C. R. Acad. Sci. Paris, t. 315, Série I, p. 845-850, 1992

Probabilités/*Probability Theory*
(Analyse fonctionnelle/*Functional Analysis*)

Metric entropy and the small ball problem for Gaussian measures

James KUELBS and Wenbo V. LI

Abstract – We establish a precise link between the small ball problem for a Gaussian measure μ on a separable Banach space, and the metric entropy of the unit ball of the Hilbert space H_μ generating μ. This link allows us to compute small ball probabilities from metric entropy results, and vice versa.

L'entropie métrique et le problème de la petite boule pour une mesure gaussienne

Résumé – *Nous établissons une relation précise entre le problème de la petite boule pour une mesure gaussienne dans un espace de Banach séparable, et l'entropie métrique dans la boule unité de l'espace de Hilbert H_μ qui engendre μ. Cette relation nous permet de calculer des probabilités pour le problème de la petite boule à partir des résultats connus pour l'entropie métrique, et vice versa.*

Version française abrégée – Soit μ une mesure gaussienne centrée dans un espace de Banach réel, séparable avec norme $\|.\|$. On suppose aussi que K désigne la boule unité dans l'espace de Hilbert H_μ qui engendre μ. Alors, K est compact dans B avec entropie métrique $H(\varepsilon, K)$ finie. Pour K compact, rappelons que $H(\varepsilon, K)=\log N(\varepsilon, K)$ où

$$N(\varepsilon, K)=\min\left\{n\geqq 1:\exists k_1, \ldots, k_n\in K \text{ telles que } \bigcup_{j=1}^{n} B_\varepsilon(k_j)\supseteqq K\right\},$$

et $B_\varepsilon(k)=\{x:\|x-k\|<\varepsilon\}$. La notation $f(x)\approx g(x)$ lorsque $x\to a$ signifie que :

$$0<\varliminf_{x\to a} f(x)/g(x)\leqq\varlimsup_{x\to a} f(x)/g(x)<\infty,$$

et $f(x)\ll g(x)$ lorsque $x\to a$ signifie que :

$$\varlimsup_{x\to a} f(x)/g(x)<\infty.$$

Nous énonçons les théorèmes suivants :

THÉORÈME 1. – *Soit* $\log\mu(B_\varepsilon(0))=-\phi(\varepsilon)$, *et supposons* $f(1/x)$ *à variation régulière à l'infini, avec des constantes strictement positives* c_1, c_2 *telles que* $c_1 f(\varepsilon)\leqq\phi(\varepsilon)\leqq c_2 f(\varepsilon)$ *pour* $\varepsilon>0$, *suffisamment petit. Si* $j(\varepsilon)=\varepsilon(4c_2 f(\varepsilon))^{-1/2}$ *et* $g(j(\varepsilon))\approx\varepsilon$ *lorsque* $\varepsilon\to 0$, *alors on a*

$$H(\varepsilon, K)\approx f(g(\varepsilon)) \quad \textit{lorsque} \quad \varepsilon\to 0.$$

THÉORÈME 2. – *Soit* $\phi(\varepsilon)$ *comme dans le théorème* 1 *avec* $\phi(\varepsilon)\ll\phi(2\varepsilon)$ *lorsque* $\varepsilon\to 0$, *et supposons que* $g(1/x)$ *est à variation régulière à l'infini avec*

$$H(\varepsilon, K)\approx g(\varepsilon) \quad \textit{lorsque} \quad \varepsilon\to 0.$$

Alors, on a

$$\phi(\varepsilon)\approx g(\varepsilon/\phi(\varepsilon)^{1/2}) \quad \textit{lorsque} \quad \varepsilon\to 0.$$

En particulier, si $g(\varepsilon)=\varepsilon^{-\beta}J(1/\varepsilon)$, *où* $0<\beta<2$ *et* $J(x)$ *est à variation lente, monotone, et telle que* $J(x)\approx J(x^\rho)$ *lorsque* $x\to\infty$ *pour tout* $\rho>0$, *alors on a*

$$\phi(\varepsilon)\approx\varepsilon^{-2\beta/(2-\beta)}(J(1/\varepsilon))^{2/(2-\beta)} \quad \textit{lorsque} \quad \varepsilon\to 0.$$

Note présentée par Paul-André MEYER.

