Analyse fonctionnelle/Functional Analysis

A Geometric Approach to Duality of Metric Entropy

Note de V. D. Milman et S. J. Szarek, présentée par Michel Talagrand.

Abstract - We introduce a new geometric approach for studying Duality of Metric Entropy (of operators, or of convex sets). We demonstrate such duality, up to a logarithmic factor, in one important case, for the hitherto not-too-well understood low "levels of resolution".

Dualité d'Entropie Métrique: une Approche Géométrique

Résumé - Nous introduisons une nouvelle approche à l'étude de Dualité d'Entropie Métrique (d'opérateurs, ou de convexes). Nous démontrons la dualité, à un facteur logarithmique près, dans un certain cas important et pour les niveaux faibles de résolution qui ont été mal compris jusqu'ici.

---

0. Version française abrégée - Pour deux sous-ensembles U et V d'un espace linéaire on définit le *nombre de récouvrement* $N(U,V) := \min\{N$: *U peut être couvert par N translates de V*$\}$. La version géométrique du problème bien connu de dualité d'entropie affirme que

Conjecture 1 *Soit n un entier positif et soient U et V deux corps convexes symétriques dans $\mathbb{R}^n$. Alors*

$$\log N(U^\circ, bV^\circ) \le a \log N(V,U). \tag{1}$$

où $a, b > 0$ sont des constantes universelles et K° désigne le polaire de K.

Cette conjecture s'énonce habituellement dans le langage des opérateurs compacts entre des epaces de Banach. Pour un tel opérateur, disons $u : X \to Y$, on definit le k-ième nombre d'entropie de u par

$$e_k(u) := \inf\{\varepsilon > 0 : N(u(B_X), \varepsilon B_Y) \le 2^k\}, \tag{2}$$

où B_Z désigne la boule unité de l'espace Z, et l'affirmation (1) devient alors

$$\forall k \in \mathbb{N} \quad b^{-1} e_{ak}(u) \le e_k(u^*) \le b e_{k/a}(u). \tag{3}$$

La conjecture 1 reste toujours ouverte, voir [1], [2], [7], [10] et leurs bibliographies pour des résultats partiels et voisins. Dans notre recherche nous avons identifié les deux affirmations géométriques suivantes qui, si elles étaient vraies, permettraient la solution du cas central où l'un des espaces X, Y est Hilbertien (ou l'un des corps U, Vest la boule unité euclidienne D).

Il existe une constante $c > 0$ telle que pour tout $n, N \in \mathbb{N}$ verifiant $\log N \le cn$ on a, pour tout polytope $K \subset \mathbb{R}^n$ (resp. $P \subset \mathbb{R}^n$),
(i) *si le nombre de sommets de K est au plus N et si $N(K,D) \le N$, alors $\frac{1}{4}D \not\subset K$.*
(ii) *si le nombre de faces de P est au plus N et si $N(D,P) \le N$, alors $P \not\subset 4D$.*

Nous avons démontré que l'affirmation (i) est vraie à un facteur logarithmique près. Plus précisément, nous avons

Théorème 3-4 *L'affirmation* (i) *au-dessus est vraie sous l'hypothèse plus forte* $\log N \leq c(1+\log n)^{-6} n$. *Par conséquent, pour tout opérateur compact à valeurs dans un espace de Hilbert et pour tout* $k \in \mathbb{N}$ *on a*

$$e_{ak}(u^*) \leq C(1+\log k)^3 e_k(u).$$

Ici, $C, c > 0$ et $a \geq 1$ *sont des constantes universelles.*

Il existe également une version géométrique de ce résultat dans l'esprit de la Conjecture 1 et des résultats formellement plus forts exprimés dans le langage d'*epaisseures moyennes* definies par (4) ci-dessous.

