“On the metric entropy”

Based on joint works with

V. Milman, A. Pajor, and N. Tomczak-Jaegermann

papers available at my webpage:
http://www.math.ualberta.ca/~alexandr/papers

Let K, L be subsets of $\mathbb{R}^n$. The covering number $N(K, L)$ of K by L is the minimal number N such that there are vectors x_1, ..., x_N in $\mathbb{R}^n$ satisfying

$$K \subset \bigcup_{i=1}^{N} (x_i + L).$$

We use notation $\bar{N}(K, L)$ if additionally $x_i \in K$.

General idea: to extend theorems from operator theory or volume inequalities to the covering number setting.

Note,

$$\|T \ : \ (\mathbb{R}^n, K) \to \ (\mathbb{R}^n, L)\| \leq a$$

means

$$TK \subset aL,$$

equivalently

$$N(TK, aL) \leq 1.$$

Usually, in operator theory we have condition

norm of an operator is bounded by, say, one,

in other words we control the diameter of a body. Is it possible to say something similar when we control the covering number?

Examples. 1. Duality conjecture.

$\|T\| = \|T^*\|, \quad i.e. \quad \|T\| \leq 1$ implies $\|T^*\| \leq 1.$

Corresponding result (conjecture) for covering numbers would be: *there exists absolute positive constants a, b such that*

$$N(K, L) \leq N^b(L^0, aK^0).$$

2. Extension Property of ℓ_∞.

If $\|T : K \cap E \to \ell_\infty\| \leq 1$ then there is an extension:

$$\|\bar{T} : K \to \ell_\infty\| \leq 1 \quad \text{and} \quad \bar{T}_{|E} = T.$$

A week version of entropy extension (LPT): *Let* $0 < a < r < A$ *and* $1 \leq k < n$. *Let* $K, L \subset \mathbb{R}^n$ *be symmetric convex bodies, and* $K \subset AL$. *Let* $codim E = k$ *and* $K \cap E \subset aL$. *Then*

$$N(K, 2rL) \leq \left(\frac{3A}{r-a}\right)^k.$$

(If we control the diameter of a body in a subspace then we control the entropy in the entire space.)

Remark. The above result was used in **LPT** to investigate the phenomena "deterministic implies random" (in context of Gelfand numbers) and later, in **LMPT**, to investigate sharpness in Sudakov inequality.

Question. Can we provide a similar statement with the control of the entropy in the subspace instead of diameter?

A strong version (LMPT):

Let $0 < a < r < A$. Let K, L be symmetric convex bodies in $\mathbb{R}^n$ such that $K \subset AL$. Let E be k-codimensional subspace of $\mathbb{R}^n$. Then

$$N(K, rL) \leq \left(\frac{3A}{r-a}\right)^k \; N(K \cap E, \frac{a}{3}L).$$

"Dual" version (entropy lifting):

Let $0 < a < r < A$. Let K be a symmetric convex body in $\mathbb{R}^n$. Let $P : \mathbb{R}^n \to \mathbb{R}^n$ be a projection of corank k.

$$N(K, rL) \leq \left(\frac{6A}{r-a}\right)^k \; N(PK, \frac{a}{2}PL).$$

3. Rogers-Shephard inequality. Let K, L be a convex body in $\mathbb{R}^n$ and E be a k-dimensional subspace of $\mathbb{R}^n$. Then

$$|K| \le |P_E K| \max_{x \in K} |(K - x) \cap E^{\perp}| \le \binom{n}{k} |K|.$$

Theorem 1 (entropy decomposition). *Let K, L_1, and L_2 be subsets of $\mathbb{R}^n$. Let E be a subspace of $\mathbb{R}^n$ and $P : \mathbb{R}^n \to \mathbb{R}^n$ be a projection with* $\ker P = E$. *Then*

$$N(K, L_1 + L_2)$$

$$\le \bar{N}(PK, PL_1) \max_{z \in K} N((K - L_1 - z) \cap E, L_2)$$

$$\le \bar{N}(PK, PL_1)\, N((K - K - L_1) \cap E, L_2).$$

Remark. In fact this theorem implies "entropy extension" and "entropy lifting".

