Convex Geometric Analysis
MSRI Publications
Volume **34**, 1998

Metric Entropy of the Grassmann Manifold

ALAIN PAJOR

Abstract. The knowledge of the metric entropy of precompact subsets of operators on finite dimensional Euclidean space is important in particular in the probabilistic methods developped by E. D. Gluskin and S. Szarek for constructing certain random Banach spaces. We give a new argument for estimating the metric entropy of some subsets such as the Grassmann manifold equipped with natural metrics. Here, the Grassmann manifold is thought of as the set of orthogonal projection of given rank.

1. Introduction and Notation

Let A be a precompact subset of a metric space (X, τ). An ε-net of A is a subset Λ of X such that any point x of A can be approximated by a point y of Λ such that $\tau(x, y) < \varepsilon$. The smallest cardinality of an ε-net of A is called the *covering number* of A and is denoted by $N(A, \tau, \varepsilon)$. The metric entropy (shortly the entropy) is the function $\log N(A, \tau, \,.\,)$.

When X is a d-dimensional normed space equipped with the metric associated to its norm $\|\,.\,\|$, we will denote by $N(A, \|\,.\,\|, \varepsilon)$ the covering number of a subset A of X and by $B(X)$ the unit ball of X. The metric entropy of a ball $A = rB(X)$ of radius r is computed by volumic method (see [MS] or [P]): for $\varepsilon \in]0, r]$,

$$\left(\frac{r}{\varepsilon}\right)^d \leq N(rB(X), \|\,.\,\|, \varepsilon) \leq \left(3\frac{r}{\varepsilon}\right)^d. \tag{1}$$

The space $\mathbb{R}^n$ is equipped with its canonical Euclidean structure and denoted by ℓ_2^n. Its unit ball is denoted by B_2^n, the Euclidean norm by $|\,.\,|$ and the scalar product by $(\,.\,,\,.\,)$. For any linear operator T between two Euclidean spaces and any p, $1 \leq p \leq \infty$, let

$$\sigma_p(T) = \left(\sum_{i\geq 1} |s_i(T)|^p\right)^{1/p}$$

Part of this work was done while the author visited MSRI. The author thanks the Institute for its hospitality.

where $s_i(T)$, $i = 1, \dots$ denote the singular values of T. In particular, σ_1 is the nuclear norm, σ_2 is the Hilbert-Schmidt norm and σ_∞ the operator norm. The Schatten trace class of linear mapping on the n-dimensional Euclidean space equipped with the norm σ_p is denoted by S_p^n. More generally, we consider a unitarily invariant norm τ on $L(\ell_2^n)$; it satisfies $\tau(USV) = \tau(S)$ for any $S \in L(\ell_2^n)$ and any isometries U, V on ℓ_2^n. It is associated to a 1-symmetric norm on $\mathbb{R}^n$ and τ is the norm of the n-tuple of singular values. For any Euclidean subspaces E, F of ℓ_2^n, τ induces a unitarily invariant norm on $L(E, F)$ still denoted by τ.

Let $G_{n,k}$ be the Grassmann manifold of the k-dimensional subspaces of $\mathbb{R}^n$. For any subspace $E \in G_{n,k}$ we denote by P_E the orthogonal projection onto E. We will denote by σ_p the metric induced on $G_{n,k}$ by the norm σ_p, when $G_{n,k}$ is considered as a subset of S_p^n. Similarly, we denote by τ the metric induced onto $G_{n,k}$ by the norm τ.

