

## MAJORIZING MEASURES WITHOUT MEASURES

BY MICHEL TALAGRAND

URA 754 AU CNRS

We give a reformulation of majorizing measures that does not involve measures, but rather special sequences of partitions. This formulation is more convenient to perform chaining with different distances.

**1. Introduction.** The theory of majorizing measures seems the appropriate tool to estimate  $E \sup_{t \in T} X_t$ , where  $(X_t)_{t \in T}$  is a stochastic process (i.e., a family of r.v.'s indexed by a set  $T$ ) such that one has some control over the tails of the increments  $X_s - X_t$  for  $s, t \in T$ . A typical such control would be of the form

$$(1.1) \quad \forall u \geq 0, \quad P(|X_s - X_t| \geq u) \leq 2 \exp\left(-\frac{u^2}{d(s, t)^2}\right),$$

where  $d$  is a distance on  $T$  (“sub-Gaussian processes”). The theory of majorizing measures is a method to measure the “size” of a metric space  $(T, d)$  in a manner appropriate to take full advantage of (1.1). This theory is explained in detail in [8]. The case (1.1) is very important, because of Gaussian processes. It is unfortunately rather exceptional that one can control the left-hand side of (1.1) in a sharp way as a function of one single distance on  $T$ . There are many cases (e.g., infinitely divisible processes) where this is best done using families of distances. The theory of majorizing measures has been extended to this setting. There does not exist a unified exposition of these results, which are scattered among a number of papers ([5, 6] and [7]). This is unfortunate, because this extension of the theory of majorizing measures to the case of families of distances requires new ideas. This paper is motivated by the observation that one of these ideas can be expressed in the familiar “one distance” setting and that this seems to have at least pedagogical interest. The present paper is therefore a complement to the expository paper [8].

Before we state our results, we recall the traditional definition of majorizing measures. Given  $\alpha > 0$  and a metric space  $(T, d)$ , we define

$$(1.2) \quad \gamma_\alpha(T) = \gamma_\alpha(T, d) = \inf_{\mu} \sup_{t \in T} \int_0^\infty \left( \log \frac{1}{\mu(B(t, \varepsilon))} \right)^{1/\alpha} d\varepsilon,$$

where  $B(t, \varepsilon)$  is the closed ball of center  $t$  and radius  $\varepsilon$  and where the infimum is taken over all the (discrete) probability measures on  $T$ . Several pages of [8] attempt to take some of the mystery out of this definition and to explain the meaning of this mysterious probability  $\mu$ . The formulation we will give here offers the advantage of dispensing with  $\mu$ .

---

Received January 2000; revised May 2000.

AMS 2000 subject classifications. 60G15, 60G17, 60G60.

Key words and phrases. chaining, increment condition, bounded sample path.

We define

$$(1.3) \quad \gamma'_\alpha(T) = \inf \sup_{t \in T} \sum_{k \geq 0} d(t, C_k) 2^{k/\alpha},$$

where the infimum is taken over all choices of finite subsets  $(C_k)_{k \geq 0}$  of  $T$  with  $\text{card } C_0 = 1$ ,  $\text{card } C_k \leq 2^{2^k}$  for  $k \geq 1$ , and where, of course,  $d(t, C_k)$  is the distance from  $t$  to  $C_k$ .

**THEOREM 1.1.** *There exists a constant  $K(\alpha)$  depending only on  $\alpha$  such that*

$$K(\alpha)^{-1} \gamma'_\alpha(T) \leq \gamma_\alpha(T) \leq K(\alpha) \gamma'_\alpha(T).$$

