Embeddings under various notions of randomness

Nicole Tomczak-Jaegermann

June 25, 2007

– 3rd Annual PHD Conference, Samos

Plan of the talk

- Subgaussian embeddings/projections: Shahar Mendelson and NT-J
- Embeddings of convex bodies in ℓ_∞^N and ε nets: Yoram Gordon, Sasha Litvak, Alain Pajor and NT-J
- Very economical "almost isometric" embeddings of ℓ_2^n into ℓ_∞^N

Notation

$\mathbb{R}^n$, $\mathbb{R}^k$, $(e_i)_i$ the unit vector basis, $|\cdot|$ the Euclidean norm, S^{n-1} – sphere

(g_i), (g_{ij}) be independent $N(0,1)$ r.v.

Fix $n, k \geq 1$; define

$$\Gamma : \mathbb{R}^n \to \mathbb{R}^k$$

by

$$\Gamma t = \sum_{i=1}^{k} \sum_{j=1}^{n} g_{ij} t_j e_i$$

for $t = (t_j)$.

In other words, $\Gamma = \left[g_{ij}\right]$ is $k \times n$ matrix
Γe_j is the jth column of Γ, for $j = 1, \ldots, n$.

Gaussian theorem

Theorem 1. *There is $c > 0$ s.t.: Let $T \subset S^{n-1}$ and $E = (\mathbb{R}^k, \| \ \|_E)$ satisfy $\|x\|_E \leq |x|$. Fix $\varepsilon > 0$ and assume that*

$$\mathbb{E} \sup_{t \in T} | \sum_{i=1}^{n} g_i t_i | \leq c\, \varepsilon\, \mathbb{E} \| \sum_{i=1}^{k} g_i e_i \|_E .$$

With probability close to 1, Γ satisfies, for every $t \in T$,

$$(1 - \varepsilon) \mathbb{E} \| \sum_{i=1}^{k} g_i e_i \|_E \leq \|\Gamma t\|_E \leq (1 + \varepsilon) \mathbb{E} \| \sum_{i=1}^{k} g_i e_i \|_E .$$

Special cases

$\Gamma : \mathbb{R}^n \to \mathbb{R}^k$. If

$$\ell_*(T) := \mathbb{E} \sup_{t \in T} |\sum_{i=1}^{n} g_i t_i| \ \le \ c\,\varepsilon\,\mathbb{E}\|\sum_{i=1}^{k} g_i e_i\|_E \ =: c\,\varepsilon\,\ell(E)$$

then

$$(1-\varepsilon)\ell(E) \le \|\Gamma t\|_E \le (1+\varepsilon)\ell(E) \quad \text{for } t \in T \subset S^{n-1}.$$

• $n \le k$, $T = S^{n-1}$: Gaussian Dvoretzky-type embedding of ℓ_2^n into E Condition $\sqrt{n} \le c\,\varepsilon \ell(E)$ appears in the familiar formulation going back to Milman, around 1970

• $n \ge k$, Gaussian projection; if $E = (\mathbb{R}^k, |\cdot|) = \ell_2^k$ and $T \subset S^{n-1}$ satisfies $\mathbb{E}\sup_{t\in T} |\sum_{i=1}^{n} g_i t_i| \le c\,\varepsilon\sqrt{k}$, then Γ is almost an isometry on T;

Johnson-Lindenstrauss Lemma: if $\log|T| \le c'\varepsilon^2 k$ then Γ provides an “almost distance-preserv.” mapping of T onto a subset of ℓ_2^k.
Generalized to *any* normed space E.

Gaussian theorem, final comments

Theorem 1 easily follows (by a known argument) from the Gaussian min-max theorem by Y. Gordon.

The present formulation was proposed in a recent paper by G. Schechtman (who gave a proof by majorizing measures approach).

A yet another proof by G. Pisier, based on his Gaussian measure concentration theorem.

All three approaches use in an *essential* way that Γ is a Gaussian operator.

We use the concentration of a random vector $\|\Gamma t\|_E$ around its mean to prove an *isomorphic* analog of Theorem 1, where the Gaussian operator is replaced by an arbitrary *subgaussian* operator, under a cotype assumption on the space E.