0764-4442/92/03150845 $ 2.00

THÉORÈME 3. — *Soit* $\mu=\mathfrak{L}(X)$, où $X=\sum_{k\geq 1}\lambda_k\xi_k e_k$ *est un vecteur gaussien à valeurs dans un espace de Hilbert* H *réel séparable,* $\lambda_k>0$ *est décroissante,* $\{\xi_k:k\geqq 1\}$ *est une suite de variables aléatoires indépendantes de loi* $N(0, 1)$, *et* $\{e_k:k\geqq 1\}$ *est une suite orthonormale dans* H. *Soit*

$$m(t)=\begin{cases}\sup_{k\geqq 1}\{k:\lambda_k\geqq t^{-1}\}, & t\geqq\lambda_1^{-1}\\ 0, & t<\lambda_1^{-1}\end{cases}$$

et, pour $t\geqq 0$,

$$I(t)=\int_0^t x^{-1}m(x)\,dx.$$

Soit $\phi(\varepsilon)$ *comme dans le théorème* 1 *avec* $\phi(\varepsilon)\ll(2\varepsilon)$ *lorsque* $\varepsilon\to\infty$. *Alors, si* $I(t)=t^{\beta}J(t)$ *où* $0<\beta<2$ *et* $J(.)$ *est à variation lente, monotone, et telle que* $J(x)\approx J(x^{\rho})$ *lorsque* $x\to\infty$ *pour tout* $\rho>0$, *nous avons*

$$\log P(\|X\|\leqq\varepsilon)\approx-\varepsilon^{-2\beta/(2-\beta)}(J(1/\varepsilon))^{2/(2-\beta)}\quad \textit{lorsque}\ \varepsilon\to 0.$$

---

1. INTRODUCTION. — Let μ denote a centered Gaussian measure on a real separable Banach space B with norm $\|.\|$ and dual B^*. If K is the unit ball of the Hilbert space H_μ which generates μ, then it is well known that $\lim_{t\to\infty} t^{-2}\log\mu(x:\|x\|\geqq t)=-(2\sigma^2)^{-1}$, where

$$\sigma^2=\sup_{\|f\|_{B^*}\leqq 1}\ \sup_{x\in K} f^2(x)=\sup_{\|f\|_{B^*}\leqq 1}\int_B f^2(x)\,d\mu(x),$$

and hence the distribution of the norm at infinity is, at the logarithmic level, a simple function of σ^2. The small ball problem studies this distribution near zero, namely, the behavior of $\log\mu(x:\|x\|\leqq\varepsilon)=-\phi(\varepsilon)$ as $\varepsilon\to 0$, and here the behavior of $\phi(\varepsilon)$ depends on much more than the single parameter σ^2. Indeed, the complexity of $\phi(\varepsilon)$ is well known, and there are only a few Gaussian measures for which $\phi(\varepsilon)$ has been determined completely as $\varepsilon\to 0$. The point of this paper is to link the behavior of $\phi(\varepsilon)$ to the metric entropy of K. Hence as a parallel to the large ball behavior being determined by the simple characteristic of K given by σ^2, the behavior of $\phi(\varepsilon)$ is governed by the more subtle metric entropy. This is a connection which is rather simple, but it links two delicate topics in a useful way. That is, once this link is obtained, then metric entropy results regarding K will yield information regarding $\phi(\varepsilon)$. Conversely, in instances when we know the behavior of $\phi(\varepsilon)$, we can establish some non-trivial and sometimes new results about the metric entropy of the various sets which appear as K. We include a sample of these applications in what follows, but our primary results are Theorems 1 and 2 below. We also mention the interesting partial result in [4] which was one of the starting points of this work.