Proofs

Extension and lifting properties of entropy.

Having $\quad N\left(K, L_1 + L_2\right)$

$\leq \bar{N}\left(PK, PL_1\right) \; N\left(\left(K - K - L_1\right) \cap E, L_2\right), \quad (*)$

want to prove

$$N(K, rL) \leq \left(\frac{3A}{r-a}\right)^k \; N(K \cap E, \frac{a}{3}L).$$

for $0 < a < r < A$ and $K \subset AL$.

Let $\varepsilon := r - a$. First, by the convexity of L,

$$N(K, rL) \leq N(K, \frac{\varepsilon}{A}K + aL).$$

Using $(*)$ with $L_1 = (\varepsilon/A)K$ and $L_2 = aL$,

$$N(K, rL) \leq \bar{N}\left(PK, \frac{\varepsilon}{A}PK\right) N\left(\left(2 + \frac{\varepsilon}{A}\right) K \cap E, aL\right).$$

Now, the first factor is bounded by $(3A/\varepsilon)^k$ (standard volume argument), the second factor is bounded by $N(3K \cap E, aL)$.

Now using $N(K, L_1 + L_2)$

$\leq \bar{N}(PK, PL_1)\; N((K - K - L_1) \cap E, L_2)\,, \quad (*)$

want to prove

$$N(K, rL) \leq \left(\frac{6A}{r-a}\right)^k \; N(PK, \frac{a}{2}PL).$$

for $a < r < A$ and projection P on $E^{\perp}$.

Using $(*)$ with $L_1 = aL$ and $L_2 = \varepsilon L$ we get

$$N(K, rL) \leq \bar{N}(PK, aPL)\; N((2K + aL) \cap E, \varepsilon L)\,.$$

$$\leq N\left(PK, \frac{a}{2}PL\right) N(((2A + a)L) \cap E, \varepsilon L)$$

$$\leq \left(\frac{6A}{r-a}\right)^k \; N(PK, \frac{a}{2}PL).$$

Proof of Theorem 1. Set $N_1 := \bar{N}(PK, PL_1)$. Then there are $z_i \in PK$, $i \le N_1$, such that

$$PK \subset \bigcup_{i=1}^{N_1} (z_i + PL_1).$$

For every $x \in K$ fix $i(x)$, $y_x \in PL_1$ such that

$$Px = z_{i(x)} + y_x$$

For $i \le N_1$ pick $\tilde{z}_i \in K$ such that $P\tilde{z}_i = z_i$, for every $y \in PL_1$ pick $\tilde{y} \in L_1$ such that $P\tilde{y} = y$.

Now for every $x \in K$ define

$$v(x) = \tilde{z}_{i(x)} + \tilde{y}_x \in \tilde{z}_{i(x)} + L_1,$$

$$w(x) = x - v(x) = x - \tilde{z}_{i(x)} - \tilde{y}_x.$$

Denote

$$T_i := K - L_1 - \tilde{z}_i, \qquad \text{for} \qquad i \le N_1.$$

Then $w(x) \in T_{i(x)}$ and

$$Pw(x) = Px - Pv(x) = Px - z_{i(x)} - y_x = 0.$$

Thus $w(x) \in T_{i(x)} \cap E$ and, hence,

$$x = w(x) + v(x) \in T_{i(x)} \cap E + \tilde{z}_{i(x)} + L_1.$$

This implies

$$K \subset \bigcup_{i=1}^{N_1} \Big(T_i \cap E + \tilde{z}_{i(x)} + L_1 \Big) .$$

Since for every $i \le N_1$,

$$N\left(T_i \cap E, L_2\right) \le \max_{z \in K} N\left(\left(K - L_1 - z\right) \cap E, L_2\right),$$

we obtain

$$N\left(K, L_1 + L_2\right) \le N_1 \max_{z \in K} N\left(\left(K - L_1 - z\right) \cap E, L_2\right).$$

Lower bound.