Subgaussian theorem; notation

The ψ_2 norm of a random variable ξ:
$\|\xi\|_{\psi_2} = \inf\left\{u > 0 : \mathbb{E}\exp\left(|\xi|^2/u^2\right) \le 2\right\}.$

μ a symmetric measure on $\mathbb{R}^n$, isotropic and L-subgaussian if:
letting $X \in \mathbb{R}^n$ be a random vector distributed according to μ

$$\mathbb{E}\langle X, t\rangle^2 = |t|^2 \text{ and } \|\langle X, t\rangle\|_{\psi_2} \le L|t|, \qquad \text{for } t \in \mathbb{R}^n.$$

(a subgaussian decay of linear functionals, $\exp(-cu^2/L^2)$).

An operator $\Gamma : \mathbb{R}^n \to \mathbb{R}^k$ is L-subgaussian if ($k \times n$ matrix)

$$\Gamma t = \sum_{i=1}^{k} \langle X_i, t\rangle e_i \qquad \text{for} \quad t \in \mathbb{R}^n,$$

where $(X_i)_{i=1}^k$ are independent random vectors as above.

Subgaussian theorem; notation II

A Banach space E has cotype $q \geq 2$ with a constant β_q if for all finite sequences (z_i) in E,

$$\Big(\sum_i \|z_i\|_E^q\Big)^{1/q} \leq \beta_q \, \mathbb{E}\Big\|\sum_i \varepsilon_i z_i\Big\|_E,$$

where $(\varepsilon_i)_i$ are independent Bernoulli random variables.

Corresponds to the "lower estimate" for $\mathbb{E}$ in the parallelogram identity.

Examples: classical and non-commutative L_p-spaces, Schatten classes S_p, for $1 \leq p < \infty$.

Subgaussian theorem

Theorem 2. $\exists c_1, c_2, c_3 > 0$ *s.t.: Let* $T \subset S^{n-1}$ *and* $E = (\mathbb{R}^k, \|\cdot\|_E)$ *with* $\|x\|_E \le |x|$, *for* $x \in \mathbb{R}^k$. *Fix* $L > 0$, *assume that* E *has cotype* q *with a constant* β_q *and that* *(recall* $\ell(E) := \mathbb{E}\|\sum_{i=1}^k g_i e_i\|_E$*)*

$$\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n g_i t_i\Big| =: \ell_*(T) \le \frac{(c_3/L^2\beta_q\sqrt{q})}{\sqrt{\log k}}\,\ell(E).$$

If Γ *is* L*-subgaussian then, with probability close to 1, for every* $t \in T$

$$(c_1/L\beta_q\sqrt{q})\,\ell(E) \le \|\Gamma t\|_E \le c_2 L\,\ell(E).$$

Moreover, if Γ *is an operator with independent Bernoulli random entries then the logarithmic factor in the hypothesis can be removed.*

Corollary for Rademacher embedding

Let $E = (\mathbb{R}^k, \|\cdot\|_E)$ with $\|x\|_E \leq |x|$, for $x \in \mathbb{R}^k$ and assume that E has cotype q with a constant β_q. Let

$$n \leq \ (c_3/\beta_q^2 q)\ (\ell(E))^2.$$

If Γ is an operator with independent Bernoulli random entries then, with probability close to 1, the random subspace $F := \Gamma(\mathbb{R}^n) \subset \mathbb{R}^k$ spanned by ± 1 vectors satisfies $\|z\|_E \sim A|z|$, for all $z \in F$ (for a certain $A > 0$).

In particular, if the Euclidean unit ball on $\mathbb{R}^k$ is the maximal volume ellipsoid for E then the assertion holds once $n \leq (c_3/\beta_q^2 q) k^{2/q}$.

Figiel-Lindenstrauss-Milman: Thus we fully recover the dimension of Euclidean subspaces in spaces with cotype q obtained in FLM.
Subspaces here have more structure (but are isomorphic).

Role of the cotype assumption I

For a Gaussian operator, with large probability, for all $t \in T$,

$$(1-\varepsilon)\ell(E) \leq \|\Gamma t\|_E \leq (1+\varepsilon)\ell(E).$$

So,

$$\mathbb{E}\|\Gamma t\|_E = \mathbb{E}\|\sum_i \sum_j g_{ij} t_j e_i\|_E = \mathbb{E}\|\sum_i g_i e_i\|_E = \ell(E)$$

does not depend on $t \in T$.