If μ is a centered Gaussian measure on B, then it is well known that there is a unique Hilbert space $H_\mu\subseteq B$ such that μ is determined by considering the pair (B, H_μ) as an abstract Wiener space (*see* [5]). For example, if $B=C[0, 1]$ and μ is Wiener measure, then the unit ball of H_μ is

$$K=\left\{f(t)=\int_0^t f'(s)\,ds,\ 0\leqq t\leqq 1:\int_0^1 |f'(s)|^2\,ds\leqq 1\right\}. \tag{1.1}$$

Lemma 2.1 in [6] presents various properties of the relationship between H_μ and B, but the most important for us at this point is that the unit ball K of H_μ is always compact in the B-topology. Hence K has finite metric entropy.

To be precise we recall that if (E, *d*) is any metric space and A is a compact subset of (E, *d*), then the *d*-metric entropy of A is denoted by $H(\varepsilon, A)=\log N(\varepsilon, A)$ where

$$N(\varepsilon, A)=\min\left\{n\geqq 1:\exists\, a_1, \ldots, a_n\in A \text{ such that } \bigcup_{j=1}^{n} B_\varepsilon(a_j)\supseteqq A\right\},$$

and $B_\varepsilon(a)=\{x:d(x, a)<\varepsilon\}$ is the open ball of radius ε centered at *a*.

To state our results we use the notation $f(x)\approx g(x)$ as $x\to a$ if

$$0<\varliminf_{x\to a} f(x)/g(x)\leqq \varlimsup_{x\to a} f(x)/g(x)<\infty.$$

and write $f(x)\ll g(x)$ *as* $x\to a$ *if* $\varlimsup_{x\to a} f\infty/g(x)<\infty$.

2. Theorems.

Theorem 1. — *Let* μ *be a centered Gaussian measure on a real separable Banach space* B *and let*

$$\log \mu(B_\varepsilon(0))=-\phi(\varepsilon) \tag{2.1}$$

where $B_\varepsilon(0)=\{x\in B:\|x\|<\varepsilon\}$. *Let* K *denote the unit ball of the Hilbert space* H_μ *generating* μ. *If* $f(1/x)$ *is regularly varying at infinity with strictly positive finite constants* c_1, c_2 *such that*

$$c_1 f(\varepsilon)\leqq\phi(\varepsilon)\leqq c_2 f(\varepsilon) \tag{2.2}$$

for $\varepsilon>0$ *small and*

$$j(\varepsilon)=\varepsilon(4c_2 f(\varepsilon))^{-1/2}, \tag{2.3}$$

then

$$H(\varepsilon, K)\approx f(g(\varepsilon)) \quad \text{as } \varepsilon\to 0 \tag{2.4}$$

provided

$$g(j(\varepsilon))\approx\varepsilon \quad \text{as } \varepsilon\to 0. \tag{2.5}$$

Remarks. — I. The most prevalent form for $f(\varepsilon)$ is

$$f(\varepsilon)=\varepsilon^{-\alpha}(\log 1/\varepsilon)^{\beta} \tag{2.6}$$

where $\alpha\geqq 0$ and $\beta\in(-\infty, +\infty)$, and hence as $\varepsilon\to 0$ we have

$$H(\varepsilon, K)\approx\varepsilon^{-2\alpha/(2+\alpha)}(\log 1/\varepsilon)^{2\beta/(2+\alpha)}. \tag{2.7}$$

When $\beta=0$ in (2.6), a one sided estimate of (2.7) was obtained in [4].

II. Perhaps it should be pointed out that the function $f(\varepsilon)$ is used in Theorem 1 because it is rare that $\phi(\varepsilon)$ is known precisely. Furthermore, if only the upper (lower) bound in (2.2) is known, then the upper (lower) bound result in (2.4) also follows from the proof of the theorem. Now we turn to the converse result.