Recall, $\quad N(K, L_1 + L_2)$

$\leq \bar{N}(PK, PL_1)\; N((K - K - L_1) \cap E, L_2)\,, \quad (*)$

Theorem 2. *Let $t \in (0,1)$, K_1, K_2 be subsets of $\mathbb{R}^n$ and L_1, L_2 be symmetric convex bodies in $\mathbb{R}^n$. Let $P : \mathbb{R} \to \mathbb{R}$ be a projection and $E = \ker P$. Then*

$$N\Big(tK_1 + (1-t)K_2, (tL_1) \cap ((1-t)L_2)\Big)$$

$$\geq \bar{N}\Big(PK_1, 2PL_1\Big)\, \bar{N}\Big(K_2 \cap E, 2L_2 \cap E\Big).$$

In particular taking $K_1 = K_2$ we have

$$N\Big(K, (tL_1) \cap ((1-t)L_2)\Big)$$

$$\geq \bar{N}\Big(PK, 2PL_1\Big)\, \bar{N}\Big(K \cap E, 2L_2 \cap E\Big).$$

and taking $L_1 = ((1-t)/t)L_2$,

$N(K, L) \geq \bar{N}(tPK, 2PL)\, \bar{N}((1-t)K \cap E, 2L \cap E).$

We will need the notion of packing numbers. For K and L in $\mathbb{R}^n$ the packing number $P(K, L)$ of K by L is the maximal number M such that there exist vectors $x_i \in K$, $i \leq M$ satisfying

$$(x_i + L) \cap (x_j + L) = \emptyset \qquad \text{for every} \qquad i \neq j.$$

In other words, $x_i - x_j \notin L_0 := L - L$. Such set of points we also call L_0-separated set. It is well known that if L is symmetric convex body then

$$\bar{N}(K, 2L) \leq P(K, L) \leq N(K, L).$$

Proof of Theorem 2. Define N_1 and N_2 by

$$N_1 := P(PK_1, PL_1) \geq \bar{N}(PK_1, 2PL_1).$$

Then there are $z_1, \ldots, z_{N_1} \in PK_1$ such that $z_i - z_j \notin 2PL_1$ whenever $i \neq j$. For $1 \leq i \leq N_1$ pick $\tilde{z}_i \in K_1$ such that $P\tilde{z}_i = z_i$.

$$N_2 := P(K_2 \cap E, L_2 \cap E) \geq \bar{N}(K_2 \cap E, 2L_2 \cap E).$$

Then there exist $w_1, \ldots, w_{N_2}$ in $K_2 \cap E$ such that $w_k - w_\ell \notin 2L_2$ if $k \neq \ell$.

For every $i \leq N_1$ and $k \leq N_2$ denote $x_{i,k} := t\tilde{z}_i + (1-t)w_k$ and consider the set

$$\mathcal{A} = \left\{x_{i,k}\right\}_{i \leq N_1,\ k \leq N_2} \subset tK_1 + (1-t)K_2.$$

We show that $\mathcal{A}$ is well separated, namely

$$x_{i,k} - x_{j,\ell} \notin (2tL_1) \cap (2(1-t)L_2) \quad \text{if} \quad (i,k) \neq (j,\ell).$$

If $i \neq j$ then $P(x_{i,k} - x_{j,\ell}) = t(z_i - z_j) \notin 2tPL_1$, so $x_{i,k} - x_{j,\ell} \notin 2tL_1$. If $i = j$ then $k \neq \ell$ and $x_{i,k} - x_{j,\ell} = (1-t)(w_k - w_\ell) \notin 2(1-t)L_2$. Thus

$$P\Big(tK_1 + (1-t)K_2, (tL_1) \cap ((1-t)L_2)\Big)$$

$$\geq N_1 N_2 \geq \bar{N}(PK_1, 2PL_1)\bar{N}(K_2 \cap E, 2L_2 \cap E).$$

Application to Euclidean metric entropy: sharpness of Sudakov inequality.

Recall, given a convex body $K \subset \mathbb{R}$ with the origin in its interior,

$$M_K = M(K) = \int_{S^{n-1}} \|x\|_K \, d\nu$$

and

$$\ell(K) = \mathbb{E} \left\| \sum_{i=1}^{n} g_i e_i \right\|_K \leq \sqrt{n} M_K.$$

Sudakov inequality:

$$N(K, tB_2^n) \leq \exp\left(5 \left(\ell(K^0)/t\right)^2\right).$$

When it is sharp? In other words, when the covering number is big?

