sqrt{N}} \sqrt{\log N (\log \log N)} \Big|\sum_{i\in I} a_i \varphi_i\Big|_1$$

→ Dvoretzky type Thm
→ Majorizing measure.

- It was known from Bourgain that $\sqrt{\log N}$ is necessary in the estimate.

Theorem. GMPT '08.

1) There exists a subset I with $\# I \geq N-n$ such that

$$\Big|\sum_{i\in I} a_i \varphi_i\Big|_2 \leq \frac{1}{\sqrt{N}} \quad \mu \left(\log \mu\right)^{5/2} \Big|\sum_{i\in I} a_i \varphi_i\Big|_1$$

where $\mu = K \sqrt{\frac{N}{n} \log n}$

2) There exists a subset I with $\frac{N}{2} - c\sqrt{N} \leq \# I \leq \frac{N}{2} + c\sqrt{N}$ s.t.

Kashin's splitting

$$\left\{ \begin{array}{l} \Big|\sum_{i\in I} a_i \varphi_i\Big|_2 \lesssim \frac{1}{\sqrt{N}} \cdot \sqrt{\log N} \left(\log\log N\right)^{5/2} \Big|\sum_{i\in I} a_i \varphi_i\Big|_1 \\ \text{and} \\ \Big|\sum_{i\notin I} a_i \varphi_i\Big|_2 \lesssim \frac{1}{\sqrt{N}} \sqrt{\log N} \left(\log\log N\right)^{5/2} \Big|\sum_{i\notin I} a_i \varphi_i\Big|_1 \end{array} \right.$$

Improvement on the study of the empirical processes via the majorizing measure theory.

→ a Banach space X is called of type 2

if $\exists C>0$, $\forall n$, $\forall x_1, .., x_n \in X$,

$$\left(\mathbb{E} \Big\| \sum_{i=1}^{n} \varepsilon_i x_i \Big\|^2\right)^{1/2} \leq C \cdot \left(\sum_{i=1}^{n} \|x_i\|^2\right)^{1/2}$$

↳ Hilbert spaces have type 2 (parallelogram identity!)

↳ L_q spaces with $q \geq 2$ have type 2

→ a Banach space X has modulus of convexity of power type 2 (with constant λ) if

$\forall x, y \in X$, $$\Big\|\frac{x+y}{2}\Big\|^2 + \lambda^{-2} \Big\|\frac{x-y}{2}\Big\|^2 \leq \frac{1}{2}\left(\|x\|^2 + \|y\|^2\right)$$

Picture : $$\delta_X(\varepsilon) = \inf \left\{ 1 - \Big\|\frac{x+y}{2}\Big\| ,\ \|x\| = \|y\| = 1 \text{ and } \|x-y\| < \varepsilon \right\}$$

$$\delta_X(\varepsilon) \geq \varepsilon^2 / 8\lambda^2.$$

→ Moreover, if X has modulus of convexity of power type 2 then X^* has modulus of smoothness of power type 2

i.e. $$\left\|\frac{x+y}{2}\right\|_*^2 + \lambda^{2}\left\|\frac{x-y}{2}\right\|^2 \geq \frac{1}{2}\left(\|x\|_*^2 + \|y\|_*^2\right)$$

<u>Rk</u>: this implies that X^* has type 2. Indeed

$$\mathbb{E}\,\|\varepsilon u + v\|_*^2 = \frac{1}{2}\left(\|u+v\|_*^2 + \|-u+v\|_*^2\right)$$

$$\leq \|v\|_*^2 + \lambda^{2}\|u\|_*^2 .$$

By induction,
$$\mathbb{E}\,\Big\|\sum_{i=1}^n \varepsilon_i v_i\Big\|_*^2 \leq \lambda^2\left(\sum_{i=1}^n \|v_i\|_*^2\right)$$

and $T_2(X^*) \leq \lambda$. □.

<u>Key theorem.</u> (GMPT).