---

1. Background. If U, V are subsets of a vector space, one defines the *covering number* $N(U,V) := \min\{N : U$ *may be covered by* N *translates of* $V\}$. The geometric form of the over 25 years old "duality conjecture" for entropy numbers of operators asserts that

Conjecture 1 *There exist universal constants* $a, b > 0$ *such that whenever* $n \in \mathbb{N}$ *and* U, V *are symmetric convex bodies in* $\mathbb{R}^n$, *then* (1) *holds.*

In (1) and in what follows, K° denotes the polar body of K and all logarithms are to the base 2. The conjecture is often expressed as a statement about compact operators acting between Banach spaces. For such an operator, say, $u : X \to Y$, one defines *entropy numbers* (see [8] for basic results on this and related concepts) by (2) (B_Z is the unit ball of Z); the assertion of the conjecture takes then the form (3). We shall often use the notation $\Phi \lesssim \Psi$ (resp. $\gtrsim$) for universal estimates $\Phi \leq C\Psi$, where $C > 0$ stands for a numerical constant independent of any parameters implicit in the quantities Φ, Ψ (notably the dimension). Accordingly, one way to restate (3) would be $e_{ak}(u) \lesssim e_k(u^*) \lesssim e_{k/a}(u)$. Here and in what follows, numbers like ak and k/a should be thought of as integers: due to the asymptotic nature of the questions we investigate, the distinction between the number and its integer part or, usually, even its double, is immaterial.

Up to now, in spite of a substantial body of work, the duality conjecture has been verified only under very strong assumptions on *both* spaces. In a way, many of the results stem from the Hilbertian case where, by a simple spectral-theoretic argument, $u(B_X)$ and $u^*(B_{Y^*})$ are just *isometric.* For the most "up-to-date" results on the matter and further details on the current state of the knowledge we refer the reader to [1], [7], [10], the survey [2] and their references. Our approach is arguably the first – in this context – attempt at an analysis of a *general* n-dimensional convex set at multiple "levels of resolution". Details of most the arguments below can be found in [6]; the remaining few are a part of a work in progress.

2. The geometric statements. In our approach to the duality conjecture, we identified the following two geometric statements which, if proved, would together imply the special case of Conjecture 1 when *one* of the spaces is a Hilbert space, and the other arbitrary (or, equivalently, when *one* of the bodies U, V in (3) is the Euclidean ball $D = D_n \subset I\!R^n$, or any ellipsoid).

There exists a universal constant $c > 0$ *such that whenever* $n, N \in I\!N$ *verify* $\log N \leq cn$, *then, for any polytope* $K \subset I\!R^n$ *(resp.* $P \subset I\!R^n$*),*
(i) *if* K *has no more than* N *vertices and if* $N(K,D) \leq N$, *then* $\frac{1}{4}D \not\subset K$
(ii) *if* P *has no more than* N *faces and if* $N(D,P) \leq N$, *then* $P \not\subset 4D$.

The "half-Hilbertian" setting is of significant interest as it is relevant, e.g., to many contexts where "maximal functions" of some kind are used (regularity of Gaussian processes, convergence almost everywhere) or to areas such as coding theory. Moreover, we believe that once that case is resolved (in the affirmative, that is), the stage will be set for proving a "nearly" general version, paralleling the developments of ideas in [13] and [1], where "weaker" duality results were established.

At the first glance, statements of the above nature may appear "trivial". Indeed, on some meta-mathematical level, we are asking whether the "complexity" of an n-dimensional Euclidean ball is smaller than exponential in n. And our intuition says "no", no matter what exactly this complexity is supposed to mean. However, the more exact formulation brings together two rather different hypotheses: the small number of vertices of the polytope K (resp., faces, P) and the small cardinality of the covering family. These data are not easy to combine; a difficulty which also arouses interest.