We are mainly interested in estimating $N(B(S_1^n), \sigma_p, \varepsilon)$. The computation of $N(G_{n,k}, \sigma_p, \varepsilon)$ was done in [S1] and is the basic tool in [S2] for solving the finite dimensional basis problem (see also [G1], [G2] and [S3]). For computing the metric entropy of a d-dimensional manifold, we first look for an *atlas*, a collection of charts $(U_i, \varphi_i)_{1 \le i \le N}$ and estimate N. The situation is particularly simple if for each of the charts, φ_i is a bi-Lipschitz correspondence with a d-dimensional ball. Locally, the entropy is computed by volumic method. To estimate the number N of charts in the atlas, we look at the Grassmann manifold in $L(\mathbb{R}^n)$ with the right metric. Such an embedding that does not reflect the dimension of the manifold cannot give directly the right order of magnitude of the entropy, but as we will see, it gives an estimate of the cardinality N of a "good" atlas. The two arguments are combined to give the right order of magnitude for the entropy of $G_{n,k}$. We did not try to give explicit numerical constant and the same letter may be used to denote different constants.

2. Basic Inequalities

Let G be a Gaussian random d-dimensional vector, with mean 0 and the identity as covariance matrix. The following result is the geometric formulation of Sudakov's minoration (See [P]).

Lemma 1. *There exists a positive constant c such that for any integer $d \ge 1$, any subset A of $\mathbb{R}^d$ and for every $\varepsilon > 0$, we have*

$$\varepsilon \sqrt{\log N(A, |\,.\,|, \varepsilon)} \le c \mathbb{E} \sup_{t \in A} (G, t). \tag{2}$$

Lemma 2. *There exists a positive constant c such that for any integer $n \ge 1$, for any p such that $1 \le p \le \infty$ and for every $\varepsilon > 0$, we have*

$$\log N(B(S_p^n), \sigma_2, \varepsilon) \le c \, \frac{n^{3-2/p}}{\varepsilon^2}. \tag{3}$$

Proof. Let A be a subset of the Euclidean space S_2^n equipped with its scalar product given by the trace. Applying the inequality (2) where G is a standard Gaussian matrix whose entries are independent $\mathcal{N}(0,1)$ Gaussian random variables, we get that

$$\varepsilon\sqrt{\log N(A,\sigma_2,\varepsilon)} \le c\mathbb{E}\sup_{T\in A}\operatorname{trace}(GT).$$

Let q such that $1\le q\le\infty$ and $\frac{1}{p}+\frac{1}{q}=1$. By the trace duality,

$$|\operatorname{trace}(GT)| \le \sigma_q(G)\sigma_p(T) \le \sigma_q(G),$$

for any $T\in B(S_p^n)$. Now it is well known that

$$\mathbb{E}\sigma_q(G) \le n^{1/q}\mathbb{E}\sigma_\infty(G) \le \alpha n^{1/q}\sqrt{n}, \tag{4}$$

for some universal constant α (see [MP], Proposition 1.5.). Therefore

$$\sqrt{\log N(B(S_p^n),\sigma_2,\varepsilon)} \le c\alpha n^{1/q}\sqrt{n}\,\frac{1}{\varepsilon},$$

which gives the estimate (3). □

The next inequality is in some sense dual to (2). Again G is a Gaussian d-dimensional random vector.

Lemma 3 (see [PT]). *There exists a positive constant c such that for any integer $d\ge 1$, any norm $\|\,.\,\|$ on $\mathbb{R}^d$ and for every $\varepsilon>0$, we have*

$$\varepsilon\sqrt{\log N(B_2^d,\|\,.\,\|,\varepsilon)} \le c\mathbb{E}\|G\|. \tag{5}$$

Lemma 4. *There exists a positive constant c such that for any integer $n\ge 1$ and for any q such that $1\le q\le\infty$, we have*

$$\log N(B(S_2^n),\sigma_q,\varepsilon) \le c\,\frac{n^{1+2/q}}{\varepsilon^2} \tag{6}$$

for every $\varepsilon>0$.