Theorem 2. — *Let* μ *be a centered Gaussian measure on a real separable Banach space* B, $\phi(\varepsilon)$ *be as in* (2.1) *with* $\phi(\varepsilon)\ll\phi(2\varepsilon)$ *as* $\varepsilon\to 0$, *and let* K *be the unit ball of the*

Hilbert space H_μ *generating* μ. *If* $g(1/x)$ *is regularly varying at infinity and*

(2.8) $$H(\varepsilon, K) \approx g(\varepsilon) \quad \text{as } \varepsilon \to 0,$$

then

$$\phi(\varepsilon) \approx g(\varepsilon/\phi(\varepsilon)^{1/2}) \quad \text{as } \varepsilon \to 0.$$

In particular, if $g(\varepsilon)=\varepsilon^{-\beta} J(1/\varepsilon)$ *where* $0<\beta<2$ *and* $J(x)$ *is slowly varying, monotonic and such that* $J(x) \approx J(x^\rho)$ *as* $x \to \infty$ *for each* $\rho>0$, *then*

$$\phi(\varepsilon) \approx \varepsilon^{-2\beta/(2-\beta)} (J(1/\varepsilon))^{2/(2-\beta)} \quad \text{as } \varepsilon \to 0.$$

Remark. — I. The restriction on β in Theorem 2 is natural since it is known (*see* [4]) that $H(\varepsilon, K)=o(\varepsilon^{-2})$ regardless of the Gaussian measure μ (and hence the subsequent K). When $\beta=2$, these idea also apply, but their application is much more delicate.

II. Putting Theorem 1 and 2 together, it is easy to see that $\phi(\varepsilon) \approx \varepsilon^{-\alpha}$ $(\alpha>0)$ iff $H(\varepsilon, K) \approx \varepsilon^{-2\alpha/(2+\alpha)}$ and $\phi(\varepsilon) \ll \phi(2\varepsilon)$ as $\varepsilon \to 0$.

III. The remark (II) following Theorem 1 has a complete analogue for Theorem 2.

The proofs of Theorems 1 and 2 depend on a number of important results for Gaussian measures including the Cameron-Martin translation formula and Borell's inequality. It is also possible to prove a random version of Theorem 1. This result gives the possibility of computer simulation for the metric entropy of the sets which appear as the unit ball of the H_μ. Since K is usually infinite dimensional this may not be terribly practical, but our random version of these results is suprisingly sharp. The proofs and further applications will appear in [7].

3. Applications. — Our first result shows how known metric entropy results can provide estimates for small ball probabilities.

Theorem 3. — *Let* $\mu=\mathfrak{L}(X)$ *where* $X=\sum_{k\geq 1} \lambda_k \xi_k e_k$ *is a centered Gaussian vector with values in a real separable Hilbert space* H, $\lambda_k>0$ *is non-increasing,* $\{\xi_k : k \geqq 1\}$ *is a sequence of independent* N(0, 1) *random variables, and* $\{e_k : k \geqq 1\}$ *is an orthonormal sequence in* H. *Let*

$$m(t)=\begin{cases} \sup_{k\geqq 1} \{k : \lambda_k \geqq t^{-1}\}, & t \geqq \lambda_1^{-1} \\ 0, & t<\lambda_1^{-1} \end{cases}$$

and define for $t \geqq 0$

$$I(t)=\int_0^t x^{-1} m(x)\, dx.$$

If $I(t)=t^\beta J(t)$ *where* $0<\beta<2$ *and* $J(.)$ *is slowly varying, monotonic, and such that* $J(x) \approx J(x^\rho)$ *as* $x \to \infty$ *for each* $\rho>0$, *then* $\phi(\varepsilon) \ll \phi(2\varepsilon)$ *as* $\varepsilon \to 0$ *implies*

$$\log P(\|X\| \leqq \varepsilon) \approx -\varepsilon^{-2\beta/(2-\beta)} (J(1/\varepsilon))^{2/(2-\beta)} \quad \text{as } \varepsilon \to 0.$$