3. The mean width. In this note we analyze the first of the two geometric conjectures stated above (i.e., part (i)), to which we shall refer as the "Geometric Lemma". We first present its more precise and quantitative (but equivalent) version, for which we need to introduce some notation. First, if $U \subset I\!R^n$ is a compact symmetric convex body containing the origin in the interior, one denotes by $\|\cdot\|_U$ its *Minkowski functional*, i.e. the norm, for which U is the unit ball. We shall use the same notation for gauges of *nonsymmetric* sets. We also set

$$M^*(U) := \int_{S^{n-1}} \sup_{y \in U} \langle x, y \rangle \, d\mu_n(x) = \int_{S^{n-1}} \|x\|_{U^\circ} \, d\mu_n(x) \tag{4}$$

where μ_n is the normalized (i.e., probability) Lebesgue measure on S^{n-1}. The quantity $M^*(U)$ equals $1/2$ of the *mean width* of U, a well known geometric parameter. We can now state (# stands for cardinality)

Conjecture 2 *Given* $S \subset I\!R^n$, *set* $K = \mathrm{conv}\, S$ *and* $k = \max\{\log \#S, \log N(K,D)\}$. *Then*

$$M^*(K \cap D) \lesssim \sqrt{k/n}\,. \tag{5}$$

Clearly Conjecture 2 $\Rightarrow$ Geometric Lemma: if $\gamma D \subset K$, then $M^*(K \cap D) \geq M^*(\gamma D) = \gamma$, which is inconsistent with (5) if k/n is small. The reverse implication is less immediate; we shall sketch its proof in §5.

4. The results. Our main result asserts that Conjecture 2 (and the Geometric Lemma) hold "up to a logarithmic factor". We have

Theorem 3 *If $n \in \mathbb{N}$, $S, K \subset \mathbb{R}^n$ and k are as in* Conjecture 2 *, then $M^*(K \cap D) \lesssim (1+\log k)^3 \sqrt{k/n}$. In particular, there exists a universal constant $c > 0$ such that if $k \leq c(1+\log n)^{-6} n$, then $M^*(K \cap D) < \frac{1}{4}$ (and hence $\frac{1}{4}D \not\subset K$).*

The Geometric Lemma as stated in previous sections remains thus a conjecture. However, we did check it for various (classes of) convex bodies, including ℓ_p^n-balls of arbitrary radii and some "random bodies" (specifically, "generic" projections of ℓ_1^N-balls, a "canonical" counterexample to many problems in high dimensional convexity, cf. [4]).

And here is a duality result for covering numbers and entropy numbers, which is a corollary of Theorem 3.

Theorem 4 *There exist universal constants $a, C > 0$ such that, for all $n \in \mathbb{N}$, all convex sets $K \subset \mathbb{R}^n$ and all k,*

$$\log N(K, D) \leq k \;\Rightarrow\; \log N(D, \omega K^\circ) \leq ak.$$

where $\omega = \omega(k) = C(1+\log k)^3$. Similarly, for a compact Hilbert-space-valued operator u,

$$e_{ak}(u^*) \leq \omega e_k(u),$$

for all $k \in \mathbb{N}$ (with the same ω).

In the remainder of this note we sketch the implications Geometric Lemma $\Rightarrow$ Conjecture 2, Conjecture 2 $\Rightarrow$ (the case $U = D$ of) Conjecture 1 (we note that a similar but more quantitative and more precise reasoning allows to deduce Theorem 4 from Theorem 3), and hint at some ingredients of the proof of Theorem 3. The arguments are quite involved, and they invoke a surprisingly wide spectrum of methods and ideas of local theory of Banach spaces. We reiterate that, on the technical level, the principal difficulty appears to lie in combining the two quite different kinds of hypotheses: the control of the number of vertices of K (say, $\log \#S \ll n$, where by $\ll$ we mean "much smaller than"; essentially a negation of $\gtrsim$) and that of the covering number $N(K, D)$ (say, $\log N(K, D) \ll n$). On the more methodological level, we believe that the complexity of the proofs reflects obstructions inherent in the general problem. Indeed, as we indicated in the paragraph following (3), the conjecture holds if *both* spaces (resp., bodies) are "sufficiently close" to the Hilbert space (resp., ellipsoids). Similarly, if $k \gtrsim \text{rank } u$ (in the language of (3), or $k \gtrsim n$ in (1)), entropy duality holds ([3], see also [10]) and is essentially a "volumetric" statement related to (the highly nontrivial) inverse Santaló and inverse Brunn-Minkowski inequalities. However, the validity of (3) for $k \ll \text{rank } u$, or (1) for $\log N(K, D) \ll n$, seems to be a much more delicate question: there is no manifest reason why convex bodies and their polars should be so closely related on so fine a scale. In this context, our *up-to-logarithmic-factor* results do suggest at least some of the needed "connections" and hopefully prepare the ground for further progress.