There are two rather distinct aspects to the theory of majorizing measures. One is, knowing the existence of a majorizing measure, to control processes such as in (1.1). We will show that the formulation (1.3) is well adapted to that purpose. The second topic, which is typically much harder, concerns construction of majorizing measures. Tools have been developed to that purpose, but the bottom line remains that one has to “guess” what the right choice of  $\mu$  is. With the formulation (1.3), this amounts to guessing what sets one should take for  $C_k$ . It is conceivable that, at least in certain situations, it is easier to guess these sets rather than to guess  $\mu$ . The present paper was motivated by discussions with K. Ball about the case of the simplex, which is as follows. Consider the basis  $\mathbf{e}_1, \dots, \mathbf{e}_n$  or  $\mathbb{R}^n$ , provided with the Euclidean distance, and denote by  $T$  the convex hull of  $\mathbf{e}_1, \dots, \mathbf{e}_n$ . It is known from the theory of Gaussian processes that  $\gamma_2(T)$  is of order  $\sqrt{\log n + 1}$ , and it is an interesting exercise to prove this directly. A solution to this exercise is to be found in [8]. It is another interesting exercise (left here as a challenge to the motivated reader) to define sets  $C_k$  that show that  $\gamma'_2(T)$  is at most of order  $\sqrt{\log n + 1}$ .

Another important theme of majorizing measures is that of “ultrametricity.” For a metric space  $(T, d)$ , consider an increasing sequence of finite partitions  $(\mathcal{A}_k)_{k \geq 0}$  of  $T$ . We assume

$$(1.4) \quad \text{card } \mathcal{A}_0 = 1, \quad \text{card } \mathcal{A}_k \leq 2^{2^k}.$$

For a point  $t$  in  $T$ , we denote by  $A_k(t)$  the unique element of  $\mathcal{A}_k$  that contains  $t$ . We denote by  $D(A)$  the diameter of a set  $A$ . We set

$$(1.5) \quad \gamma''_\alpha(T) = \inf \sup_{t \in T} \sum_{k \geq 0} 2^{k/\alpha} D(A_k(t)).$$

The infimum is taken over all choices of the sequence  $(\mathcal{A}_k)_{k \geq 0}$ . It should be obvious that  $\gamma''_\alpha(T) \geq \gamma'_\alpha(T)$ , and the next result shows that, in fact, these two quantities are of the same order.

**THEOREM 1.2.** *There exists a constant  $K(\alpha)$  depending on  $\alpha$  only such that*

$$(1.6) \quad K(\alpha)^{-1} \gamma''_\alpha(T) \leq \gamma_\alpha(T) \leq K(\alpha) \gamma''_\alpha(T).$$

To illustrate the fact that the previous formulation is well adapted to the proof of upper bounds, we will prove the following.

**THEOREM 1.3.** *Consider a set  $T$  provided with two distances  $d_1, d_2$  and a process  $(X_t)_{t \in T}$  such that, for each  $u > 0$ .*

$$(1.7) \quad P(|X_s - X_t| \geq u) \leq 2 \exp\left(-\min\left(\frac{u}{d_1(s, t)}, \frac{u^2}{d_2^2(s, t)}\right)\right).$$

Then

$$E \sup_{s, t \in T} |X_t - X_s| \leq C(\gamma_1(T, d_1) + \gamma_2(T, d_2)),$$

where  $C$  is universal.

There are all kinds of possible variations (different powers, etc.) with the same proof. While proofs of upper bounds are overall easy, there seems to be a genuine difficulty in the proof of Theorem 1.3. In fact, the proof given in [1], page 327, contains a serious error. This gap was apparently first discovered by the referee of [2] (as explained there). In the interval between the publication of [1] and this observation, a correct proof was given in [5] and [7] using the framework of the theory of majorizing measures “with a family of distances.”

Even though Theorem 1.3 follows by combining the results of [5] and [7], the theory of majorizing measures for families of distances is a bit impressive at first sight, and the reader might enjoy the simple proof given here. Theorem 1.3 has been applied in [3] and [4].

**2. Proofs.** We denote by  $K$  a constant depending on  $\alpha$  only, not necessarily the same at each occurrence.

**LEMMA 2.1.** *We have  $\gamma_\alpha(T) \leq K\gamma'_\alpha(T)$ .*

**PROOF.** Consider sets  $C_k$  with  $\text{card } C_0 = 1$ ,  $\text{card } C_k \leq 2^{2^k}$  such that

$$(2.1) \quad \forall t \in T, \quad \sum_{k \geq 0} d(t, C_k) 2^{k/\alpha} \leq 2\gamma'_\alpha(T).$$

Consider a probability measure  $\mu$  on  $T$  such that

$$(2.2) \quad \forall x \in C_k, \quad \mu(\{x\}) \geq 2^{-k-1} 2^{-2^k}.$$

We will prove that

$$(2.3) \quad \forall t \in T, \quad \int_0^\infty \left( \log \frac{1}{\mu(B(t, \varepsilon))} \right)^{1/\alpha} d\varepsilon \leq K\gamma'_\alpha(T)$$

and this will prove the lemma.