Remarks. — If $\lambda_k=k^{-\alpha}(\log k)^\beta$ with $\alpha>1/2$ and β a real number, then as $t \to \infty$

$$m(t) \approx t^{1/\alpha} (\log t)^{\beta/\alpha}.$$

Hence in these situations,

(3.1) $$\log P(\|X\| \leqq \varepsilon) \approx -\varepsilon^{-2(2\alpha-1)} (\log(1/\varepsilon))^{2\beta/(2\alpha-1)}$$

provided $\phi(\varepsilon) \ll \phi(2\varepsilon)$ as $\varepsilon \to 0$.

The small ball probabilities calculated previously do not usually involve the logarithmic factors as in (3.1), as they make the estimates which were obtained via a detailed analysis of the Laplace transform of $\|X\|$ rather delicate (*see* [9] for further references). The method applied here is much simpler provided $\phi(\varepsilon) \ll \phi(2\varepsilon)$ as $\varepsilon \to 0$, and holds because in this case

$$K = \{ x \in H : x = \sum_{k \geq 1} (x, e_k) e_k, \sum_{k \geq 1} (x, e_k)^2 / \lambda_k^2 \leqq 1 \},$$

and hence the metric entropy of K can be determined by the function $I(t)$, *see* [11]. The result then follows by an easy application of Theorem 2. On the other hand, sometimes $\phi(\varepsilon)$ is known very precisely, and then there are correspondingly precise estimates of $H(\varepsilon, K)$ which are better than what is in the literature.

If $X = \sum_{k \geq 1} \lambda_k \xi_k e_k$ where $\{ e_k : k \geqq 1 \}$ is the canonical basis in the l^p spaces, $1 \leqq p < \infty$, then $P(X \in l^p) = 1$ iff $\sum_{k \geq 1} |\lambda_k|^p < \infty$ and

$$K = \{ x \in l^p : \sum_{k \geq 1} x_k^2 / \lambda_k^2 \leqq 1 \}.$$

If $P(X \in l^p) = 1$, then K is compact in l^p. For $p \neq 2$, the metric entropy of K in the l^p-norm is not so trivial to compute. The basic reason for this is that the volumes of finite dimensional projections of K do not compare well with the volumes of the same finite dimensional projection of the unit ball of l^p when $p \neq 2$. However, when $1 \leqq p < \infty$ and $\lambda_k = k^{-\alpha/p}$ for $\alpha > 1$ then [10] yields

$$\log P\Big(\sum_{k \geq 1} |\lambda_k|^p |\xi_k|^p)^{1/p} \leqq \varepsilon \Big) \approx -\varepsilon^{-p/(\alpha - 1)}.$$

Hence the corresponding ellipsoids

$$K = \{ x \in l^p : \sum_{k \geq 1} k^{2\alpha/p} x_k^2 \leqq 1 \}$$

have metric entropy in the l^p norm

$$H(\varepsilon, K) \approx \varepsilon^{-2p/(2\alpha + p - 2)}.$$

Another interesting class of examples arises when μ denotes Wiener measure on $C[0, 1]$. In this case, K is given in (1.1) and is a compact subset of $C[0, 1]$ for any of the norms

$$\| f \|_p = \begin{cases} \left(\int_0^1 | f(s) |^p \, ds \right)^{1/p}, & 1 \leqq p < \infty \\ \sup_{0 \leqq s \leqq 1} | f(s) |, & p = \infty. \end{cases}$$