5. The implication Geometric Lemma $\Rightarrow$ Conjecture 2 is based on the following "relative" of the Dvoretzky Theorem (see, e.g., [5], Lemma 2.1 and its proof), needed here *without* the usual symmetry hypothesis.

Proposition 5 *There exists a numerical constant $c > 0$ such that if $m \le n$ and a convex set $V \subset D \subset \mathbb{R}^n$ verify $\sqrt{m/n} \le cM^*(V)$, then, for a generic orthogonal projection P of rank m, $\frac{1}{2}M^*(V)PD \subset PV \subset 2M^*(V)PD$.*

Above and elsewhere in the paper, "generic" means "except on a small exceptional set". More specifically, the projections are considered here as elements of the Grassmann manifold $G_{n,m}$ and "small" may mean "of (the normalized Haar) measure $\le \exp(-c'm)$", where $c' > 0$ is a universal numerical constant.

Assuming Proposition 5, we now argue as follows. Suppose K is a "counterexample" to Conjecture 2, i.e., $1 \ge M^*(K \cap D) \gg \sqrt{k/n}$. Without loss of generality we may assume that $M^*(K \cap (D + x))$ attains maximum at $x = 0$ (which is necessarily the case if $K = -K$). Now apply the Proposition with $V = K \cap D$ and $m = \lfloor (cM^*(K \cap D))^2 n \rfloor$; note that $m \gg k$. This yields $K_0 = PK = \operatorname{conv} PS$, of which we may think to be contained in $\mathbb{R}^m$, such that $K_0 \supset P(K \cap D_n) \supset \frac{1}{2}M^*(K \cap D_n)D_m$ while, for each $x \in \mathbb{R}^n$, $P(K \cap (D_n + x)) \subset Px + 2M^*(K \cap D_n)D_m$ and so K_0 can be covered by 2^k balls of radius $2M^*(K \cap D_n)$. Accordingly, if we set $K_1 = (2M^*(K \cap D_n))^{-1}K_0 \subset \mathbb{R}^m$, then $\log N(K_1, D_m) \le k$ and $K_1 \supset \frac{1}{4}D_m$ although $m \gg k$, contradicting the Geometric Lemma for $n = m$.

We note that the "dimension reduction" trick of the type presented above, and particularly its use of the simultaneous control of the covering number and of the parameter $M^*(\cdot)$, is quite representative for our arguments.

6. The implication Conjecture 2 $\Rightarrow$ Conjecture 1 (the case $U = D$). It is here where an entropy duality result is deduced from a statement about the mean width. We need the following auxiliary result in the spirit of [5], [7] and [12].

Proposition 6 *Let $k \le n$, and let $K \subset \mathbb{R}^n$ be a symmetric convex body verifying $\log N(K, D) \le k$ and $M^*(K \cap D) \le \sqrt{k/n}$. Then, for a generic rank k orthogonal projection P, we have*

$$K \cap \sqrt{k/n}\, P^{-1}D \subset C_0 D\,, \tag{6}$$

where $C_0 > 0$ is a universal numerical constant.