For a *subgaussian* operator, $\mathbb{E}\|\Gamma t\|_E$ *does* depend on $t \in T$. Not good!

Role of the cotype assumption II

In spaces of cotype $q < \infty$, Rademacher and Gaussian averages are equivalent.

Same is true for averages with respect to arbitrary independent symmetric L-subgaussian r.v. In particular,

$$c\,\ell(E) = \mathbb{E}\| \sum_{i=1}^{k} g_i e_i \|_E \le \inf_{t\in T} \mathbb{E}\|\Gamma t\|_E \le \sup_{t\in T} \mathbb{E}\|\Gamma t\|_E \le C\,\ell(E),$$

where c, C depend on q, β_q and L (but not on E!)

Concentration for individual vectors

Concentration result for each r.v., $\| \sum_{i=1}^{k} \langle X_i, t\rangle \, e_i\|_E$ around its mean.

Let $(\xi_i)_{i=1}^k$ be i.i.d. symmetric L-subgaussian r.v. Let $E = (\mathbb{R}^k, \| \cdot \|_E)$, such that for every $x \in E$, $\|x\|_E \le |x|$. Then for every $u \ge 2$,

$$\mathbb{P}\left(\left| \| \sum_{i=1}^{k} \xi_i e_i \|_E - \mathbb{E} \| \sum_{i=1}^{k} \xi_i e_i \|_E \right| \ge u \right) \le 2 \exp(-c\, u^2 / L^2 \log k).$$

For Gaussian or bounded r.v., the estimate is valid without $\log k$ factor.

We use for $\xi_i = \langle X_i, t\rangle$ for $i = 1, \ldots, k$, for a fixed $t \in T$.

Approximation

Let $\Gamma : \mathbb{R}^n \to \mathbb{R}^k$ be L-subgaussian.

Let $T \subset S^{n-1}$ and $E = (\mathbb{R}^k, \| \cdot \|_E)$, such that $\|x\|_E \leq |x|$, and, for each $t \in T$, the r.v. $\xi_i = \langle X_i, t \rangle$ satisfy the concentration inequality.

Standard approximation: For $\varepsilon > 0$, let Λ be an ε net for T with respect to $| \cdot |$, so that

$$T \subset \bigcup_{y \in \Lambda} y + \varepsilon B_2^n.$$

We require the Lipschitz constant of $\|\Gamma(\cdot)\|_E : S^{n-1} \to \mathbb{R}$

Instead, let $U := T \cap \varepsilon B_2^n$. Then, $T \subset \Lambda + U$ and so sufficient to require an upper bound for $\sup_{u \in U} \|\Gamma u\|_E$ with large probability.

Norms of random operators

For any $s \geq 1$,

$$\mathbb{P}\left(\sup_{u \in U} \|\Gamma u\|_E \leq cs\,H\right) \geq 1 - 2\exp(-s^2),$$

where $H := L\Big(\ell_*(T) + \varepsilon\ell(E)\Big)$ and $c > 0$ is an absolute constant.

Norm of a random operator:

$$\sup_{u \in U} \|\Gamma u\| = \|\Gamma : (\mathbb{R}^n, \mathrm{conv}\ U) \to E\|.$$

The Gaussian case contained in Chevet-Gordon inequality.

Subgaussian: by majorizing measure of Talagrand, for example by generic chaining.

Almost isometric embeddings into ℓ_∞^N

ℓ_∞ the space of all bounded sequences endowed with

$$\|t\|_\infty = \sup_i |t_i|, \quad \text{for } t = (t_i) \in \ell_\infty.$$

Every separable Banach space is isometric to a subspace of ℓ_∞.

$\ell_\infty^N = (\mathbb{R}^N, \|\cdot\|_\infty)$: the unit ball is the cube

Fact. *For every n and $\varepsilon > 0$ there is $N = N(n, \varepsilon)$ s.t. every n-dimensional normed space E is $1+\varepsilon$ isomorphic to a subspace of ℓ_∞^N.*

Let $E = (\mathbb{R}^n, \|\cdot\|)$, K the unit ball (a symmetric convex body in $\mathbb{R}^n$).
sometimes we denote $\|\cdot\| = \|\cdot\|_K$.