If E is a Banach space with modulus of convexity of power type 2 with constant λ then $\forall\, Y_1, \dots, Y_n \in E^*$,

$$\mathbb{E}\sup_{y \in B_E}\left|\sum_{i=1}^n g_i \langle Y_i, y\rangle^2\right|$$

$$\lesssim \lambda^5 \sqrt{\log n}\ \max_{1\leq i\leq n} |Y_i|_* \ \sup_{y\in B_E}\left(\sum_{i=1}^n \langle Y_i, y\rangle^2\right)^{1/2} .$$

As before we deduce from this result that

$$\mathbb{E} \sup_{y\in B_E} \left| \sum_{i=1}^{n} \langle Y_i, y\rangle^2 - \mathbb{E}\langle Y_i, y\rangle^2 \right| \lesssim \max\left(\sqrt{n}\, \sigma_{B_E}\, U_n,\ U_n^2\right)$$

where $U_n = \lambda^5 \sqrt{\log n} \left(\mathbb{E} \max |Y_i|_*^2\right)^{1/2}$.

$$\sigma_{B_E} = \sup_{y\in B_E} \left(\mathbb{E}\langle Y, y\rangle^2\right)^{1/2}$$

$$\to \quad B_E \approx B_p^N \cap \rho B_2^N .$$

$$|Y_i|_* \lesssim |Y_i|_q \leq N^{1/q} |Y_i|_\infty \leq \frac{K N^{1/q}}{\sqrt{N}}$$

For any $p > 1$, B_p^N has modulus of convexity of power type 2 with $\lambda^2 \approx \frac{1}{p-1}$ i.e. $\lambda \approx \frac{1}{(p-1)^{1/2}}$.

So $$\mathbb{E} \sup_{y\in B_p^N \cap \rho S^{N-1}} \left| \sum_{i=1}^{n} \langle Y_i, y\rangle^2 - \frac{n\rho^2}{N} \right|$$

$$\lesssim \max\left(\underbrace{\sqrt{n}\, \frac{\rho}{\sqrt{N}}}_{''\ n\rho^2/N}\, N^{1/q} \frac{\sqrt{\log n}}{\sqrt{N}} \lambda^5,\ \underset{=}{\left(\frac{N^{1/q}\lambda^5 \sqrt{\log n}}{\sqrt{N}}\right)^2} \right)$$

and you choose $\rho \approx \sqrt{\frac{N}{n}} \cdot N^{1/q} \frac{\sqrt{\log n}}{\sqrt{N}} \lambda^5$

Hence there exists vectors $Y_1, \dots, Y_n$ s.t.

$$\operatorname{diam}\left(\operatorname{Ker} \Phi \cap B_p^N\right) \leq \rho$$

i.e. $$\left|\sum_{i\in I} a_i \varphi_i\right|_2 \leq \sqrt{\frac{N}{n}} \sqrt{\log n}\, \frac{N^{1/q}\lambda^5}{\sqrt{N}} \left|\sum_{i\in I} a_i\varphi_i\right|_p$$

But $|t|_p \leq |t|_1^{\theta} \, |t|_2^{1-\theta}$ where $\theta = \frac{2-p}{p}$ (Hölder)

So $\left| \sum_{i \in I} a_i \varphi_i \right|_2 \leq \mu \, \frac{N^{1/q}}{\sqrt{N}} \lambda^5 \, \left| \sum a_i \varphi_i \right|_1^{\theta} \cdot \left| \sum a_i \varphi_i \right|_2^{1-\theta}$

and $\left| \sum_{i \in I} a_i \varphi_i \right|_2 \leq \frac{1}{\sqrt{N}} \mu^{1/\theta} \lambda^{5/\theta} \cdot \left| \sum_{i \in I} a_i \varphi_i \right|_1$

So you choose $p = 1 + \frac{1}{\log \mu}$

So that $\theta = 1 - \frac{1}{\log \mu}$ and $\lambda = \sqrt{\log \mu}$

and you conclude

$$\left| \sum_{i \in I} a_i \varphi_i \right|_2 \leq \frac{1}{\sqrt{N}} \mu \cdot (\log \mu)^{5/2} \left| \sum_{i \in I} a_i \varphi_i \right|_1 .$$