Proof. The proof follows from the formulae (4) and (5) applied with the norm σ_q. □

Proposition 5. *There exists a positive constant c such that for any integer $n\ge 1$, for any p,q such that $1\le p\le\infty$, $2\le q\le\infty$ and for every $\varepsilon>0$, we have*

$$\log N(B(S_p^n),\sigma_q,\varepsilon) \le c\,\frac{n^{(2-1/p-1/q)q'}}{\varepsilon^{q'}}\,, \tag{7}$$

where $1/q+1/q'=1$.

Proof. We observe that

$$N(B(S_p^n),\sigma_q,\varepsilon)\le N\left(B(S_p^n),\sigma_2,\frac{\varepsilon}{\theta}\right)N(B(S_2^n),\sigma_q,\theta).$$

Therefore inequalities (3) and (6) give us

$$\log N(B(S_p^n),\sigma_q,\varepsilon)\le c\left(\frac{n^{3-2/p}\theta^2}{\varepsilon^2}+\frac{n^{1+2/q}}{\theta^2}\right).$$

Optimizing with $\theta^2=\varepsilon n^{-1+1/p+1/q}$ we arrive at

$$\log N(B(S_p^n),\sigma_q,\varepsilon)\le 2c\,\frac{n^{2-1/p+1/q}}{\varepsilon}\,. \tag{8}$$

In particular for $q=\infty$, we get

$$\log N(B(S_p^n),\sigma_\infty,\varepsilon)\le 2c\,\frac{n^{2-1/p}}{\varepsilon}\,. \tag{9}$$

Let now $2\le q\le\infty$, $\lambda=\frac{2}{q}$, so that $0\le\lambda\le 1$ and $\frac{1}{q}=\lambda\frac{1}{2}+(1-\lambda)\frac{1}{\infty}$. By Hölder's inequality, for every $x,y>0$, we have

$$N(B(S_p^n),\sigma_q,2x^\lambda y^{1-\lambda})\le N(B(S_p^n),\sigma_2,x)N(B(S_p^n),\sigma_\infty,y).$$

This relation and inequalities (3) and (9) yield

$$\log N(B(S_p^n),\sigma_q,\varepsilon)\le 8c\left(\frac{n^{3-2/p}}{z^{2/\lambda}}+\frac{n^{2-1/p}}{(\varepsilon/z)^{1/1-\lambda}}\right),\ z>0$$

and the optimal choice $z=(q-2)^{(q-2)/q(q-1)}\,\varepsilon^{1/(q-1)}\,n^{(1-1/p)(q-2)/q(q-1)}$ gives the estimate (7). □

Remarks. 1) Estimates (7) and (8) are relevant when $p\le 2\le q$, their accuracy depends on the range of ε, particularly with respect to $1/n^{(1/p-1/q)}$.

2) Note that when $p=1$, inequality (7) becomes

$$\log N(B(S_1^n),\sigma_q,\varepsilon)\le c\,\frac{n}{\varepsilon^{q'}}.$$

3) All the computation of entropy above, could have be done by the same method for the trace classes of operators between two different finite dimensional Euclidean spaces. One can use Chevet inequality [C] to get the relation corresponding to (4).

3. Metric Entropy of the Grassmann Manifold

We consider now the Grassmann manifold $G_{n,k}$ as a subset of S_q^n, which means that $G_{n,k}$ is equipped with the metric $\sigma_q(E,F)=\sigma_q(P_E-P_F)$. In view of evaluating the cardinality of a "good" atlas of the Grassmann manifold, we begin with a first estimate for its entropy.

Proposition 6. *There exists a positive constant c such that for any integers $1 \le k \le n$ and for every $\varepsilon > 0$, we have*

$$\log N(G_{n,k}\,, \sigma_\infty, \varepsilon) \le c\,\frac{d}{\varepsilon}, \tag{10}$$

where $d = k(n-k)$.