For Wiener measure and $1 \leqq p \leqq \infty$, it is known that

$$\log \mu(\| x \|_p \leqq \varepsilon) \approx -\varepsilon^{-2} \tag{3.2}$$

and it is also known that

$$H(\varepsilon, K, \| \cdot \|_p) \approx \varepsilon^{-1}. \tag{3.3}$$

In fact, more than (3.2) is known for the small ball probabilities, but for $1 \leqq p < 2$ these results have only been obtained recently in [1] and are quite delicate. On the other side, the metric entropy results in (3.3) were obtained in [2] for $p = \infty$, and in [3] for $p = 1$. The remaining cases are then obvious. Of course, in view of our results, (3.2) and (3.3) are in complete duality. Furthermore, since in this case the logarithm of the small ball probabilities are known asymptotically, especially for $p=2$ and $p=\infty$, and hence we then have correspondingly better estimates for (3.3). For example, we have as $\varepsilon \to 0$

$$(2-\sqrt{3})/4 \leqq \varepsilon . H(\varepsilon, K, \|\cdot\|_2) \leqq 1 \qquad \text{and} \qquad (2-\sqrt{3})\,\pi/4 \leqq \varepsilon . H(\varepsilon, X, \|\cdot\|_\infty) \leqq \pi.$$

For $p=2$, the estimates are better than those in Theorem XVI of [8], and for $p=\infty$, there are no constant bounds in [2].

Both authors were supported in part by N.S.F. grant number DMS-9024961.

Note remise le 16 juillet 1992, acceptée le 10 août 1992.

REFERENCES

[1] A. A. BOROVKOV and A. A. MOGULSKII, On probabilities of small deviations for stochastic processes, *Siberian Adv. in Math.*, 1, 1991, pp. 39-63.

[2] M. S. BIRMAN and M. Z. SOLOMJAK, Piecewise polynomial approximation of the functions of the classes W_p^α, *Mat. Sb.*, 73, 1967, p. 331-355; English transl. in *Math. U.S.S.R.-Sb.*, 2, 1967, pp. 295-317.

[3] G. F. CLEMENTS, Entropies of several sets of real valued functions, *Pacific J. Math.*, 13, 1963, pp. 1085-1095.

[4] V. GOODMAN Some probability and entropy estimates for Gaussian measures, in *Proceedings of the Sixth International Conference on Probability in Banach Spaces*, Progr. Probab. Statist., 20, Birkhauser, Berlin, 1990.

[5] L. GROSS, *Lectures in modern analysis and applications II*, Lecture Notes in Math., 140, Springer, Berlin, 1970.

[6] J. KUELBS, A strong convergence theorem for Banach space valued random variables, *Ann. Probab.*, 4, 1976, pp. 744-771.

[7]. J. KUELBS and W. V. LI, *Metric entropy and the small ball problem for Gaussian measure* (in preparation)

[8] A. N. KOLMOGOROV and V. M. TIHOMIROV, ε-entropy and ε-capacity of sets in function spaces, *Uspekhi Mat. Nauk.*, 14, 1959, pp. 3-86; English transl. in *Amer. Math. Soc. Transl.*, 17, 1961, pp. 277-364.

[9] W. V. LI, Comparison results for the lower tail of Gaussian seminorms, *J. Theory Probab.*, 5, 1992, pp. 1-31.

[10] W. V. LI, On the lower tail of Gaussian measures on l^p, in *Proceedings of the Eighth International Conference on Probability in Banach Spaces*, Progr. Probab. Statist., Birkhauser, Berlin (to appear).

[11] B. S. MITYAGIN, The approximate dimension and bases in nuclear spaces, *Uspekhi Mat. Nauk.*, 16, 1961, pp. 63-132; English transl. in *Russian Math. Surveys*, 16, 1961, pp. 59-127

J. K. : *Department of Mathematics, University of Wisconsin-Madison, Madison, WI* 53706, *U.S.A.*,

W. V. L. : *Department of Mathematics, University of Wisconsin-Madison, Madison, WI* 53706, *U.S.A.*
and *Department of Mathematical Sciences, University of Delaware, Newark, DE* 19716, *U.S.A.*