Now consider the setting of Conjecture 1 with $U = D$. Putting $\kappa := \log N(V, D)$, we need to show that $\log N(D, bV^\circ) \lesssim \kappa$. Since, as we noted at the end of §4, entropy duality *does* hold if $\log N(\cdot,\cdot) \gtrsim$ the dimension, we only need to consider the case $\kappa \ll n$. Let S be a set of cardinality $\le 2^\kappa$ such that $V \subset S + D$; we may assume that $S \subset V$ and, at the price of replacing κ by $\kappa + 1$, that S is symmetric. Let $K = \operatorname{conv} S$. An elementary argument shows that, for any $\rho > 0$, $N(D, (\rho + 2)V^\circ) \le N(D, \rho K^\circ)$ and so we just need to establish that $\log N(D, \rho K^\circ) \lesssim \kappa$ for some (universal) $\rho > 0$. Noting that $N(K, D) \le N(V, D)$ and appealing to Conjecture 2 (which we assume), we see that the assumptions of Proposition 6 hold for the present choice of K and some $k \lesssim \kappa$. Let P be a (generic rank k) projection guaranteed by the Proposition and let F be its range. We then have

$$C_0^{-1}D \subset \operatorname{conv}\left(K^\circ \cup \sqrt{n/k}\,(D \cap F)\right) \;\subset\; K^\circ + \sqrt{n/k}\,(D \cap F),$$

the first inclusion being just the dualized assertion of the Proposition. Hence

$$\begin{aligned} N(D, \rho K^\circ) &= N(C_0^{-1}D, C_0^{-1}\rho K^\circ) \\ &\leq N(K^\circ + \sqrt{n/k}(D \cap F), C_0^{-1}\rho K^\circ) \\ &= N(\sqrt{n/k}(D \cap F), (\rho/C_0 - 1)K^\circ). \end{aligned} \tag{7}$$

Set $\tau = 2(\rho/C_0 - 1)^{-1}$, then the last quantity in (7) is $\leq N(\tau\sqrt{n/k}(D \cap F), K^\circ \cap F)$, which involves covering numbers *inside* the k-dimensional space F. Using again the fact that entropy duality does hold if $\log N(\cdot,\cdot) \gtrsim$ the dimension, and observing that the polar of $K^\circ \cap F$ (inside F) is PK, we see that the logarithm of that quantity is $\lesssim \kappa$ whenever $\log N(PK, \beta\sqrt{k/n}\, D) \lesssim \kappa$ (for some $\beta > 0$; with τ, hence ρ, depending on β), which in turn can be deduced by an argument similar to that of §5.

7. The proof of Theorem 3. This is by far the most involved part of the argument and so we shall only hint the "highlights". The functional $M^*(\cdot)$ can be related to Gaussian processes and ideal norms. Specifically, for $U \subset I\!R^n$, $M^*(U)$ is, up to a normalizing factor, the same as the expected value of the supremum of the Gaussian process indexed by U (or the so-called ℓ-norm of the formal identity considered as acting from ℓ_2^n to $(I\!R^n, \|\cdot\|_{U^\circ})$). Not surprisingly, our arguments parallel those used in the study of regularity of Gaussian processes (e.g., [11], [12]). The tools include the so-called "Sudakov minoration", which – in our setting – asserts that, for $U \subset I\!R^n$, $\log N(U, \varepsilon D) \lesssim (M^*(U^\circ)/\varepsilon)^2 n$ and its "constructive" variant, "Maurey's lemma" (see [9]), which says that, additionally, the covering family of balls may be chosen so that the centers are averages of "few" extreme points of U. At the first glance it may seem that these estimates "go the wrong way" as it is the $M^*(\cdot)$ quantity that needs to be majorized, but in fact we do use them as ingredients in a "bootstrap" scheme, involving iteration and a priori estimates. The iteration procedure implements a "multiresolution analysis" of the set K, the aim of which is to combine contributions to $M^*(K \cap D)$ coming from different levels of resolution. A single step of the iteration uses the following fact, which we shall state as it is exactly there where the two assumptions, small number of vertices of K and smallness of $N(K, D)$, are being put together.