Recall, if we control diameter of a section then we control the covering number. Thus, if the covering number is big then every subspaces has a big diameter. We can quantify it as

Proposition. *Let* $R > 1$, $\eta > 0$, *and* $K \subset RL$ *be symmetric convex bodies in* $\mathbb{R}^n$ *such that*

$$N(K, L) \geq \exp(\eta n) .$$

Then for every k*-codimensional subspace* E *of* $\mathbb{R}^n$ *with*

$$k = \left[\frac{\eta\, n}{\ln(12R)} \right]$$

one has

$$K \cap E \not\subset \tfrac{1}{4}\, L.$$

Thus, if Sudakov inequality is almost sharp, i.e., if

$$N(K, B_2^n) \geq \exp\left(\varepsilon (M_K^*)^2 n \right),$$

for some $\varepsilon > 0$, then for

$$k = k_0 := \left[\frac{\varepsilon\, (M^*)^2 n}{\ln(12R(K))} \right]$$

every k-codimensional section of K has diameter at least $1/4$.

Usually, to improve covering of K by Euclidean balls we use truncations $K_\beta = K \cap \beta B_2^n$. The ideas above leads to the following intuition:

Let $\varepsilon > 0$, $\beta > 1$, and

$$N(K, B_2^n) \geq \exp\left(\varepsilon (M_K^*)^2 n\right),$$

then we have two distinct possibilities:

I. Either the covering number $N(K, B_2^n)$ can be significantly decreased by cutting K on the level β, (which means that "most" of the entropy of K comes from parts far from B_2^n);

II. or every k'-codimensional section of K has large diameter, for an appropriate choice of $k' > k_0$ depending on β.

Below we discuss the first case.

Theorem 3. *Let $K \subset \mathbb{R}$ be a symmetric convex body. Let $\rho > 0$ and $\beta \geq 4\rho$. Then*

$$N(K_\beta, 4\,\rho\, B_2^n) \leq \exp\left(2\,\left(\frac{\ell\left(K_\rho^0\right)}{\rho}\right)^2 \ln\frac{3\beta}{\rho}\right).$$

Remark. Note that $\ell\left(K_\rho^0\right)$ can be much smaller than $\ell\left(K_\beta^0\right)$ given by Sudakov inequality.

Idea of proof. To choose

$$k \approx (\ell(K_\rho^0)/\rho)^2$$

and to apply so-called "low M^*-estimate" saying that there exists k-codimensional subspace E such that

$$K \cap E \subset \rho B_2^n.$$

Then to apply entropy extension theorem.

Now we consider coverings without truncations. Define $F = F_K$ by

$$F(\rho) = \frac{\ell(K^0)}{\ell\left(K_\rho^0\right)}.$$

This function can be used to measure a possible gain in Sudakov estimates. Rewriting Sudakov inequality we get

$$N(K, 8\,\rho\, B_2^n) \le \exp\left(5\left(\frac{\ell(K_\rho^0)}{8\rho}\right)^2 F(\rho)^2\right),$$

which should be compared with the following:

Theorem 4. Let K be a symmetric convex body and $\rho > 0$. Then

$$N(K, 8\,\rho\, B_2^n) \le \exp\left(2\,\left(\frac{\ell\left(K_\rho^0\right)}{\rho}\right)^2\, \ln\left(6\, F(\rho)\right)\right).$$

Proof:

$$N(K, 8\,\rho\, B_2^n)$$

$$\leq N(K, (2K) \cap 2\beta B_2^n)\, N(2K_\beta, 8\,\rho\, B_2^n)$$

$$= N(K, 2\beta B_2^n)\, N(K_\beta, 4\,\rho\, B_2^n).$$

Applying Sudakov inequality to the first factor and Theorem 3 to the second one, we observe

$$N(K, 8\,\rho\, B_2^n)$$

$$\leq \exp\left(5\,\left(\frac{\ell(K^0)}{2\beta}\right)^2 + 2\,\left(\frac{\ell\left(K_\rho^0\right)}{\rho}\right)^2 \ln\frac{3\beta}{\rho}\right).$$

Optimizing in β, we obtain the result.