We fix  $t$ , and we set

$$d_k = \min_{l \leq k} d(t, C_l).$$

Thus, for  $k \geq 0$ ,

$$\varepsilon \geq d_k \Rightarrow \mu(B(t, \varepsilon)) \geq 2^{-k-1} 2^{-2^k} \geq 2^{-2^{k+1}}$$

so, for  $k \geq 1$ ,

$$(2.4) \quad \int_{d_k}^{d_{k-1}} \left( \log \frac{1}{\mu(B(t, \varepsilon))} \right)^{1/\alpha} d\varepsilon \leq K d_{k-1} 2^{k/\alpha}.$$

Since, obviously,  $\mu(B(t, \varepsilon)) = 1$  for  $\varepsilon \geq D(T)$  [where  $D(T)$  is the diameter of  $T$ ], we have

$$(2.5) \quad \int_{d_0}^{\infty} \left( \log \frac{1}{\mu(B(t, \varepsilon))} \right)^{1/\alpha} d\varepsilon \leq KD(T).$$

Now  $C_0$  consists of one single point, and, by (2.2),

$$\forall t \in T, \quad d(t, C_0) \leq 2\gamma'_\alpha(T)$$

so that

$$D(T) \leq 4\gamma'_\alpha(T).$$

Summation of the relations (2.4) and (2.5) yields the result.  $\square$

Since, obviously,  $\gamma''_\alpha \geq \gamma'_\alpha$ , to prove Theorem 1.2, it suffices to prove the following.

**THEOREM 2.2.** *We have  $\gamma''_\alpha(T) \leq K\gamma_\alpha(T)$ .*

**PROOF.** We fix a probability measure  $\mu$  on  $T$  such that

$$(2.6) \quad \forall t \in T, \quad \int_0^\infty \left( \log \frac{1}{\mu(B(t, \varepsilon))} \right)^{1/\alpha} d\varepsilon \leq 2\gamma_\alpha(T).$$

Given  $k \geq 1$ , we proceed to the basic construction. We set  $D = D(T)$ . For  $l \geq 0$ , we consider the set

$$(2.7) \quad C'_l = \{t \in T; \mu(B(t, 2^{-l}D)) \geq 2^{l+1}2^{-2^k}\}.$$

Observe that  $C'_0 = T$ . Consider

$$C_l = C'_l \setminus \bigcup_{m>l} C'_m.$$

The sets  $C_l$  are disjoint, and since  $C'_0 = T$ , form a partition of  $T$ . Moreover,

$$(2.8) \quad t \in C_l \Rightarrow \mu(B(t, 2^{-l-1}D)) \leq 2^{l+2}2^{-2^k}$$

because  $t \notin C'_{l+1}$ . Consider a subset  $N_l$  of  $C'_l$  that is maximal subject to the condition

$$s, t \in N_l, \quad s \neq t \Rightarrow d(s, t) > 2^{-l+1}D.$$

The balls  $B(t, 2^{-l}D)$ , for  $t \in N_l$ , are disjoint, so that, by (2.7),

$$(2.9) \quad \text{card}N_l \leq 2^{-l-1}2^{2^k}.$$

Moreover, the maximality of  $N_l$  implies that the balls  $B(t, 2^{-l+1}D)$ , for  $t \in N_l$ , cover  $C'_l$ . Thus, we can partition  $C_l$  in at most  $\text{card}N_l$  sets of diameter at most  $2^{-l+2}D$ .