$K^\circ = \{z \in \mathbb{R}^n : |\langle x, z\rangle| \le 1, \text{ for all } x \in K\}$, the polar body.
Banach space duality: $(\mathbb{R}^n, K^\circ)$ can be identified to the dual space E^*.

Almost isometric embeddings into ℓ_∞^N II

Let $\Lambda \subset K^\circ$ be an ε net in K°, that is, $K^\circ \subset \bigcup_{y\in\Lambda} y + \varepsilon K^\circ$.

Set $N = |\Lambda|$, say $\Lambda = (y_i)_{i=1}^N$.

Set $\Gamma : \mathbb{R}^n \to \mathbb{R}^N$ by $\Gamma t = \big(\langle y_i, t\rangle\big)_{i\le N} \in \mathbb{R}^N$. Rows are y_i.

Then $\|\Gamma t\|_\infty = \max_i |\langle y_i, t\rangle| \le \|t\|$. Converse direction also easy, Together:

$$(1-\varepsilon)\|t\| \le \|\Gamma t\|_\infty \le \|t\| \qquad \text{for all} \quad t \in \mathbb{R}^n.$$

So the subspace $\Gamma(\mathbb{R}^n) \subset \ell_\infty^N$ is $(1-\varepsilon)^{-1}$ isomorphic to $E = (\mathbb{R}^n, \|\cdot\|)$, for $N = |\Lambda|$. Thus $N(n,\varepsilon) = |\Lambda|$.

ε nets

Let $K \subset \mathbb{R}^n$ symmetric convex body. For every ε net $\Lambda \subset K$, $\big(1/\varepsilon\big)^n \leq |\Lambda|$ oppositely, there is always an ε net $\Lambda' \subset K$ with $|\Lambda'| \leq \big(3/\varepsilon\big)^n$ (we take a maximal ε-separated subset of K).

For a set of vectors in $\mathbb{R}^n$, being an ε net is not such a rare occurrence.

Randomness determined by K: a vector X is uniformly distributed on K if

$$\mathbb{P}\left(\{X \in A\}\right) = \frac{|K \cap A|}{|K|}$$

for every measurable $A \subset \mathbb{R}^n$.

Theorem 3. *Let $n \geq 1$, $0 < \varepsilon \leq 1$, and $N = (4/\varepsilon)^{2n}$. Let $K \subset \mathbb{R}^n$ symmetric convex body, and $X_1, \ldots, X_N$ be independent random vectors uniformly distributed on K. Then with a probability close to 1, the set $\Lambda = \{X_1, \ldots, X_N\}$ forms an ε-net in K.*

Random ε-nets, sketch of proof

Lemma. *Let K be a symmetric convex body in $\mathbb{R}^n$. For every $x \in K$ and for $0 < \varepsilon \leq 1$ one has*

$$|K \cap (x + \varepsilon K)| \geq \left|\frac{\varepsilon}{2}\, K\right|.$$

Set $\alpha = 1 - \frac{\varepsilon}{2}$, $\beta = \frac{\varepsilon}{2}$. Fix $x \in K$. Enough to show that

$$K \cap (x + \varepsilon K) \supset \alpha x + \beta K.$$

Let $z = \alpha x + \beta y$, where $y \in K$. Clearly, $z \in K$. Write $z = x + \beta(y - x)$. Then $y - x \in 2K$, so $z \in x + \beta 2K = x + \varepsilon K$, as required.

Random embedding into ℓ_∞^N

Combining Theorem 3 and Fact we get

Theorem 4. *Let $0 < \varepsilon < 1$ and $n \leq \log N / 2\ln(4/\varepsilon)$. Let K be a symmetric convex body in $\mathbb{R}^n$. Let $X_1, \ldots, X_N$ be independent random vectors uniformly distributed on K°. Consider $\Gamma : \mathbb{R}^n \to \mathbb{R}^N$ whose rows are $X_1, \ldots, X_N$*
Then with probability close to 1 we have, for $x \in \mathbb{R}^n$,

$$(1 - \varepsilon)\ \|x\|_K \leq \|\Gamma x\|_\infty \leq \|x\|_K.$$

For sections of the cube $(1+\varepsilon)$-isomorphic to K:
there exists a section if $N \geq \left(\frac{3}{\varepsilon}\right)^n$, equivalently, $n \leq \dfrac{\log N}{\log(\frac{3}{\varepsilon})}$;
a *random* section if $n \leq \dfrac{\log N}{2\log(\frac{4}{\varepsilon})}$.