Proof. We may suppose that $k \le n-k$. Since for any rank k orthogonal projection $\sigma_1(P) = k$, inequality (7) with $p = 1$ gives by homogeneity,

$$\begin{aligned}\log N(G_{n,k}\,, \sigma_\infty, \varepsilon) &\le \log N(B(S_1^n), \sigma_\infty, \varepsilon k^{-1})\\ &\le c\,\frac{nk}{\varepsilon} \le 2c\,\frac{d}{\varepsilon}.\end{aligned}$$

□

Let $E, F \in G_{n,k}$ and $U \in \mathcal{O}_n$ such that $U(F) = E$, then clearly

$$(P_{F^\perp} P_E)^* = (P_{F^\perp} U P_F U^*)^* = U P_F U^* P_{F^\perp} = U P_F P_{E^\perp} U^*.$$

Therefore $P_{F^\perp} P_E$ and $P_F P_{E^\perp}$ have the same singular values. Moreover, since $P_E - P_F = P_{F^\perp} P_E - P_F P_{E^\perp}$ is an orthogonal decomposition, for any unitarily invariant norm τ we have $\tau(P_{F^\perp} P_E) = \tau(P_F P_{E^\perp})$ and

$$\tau(P_{F^\perp} P_E) \le \tau(E, F) \le 2\tau(P_{F^\perp} P_E)\,.$$

Denote by $P_{F|E}$ the restriction over E of the orthogonal projection onto F and consider its polar decomposition

$$P_{F|E} = \sum_{i=1}^k s_i e_i \otimes f_i,\ \ s_i \ge 0,\ \ (e_i, f_j) = \delta_{ij} s_i, 1 \le i, j \le k,$$

with $(e_i)_{1\le i\le k}$ an orthonormal basis of E and $(f_i)_{1\le i\le k}$ an orthonormal basis of F. Let $e_i' = (-s_i e_i + f_i)/\sqrt{1-s_i^2}$ and $f_i' = (-e_i + s_i f_i)/\sqrt{1-s_i^2}$ when i is such that $s_i \ne 1$. The families (e_i') and (f_i') are respectively orthonormal systems in $E^\perp$ and $F^\perp$ and we have the following polar decomposition:

$$P_E - P_F = P_F P_{E^\perp} - P_E P_{F^\perp} = \sum_{s_i \ne 1} \sqrt{1-s_i^2}\, e_i' \otimes f_i + \sum_{s_i\ne 1} \sqrt{1-s_i^2}\, e_i \otimes f_i'\,.$$

Consequently, for $1 \le q \le \infty$

$$\sigma_q(E,F) = (\sigma_q(P_{F^\perp|E}))^q + \sigma_q(P_{F|E^\perp})^q)^{1/q} = \left(2\sum_{i=1}^k (1-s_i^2)^{q/2}\right)^{1/q}.$$

Note that the Riemannian metric is given by $\sigma_g(E,F) = \left(\sum_{i=1}^k \arccos^2 s_i\right)^{1/2}$ and therefore

$$\sigma_2(E,F) \le \sqrt{2}\sigma_g(E,F) \le \frac{\pi}{2}\,\sigma_2(E,F).$$

Let E be a k-dimensional Euclidean subspace of $\mathbb{R}^n$. For any $0 < \rho < 1$, let

$$V_\rho(E) = \{F \in G_{n,k} : \sigma_\infty(E,F) \le \rho\}.$$

Let us recall a standard chart: $(V_\rho(E), \varphi_E)$ where $\varphi_E : V_\rho(E) \longrightarrow L(E, E^\perp)$ is defined by

$$\varphi_E(F) = P_{E^\perp|F} \circ (P_{E|F})^{-1}.$$

In other words, F is the graph of $u = \varphi_E(F)$ and $F = \{x + u(x) : x \in E\}$. With this notation, we have:

Lemma 7. *Let $0 < \rho < 1$, $E, F, G \in G_{n,k}$ and $F, G \in V_\rho(E)$, let $u = \varphi_E(F)$ and $v = \varphi_E(G)$, let τ be a unitarily invariant norm on $L(\ell_2^n)$, then we have*

$$\varphi_E(V_\rho(E)) = \left\{ w \in L(E, E^\perp) : \sigma_\infty(w) \le \frac{\rho}{\sqrt{1-\rho^2}} \right\}, \tag{11}$$

$$2^{-1}\tau(F,G) \le \tau(u-v) \le \frac{2^{1/2}}{1-\rho^2}\, \tau(F,G). \tag{12}$$

Proof. Let $H \in V_\rho(E)$ and $w = \varphi_E(H)$. The relation (11) follows immediately from

$$\sigma_\infty(P_{E^\perp}P_H) = \sigma_\infty(E,H) = \sup_{|x|=1} |w(x)|/\sqrt{1+|w(x)|^2} = \sigma_\infty(w)/\sqrt{1+\sigma_\infty(w)^2}.$$

Recall that $F = \{x + u(x) : x \in E\}$ and $G = \{x + v(x) : x \in E\}$. To prove the second relation, let $x, y \in E$ such that $P_G(x + u(x)) = y + v(y)$, then

$$|P_{G^\perp}(x + u(x))| = |x - y + u(x) - v(y)| \ge |x - y|.$$

Therefore

$$\begin{aligned} |u(x)-v(x)|^2 &= |(x+u(x))-(x+v(x))|^2 \\ &= |P_{G^\perp}(x+u(x))|^2 + |(y-x)+v(y-x)|^2 \\ &\le (2+\sigma_\infty(v)^2)|P_{G^\perp}(x+u(x))|^2, \end{aligned}$$

and for every x in E we have

$$|P_{G^\perp}(x + u(x))| \le |u(x) - v(x)| \le (2 + \sigma_\infty(v)^2)^{1/2}|P_{G^\perp}(x + u(x))|.$$

The left-hand side inequality means that

$$|P_{G^\perp}P_F(z)| \le |(u - v)(P_E z)| \quad \text{for any} \quad z \in F.$$

It is well known that if $S, T \in L(\ell_2^n)$ satisfy $|Sx| \le |Tx|$ for every x then $\tau(S) \le \tau(T)$. Hence

$$\tau(P_{G^\perp}P_F) \le \tau(u - v).$$

Applying the same observation to the operators $S = (u - v)P_E$ and

$$T = (2 + \sigma_\infty(v)^2)^{1/2} P_{G^\perp}P_F(P_E + uP_E)$$

and using the right-hand side of the same inequality above, one gets

$$\tau(u - v) \le (2 + \sigma_\infty(v)^2)^{1/2}(1 + \sigma_\infty(u)^2)^{1/2}\tau(P_{G^\perp}P_F).$$

We conclude using (11) and the relation between $\tau(P_{G^\perp}P_F)$ and $\tau(E,F)$. $\square$

We now give a new proof of a result of Szarek.

PROPOSITION 8 (see [S2]). *For any integers $1 \le k \le n$ such that $k \le n-k$, for any q such that $1 \le q \le \infty$ and for every $\varepsilon > 0$, we have*

$$\left(\frac{c}{\varepsilon}\right)^d \le N(G_{n,k}, \sigma_q, \varepsilon k^{1/q}) \le \left(\frac{C}{\varepsilon}\right)^d, \tag{13}$$

where we set $d = k(n-k)$ and $c, C > 0$ are universal constants.

PROOF. The relation (11) of lemma 7 shows that if we fix ρ, say $\rho = 1/2$, then φ_E is a bi-Lipschitz correspondence from the "ball" $V_{1/2}(E)$ onto the ball of $L(E, E^\perp)$ of radius $1/\sqrt{3}$ (in the operator norm) and from (12), the Lipschitz constants are universal. Therefore the metric entropy of $V_{1/2}(E)$ for the metric σ_∞ is equivalent to the entropy of a d-dimensional ball of radius $1/\sqrt{3}$ for its own metric. From (1) we get