We now partition  $T$  as follows. First, we partition  $T$  into the sets  $C_l$ . Then we partition each  $C_l$  as explained. The resulting partition  $\mathcal{B}_k$  contains at most

$$2^{2^k} \sum_{l \geq 0} 2^{-l-1} \leq 2^{2^k}$$

sets. To each  $t, k$ , we associate the index  $l(k, t)$ , defined by  $t \in C_{l(k, t)}$ . We have

$$(2.10) \quad D(B_k(t)) \leq 2^{-l(k, t)+2}D$$

and, from (2.8),

$$(2.11) \quad \mu(B(t, 2^{-l(k, t)-1}D)) \leq 2^{l(k, t)+2}2^{-2^k}.$$

We set  $\mathcal{A}_0 = \mathcal{A}_1 = \{T\}$ , and for  $k \geq 2$ , we consider the partition  $\mathcal{A}_k$  generated by  $\mathcal{B}_1, \dots, \mathcal{B}_{k-1}$ , so that  $\text{card}\mathcal{A}_k \leq 2^{2^k}$ , and the sequence  $(\mathcal{A}_k)$  increases.

Let us fix  $t \in T$ . We have, since  $A_k(t) \subset B_{k-1}(t)$ ,

$$(2.12) \quad \begin{aligned} \sum_{k \geq 0} D(A_k(t))2^{k/\alpha} &\leq KD + \sum_{k \geq 2} D(B_{k-1}(t))2^{k/\alpha} \\ &\leq KD \left( 1 + \sum_{k \geq 1} 2^{k/\alpha} 2^{-l(k, t)} \right) \end{aligned}$$

using (2.10). Consider the set

$$L = \{l \in \mathbb{N}; \exists k, l = l(k, t)\}$$

and for  $l \in L$ , define

$$k(l) = \max\{k; l = l(k, t)\}.$$

It should be obvious that

$$\sum_{k \geq 1} 2^{k/\alpha} 2^{-l(k, t)} \leq K \sum_{l \in L} 2^{-l} 2^{k(l)/\alpha}.$$

Now, for  $l \in L, k = k(l)$ , (2.11) reads as

$$\mu(B(t, 2^{-l-1}D)) \leq 2^{l+2}2^{-2^{k(l)}}$$

and thus

$$\begin{aligned}
 (2.13) \quad & \int_{2^{-l-2}D}^{2^{-l-1}D} \left( \log \frac{1}{\mu(B(t, \varepsilon))} \right)^{1/\alpha} d\varepsilon \\
 & \geq 2^{-l-2} D \log(\max(1, 2^{2^{k(l)}-l-2}))^{1/\alpha} \\
 & \geq \frac{1}{K} 2^{-l} D ((2^{k(l)} - l - 2)^+)^{1/\alpha} \\
 & \geq \frac{1}{K} 2^{-l} D (2^{k(l)/\alpha} - K(l+2))^{1/\alpha},
 \end{aligned}$$

using the fact that, for  $a, b > 0$ ,  $((a-b)^+)^{1/\alpha} \geq (1/K)\alpha^{1/\alpha} - Kb^{1/\alpha}$ .

Summing the inequalities (2.13) and combining with (2.6) and (2.12) gives

$$\sum_{k \geq 0} D(A_k(t)) 2^{k/\alpha} \leq K(D + \gamma_\alpha(T)) \leq K\gamma_\alpha(T).$$

Since  $t$  was arbitrary, the proof is complete.  $\square$

Theorem 1.1 is a consequence of Lemma 2.1 and Theorem 2.3.

PROOF OF THEOREM 1.3. For  $j = 1, 2$ , we consider an increasing sequence  $(\mathcal{A}_k^j)_{k \geq 0}$  of partitions of  $T$  such that  $\text{card } \mathcal{A}_0^j = 1$ ,  $\text{card } \mathcal{A}_k^j \leq 2^{2^k}$  and

$$\forall t \in T, \quad \sum_{k \geq 0} D_j(A_k^j(t)) 2^{k/j} \leq K\gamma_j(T),$$

where  $D_j(A)$  denotes the diameter of  $A$  for  $d_j$ . This is possible by Theorem 1.2. We consider the partition  $\mathcal{A}_k$  generated by  $\mathcal{A}_k^1$  and  $\mathcal{A}_k^2$ , so that  $\text{card } \mathcal{A}_0 = 1$ ,  $\text{card } \mathcal{A}_k \leq 2^{2^{k+1}}$  and

$$(2.14) \quad \forall t \in T, \quad \sum_{k \geq 0} D_1(A_k(t)) 2^k + D_2(A_k(t)) 2^{k/2} \leq K(\gamma_1(T) + \gamma_2(T)).$$