Random Euclidean sections of the cube

X_1, . . . , X_N independent random vectors uniformly distributed on S^{n-1}.

Let $\Lambda = \{X_1, \ldots, X_N\}$. $\Gamma : \mathbb{R}^n \to \mathbb{R}^N$, $\Gamma x = \sum_{i=1}^N \langle x, X_i \rangle e_i$.

For any $y \in S^{n-1}$ the normalized Lebesgue measure of a cap
$$\left\{x \in S^{n-1} \;\mid\; |x - y| \leq \varepsilon\right\}$$
is larger than or equal to $(\varepsilon/2)^n$.

Theorem 5. *Let* $0 < \varepsilon < 1$ *and* $n \leq \log N / 2 \log(4/\varepsilon)$.
With probability close to 1, Λ *forms an* ε*-net on* S^{n-1}*, and the matrix* Γ *satisfies, for all* $x \in \mathbb{R}^n$,

$$(1 - \varepsilon)\; |x| \leq \|\Gamma x\|_\infty \leq |x|.$$

Random spherical sections of the cube

Q the unit ball of ℓ_∞^N (i.e. the N-dimensional cube).
By Theorem 5, a random section (by $E := \Gamma\mathbb{R}^n$) is almost an ellipsoid.

$$\Gamma B_2^n \subset Q \cap E \subset (1-\varepsilon)^{-1}\, \Gamma B_2^n.$$

On ℓ_∞^N there is the natural Euclidean norm, we want to compare $Q \cap E$ to a standard Euclidean ball of a certain radius.

We show that ΓB_2^n is, up to $\frac{1+\varepsilon}{1-\varepsilon}$, equivalent to the standard Euclidean ball of radius $\sqrt{N/n}$.
Thus a random section is almost spherical $\qquad E := \Gamma\mathbb{R}^n$

$$(1-\varepsilon)\sqrt{N/n}\, B_2^n \cap E \subset Q \cap E \subset \frac{1+\varepsilon}{1-\varepsilon}\sqrt{N/n}\, B_2^n \cap E.$$

Random spherical sections of the cube II

Theorem 6. *Under the assumptions of Theorem 5, with probability close to 1, for all $z \in E = \Gamma\mathbb{R}^n$ we have*

$$\frac{1-\varepsilon}{1+\varepsilon}\ |z| \le \sqrt{\frac{N}{n}}\,\|z\|_\infty \le \frac{1}{1-\varepsilon}\ |z|.$$

Lemma. *Let $0 < \varepsilon < 1$ and let $N \ge n^3/\varepsilon^4$. With probability close to 1,*

$$(1-\varepsilon)|x| \le |\Gamma x|\sqrt{n/N} \le (1+\varepsilon)|x|,$$

for every $x \in \mathbb{R}$.

We show: for any $\delta > 0$, $\quad \mathbb{P}\{\|\frac{n}{N}\Gamma^*\Gamma - I\| < \delta\} \ge 1 - \delta'$. Thus $\sqrt{\frac{n}{N}}\Gamma$ is almost an isometry of the Euclidean norm.

Dimension of random Euclidean sections I

Let $0 < \varepsilon < 1$. Whenever

$$n \leq \log N/2\log(4/\varepsilon),$$

then we produced a *random* embedding of ℓ_2^n in ℓ_∞^N.
It is given by $\Gamma : \mathbb{R}^n \to \mathbb{R}^N$.

Known: if $E \subset \ell_\infty^N$ is $(1+\varepsilon)$-Euclidean then

$$\dim E \leq C\log N/\log(1/\varepsilon).$$

So the dimension n of a *random* embedding Γ – as a function of ε – is (asymptotically) as large as a dimension of a Euclidean embedding can be, for *any* embedding.

Dimension of random Euclidean sections II

If G a matrix with independent $N(0,1)$ Gaussian entries, then n satisfying

$$n \leq \varepsilon \log N$$

is sufficient to get $(1+\varepsilon)$-embedding with large probability (Schechtman).

This is optimal if one requires large probability in the Haar – rotational invariant – measure on the Grassman manifold)

Randomness given by Γ allows as large dimension of sections as can be (even for deterministic embeddings). Rotational invariant randomness much more restrictive.