Proposition 7 *If $R > 0$, $n \in I\!N$ and $K = \mathrm{conv}\, S \subset RD \subset I\!R^n$, then*

$$M^*(K \cap D) \lesssim \left(R\sqrt{\frac{\log N(K,D)}{n}}\sqrt{\frac{\log \#S}{n}} \right)^{1/2}.$$

Acknowledgment. This research was partially supported by grants from the US-Israel BSF (both authors) and the NSF [U.S.A.] (the second named author).

References

[1] J. Bourgain, A. Pajor, S. J. Szarek and N. Tomczak-Jaegermann, *On the duality problem for entropy numbers of operators.* In *Geometric aspects of functional analysis (1987–88)*, 50–63, Edited by J. Lindenstrauss and V. D. Milman. Lecture Notes in Math., **1376**, Springer, Berlin-New York, 1989.

[2] E. D. Gluskin, Y. Gordon and A. Pajor, *Entropy, approximation numbers and other parameters.* In *Handbook on the Geometry of Banach spaces*, W. B. Johnson, J. Lindenstrauss eds., Elsevier Science, *to appear*

[3] H. König and V. Milman, *On the covering numbers of convex bodies.* In *Geometrical aspects of functional analysis (1985/86)*, 82–95, Lecture Notes in Math., **1267**, Springer, Berlin-New York, 1987.

[4] P. Mankiewicz and N. Tomczak-Jaegermann, *Random Banach spaces.* In *Handbook on the Geometry of Banach spaces*, W. B. Johnson, J. Lindenstrauss eds., Elsevier Science, *to appear*

[5] V. Milman, *A note on a low M^*-estimate.* Geometry of Banach spaces (Strobl, 1989), 219–229, London Math. Soc. Lecture Note Ser., **158**, Cambridge Univ. Press, Cambridge, 1990.

[6] V. Milman and S. J. Szarek, *A Geometric Lemma and Duality of Entropy Numbers.* Erwin Schrödinger Institute, Vienna, preprint # 687, April 1999, 40 p. Available at *http://www.esi.ac.at/ESI-Preprints.html* To appear.

[7] A. Pajor and N. Tomczak-Jaegermann, *Volume ratio and other s-numbers of operators related to local properties of Banach spaces.* J. Funct. Anal. **87** (1989), no. 2, 273–293.

[8] A. Pietsch, *Eigenvalues and s-numbers.* Cambridge Studies in Advanced Mathematics, **13**. Cambridge University Press, Cambridge-New York, 1987.

[9] G. Pisier, *Remarques sur un résultat non publié de B. Maurey.* (French) Seminar on Functional Analysis, 1980–1981, Exp. No. V, 13 pp., Ecole Polytech., Palaiseau, 1981.

[10] G. Pisier, *A new approach to several results of V. Milman.* J. Reine Angew. Math. **393** (1989), 115–131.

[11] M. Talagrand, *A simple proof of the majorizing measure theorem.* Geom. Funct. Anal. **2** (1992), no. 1, 118–125.

[12] M. Talagrand, *Sudakov-type minoration for Gaussian chaos processes.* Israel J. Math. **79** (1992), no. 2-3, 207–224.

[13] N. Tomczak-Jaegermann, *Dualité des nombres d'entropie pour des opérateurs à valeurs dans un espace de Hilbert.* C. R. Acad. Sci. Paris Sér. I Math. **305** (1987), no. 7, 299–301.

Department of Mathematics, Tel Aviv University, Ramat Aviv, Israel
E-mail: vitali@math.tau.ac.il

Department of Mathematics, Case Western Reserve University, Cleveland, OH 44106-7058, U.S.A. and *Université Pierre et Marie Curie, Equipe d'Analyse, Bte 186, 4, Place Jussieu, 75252 Paris, France*
E-mail : sjs13@cwru.edu