(All the difficulty of dealing with different distances is gone now that we have the correct normalization, and the rest is standard.) For  $k \geq 0$ , we consider an arbitrary set  $B_k$  that contains exactly one point in each set of  $\mathcal{A}_k$ .

Consider a parameter  $v \geq 1$ . For each  $k \geq 0$ ,  $x \in B_{k+1}$ ,  $y \in B_k$ , consider the event

$$\Omega(x, y, v) : |X_x - X_y| \geq v(2^k d_1(x, y) + 2^{k/2} d_2(x, y)).$$

Thus, by (1.7), we have

$$P(\Omega(x, y, v)) \leq 2 \exp(-v2^k).$$

Thus, if we denote by  $\Omega(v)$  the union of the events  $\Omega(x, y, v)$  over all choices of  $x, y$ , we have

$$(2.15) \quad \begin{aligned} P(\Omega(v)) &\leq \sum_{k \geq 0} 2^{2^{k+2}} \exp(-v2^k) \\ &\leq \exp\left(-\frac{v}{2}\right) \end{aligned}$$

for  $v$  large enough.

Next, if  $\Omega(v)$  does not occur, we bound  $\sup_t |X_t - X_{t_0}|$ , where  $B_0 = \{t_0\}$ . Given  $t$  in  $T$ , for  $k \geq 1$ , let  $\pi_k(t) = B_k \cap A_k(t)$ . Since  $\Omega(x, y, v)$  does not occur, we have, since  $\pi_k(t), \pi_{k+1}(t) \in A_k(t)$ ,

$$(2.16) \quad |X_{\pi_{k+1}(t)} - X_{\pi_k(t)}| \leq v(2^k D_1(A_k(t)) + 2^{k/2} D_2(A_k(t))).$$

Summation of the relations (2.16) for  $k \geq 0$ , together with (2.14), shows that

$$|X_t - X_{t_0}| \leq Kv(\gamma_1(T) + \gamma_2(T)).$$

Combining this with (2.15), we see that

$$P\left(\sup_t |X_t - X_{t_0}| \geq Kv(\gamma_1(T) + \gamma_2(T))\right) \leq \exp\left(-\frac{v}{2}\right),$$

which implies the result.  $\square$

The extension of Theorem 1.3 to situations involving several distances is immediate. Such an extension is given in [2].

## REFERENCES

- [1] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces*. Springer, Berlin.
- [2] MARCUS, M. (1998). A sufficient condition for the continuity of high order Gaussian chaos processes: high dimensional probability. *Progr. Probab.* **43** 263–276.
- [3] MARCUS, M. and ROSEN, J. (1996). Gaussian chaos and sample path properties of additive functionals of symmetric Markov processes. *Ann. Probab.* **24** 1130, 1177.
- [4] MARCUS, M. and TALAGRAND, M. (1998). Continuity conditions for a class of Gaussian chaos processes related to continuous additive functionals of Lévy processes. *J. Probab.* **11** 157–179.
- [5] TALAGRAND, M. (1993). Regularity of infinitely divisible processes. *Ann. Probab.* **21** 362–432.
- [6] TALAGRAND, M. (1994). Construction of majorizing measures, Bernoulli processes and cotype. *Geom. Funct. Anal.* **4** 660–717.
- [7] TALAGRAND, M. (1994). The supremum of certain canonical processes. *Amer. J. Math.* **116** 283–325.
- [8] TALAGRAND, M. (1996). Majorizing measures: the generic chaining. *Ann. Probab.* **24** 1049–1103.

EQUIPE D'ANALYSE  
URA 754 AU CNRS  
4 PLACE JUSSIEU  
BOITE 186  
75252 PARIS CEDEX 05  
FRANCE