$$\left(\frac{c_1}{\varepsilon}\right)^d \le N(V_{1/2}(E), \sigma_\infty, \varepsilon) \le \left(\frac{c_2}{\varepsilon}\right)^d,$$

for some positive universal constants c_1 and c_2. From inequality (10) of Proposition 6, there is an atlas $(V_{1/2}(E_i), \varphi_{E_i})_{1 \le i \le N}$ with

$$\log N \le \log N(G_{n,k}, \sigma_\infty, 1/2) \le 2c\, d.$$

Since for a fixed k-dimensional subspace E,

$$N(G_{n,k}, \sigma_\infty, \varepsilon) \le N(G_{n,k}, \sigma_\infty, 1/2) N(V_{1/2}(E), \sigma_\infty, \varepsilon),$$

we get

$$\left(\frac{c_1}{\varepsilon}\right)^d \le N(G_{n,k}, \sigma_\infty, \varepsilon) \le \left(\frac{c_2 e^{2c}}{\varepsilon}\right)^d. \tag{14}$$

The computation of $\sigma_q(\,.\,)$ on $G_{n,k}$ shows that $(2k)^{-1/q}\sigma_q(E,F)$ is an increasing function of $q \in [1, \infty)$ if $E, F \in G_{n,k}$ and so the same is true about $\log N(G_{n,k}\,, \sigma_q, \varepsilon(2k)^{1/q})$. Therefore the upper bound in (13) is a consequence of (14).

To get a lower bound of the entropy, it is sufficient to look at only one chart, say $(V_{1/2}(E), \varphi_E)$ and for the nuclear norm. Using lemma 7 with $q = 1$, we reduce the problem to a minoration, for the nuclear metric, of the entropy of the unit ball of $L(E, E^\perp)$ with the operator norm. Now we join the method of [S2]; a lower bound for the covering number is obtained by evaluating the ratio of the volume of the operator norm unit ball of $L(E, E^\perp)$ and the volume of the nuclear norm unit ball. This concludes the proof. □

References

[C] S. Chevet, *Séries de variables aléatoires Gaussiennes à valeurs dans $E \otimes_\varepsilon F$. Application aux produits d'espaces de Wiener,* Séminaire Maurey-Schwartz, exposé XIX (1977–1978).

[G1] E. D. Gluskin, *The diameter of the Minkowski compactum is roughly equal to n,* Functional Anal. Appl. **15** (1981), 72–73.

[G2] E. D. Gluskin, *Finite-dimensional analogues of spaces without basis,* Dokl. Akad. Nauk. SSSR **261** (1981), 1046–1050.

[MP] M. Marcus and G. Pisier, *Random Fourier series with applications to harmonic analysis,* Annals of Math. Studies **101** (1981), Princeton Univ. Press.

[MS] V. Milman and G. Schechtmann, *Asymptotic theory of finite dimensional normed spaces,* Springer Lectures Notes **1200** (1986).

[P] G. Pisier, *The volume of convex bodies and Banach space geometry*, Cambridge University Press, 1989.

[PT] A. Pajor and N. Tomczak-Jaegermann, *Subspaces of small codimension of finite-dimensional Banach spaces*, Proc. Amer. Math. Soc. **97** (1986), 637–642.

[S1] S. Szarek, *Metric entropy of homogeneous spaces*, Preprint.

[S2] S. Szarek, *Nets of Grassmann manifold and orthogonal groups*, Proceedings of Banach Spaces Workshop, University of Iowa Press 1982, 169–185.

[S3] S. Szarek, *The finite dimensional basis problem with an appendix on nets of Grassmann manifold*, Acta Math. **151** (1983), 153–179.

Alain Pajor
Equipe d'Analyse et Mathématiques Appliquées
Université de Marne la Vallée
5 boulevard Descartes, Champs sur Marne
77454 Marne la Vallée Cedex 2
France
pajor@math.univ-mlv.fr