C. R. Acad. Sci. Paris, Ser. I 356 (2018) 692–695

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Probability theory

Scaling and non-standard matching theorems

Mise à l'échelle et théorèmes d'appariement non-standard

Michel Talagrand

23, rue Louis-Pouey, 92800 Puteaux, France

ARTICLE INFO

Article history:
Received 11 April 2018
Accepted 13 April 2018

Presented by Gilles Pisier

ABSTRACT

Consider the standard Gaussian measure μ on $\mathbb{R}^2$. Consider independent r.v.s $(X_i)_{i\leq N}$ distributed according to μ, and an independent copy $(Y_i)_{i\leq N}$ of these r.v.s. We prove that, for some number C and N large, we have

$$\frac{(\log N)^2}{C} \leq \mathsf{E}\inf_{\pi}\sum_{i\leq N} d(X_i, Y_{\pi(i)})^2 \leq C(\log N)^2 \,, \tag{1}$$

where the infimum is over all permutations π of $\{1, \ldots, N\}$. The striking point of this result is the factor $(\log N)^2$. Indeed, if instead of μ we consider the uniform distribution on the unit square, it is well known that the proper factor is $\log N$. The upper bound was proved by Michel Ledoux (2017) [3].

RÉSUMÉ

Considérons une suite indépendente $(X_i)_{i\leq N}$ de variables aléatoires distribuées comme la mesure gaussienne canonique μ sur $\mathbb{R}^2$ et une copie independente $(Y_i)_{i\leq N}$ de cette même suite. Pour une certaine constante universelle C et $N \geq 2$, nous avons les inégalités

$$\frac{(\log N)^2}{C} \leq \mathsf{E}\inf_{\pi}\sum_{i\leq N} d(X_i, Y_{\pi(i)})^2 \leq C(\log N)^2 \tag{1}$$

où l'infimum est pris sur toutes les permutations π de $\{1, \ldots, N\}$. La borne supérieure a été prouvée par Michel Ledoux (2017) [3], qui conjecturait que l'inégalité (1) était correcte avec un facteur $\log N$ et non pas $(\log N)^2$. C'est précisement l'apparence de ce facteur $(\log N)^2$ qui est non standard.

E-mail address: michel.talagrand@gmail.com.

https://doi.org/10.1016/j.crma.2018.04.018

1. Introduction

Consider independent r.v.s $(X_i)_{i\le N}$ uniformly distributed on the unit square of $\mathbb{R}^2$ and an independent copy $(Y_i)_{i\le N}$ of these variables. It is well known that, for each $p>1$, one has

$$\frac{N^{1-p/2}(\log N)^{p/2}}{C_p} \le \mathsf{E}\inf_\pi \sum_{i\le N} d(X_i, Y_{\pi(i)})^p \le C_p N^{1-p/2}(\log N)^{p/2} \tag{2}$$

for some number C_p depending on p only. An important idea here is that the lower bound follows from Hölder's inequality and the case $p=1$, and that this lower bound also holds if the distribution of the X_i is not too different from uniform, say its density is constant within a multiplicative factor 2.

One may like, of course, to investigate what happens when the uniform distribution is replaced by an unbounded distribution μ. In the case $p=1$, this was done, in particular, by J. Yukich [5].

In this note, we bring forward a completely elementary scaling property that does not seem to have been previously noticed. Since this property is not specific to dimension 2, we will explain it in its proper setting where it is much clearer. Consider $k\ge 3$ and assume now that the r.v.s X_i and Y_i are uniform over $[0,1]^k$. Then one should replace (1) by the (much easier) inequality

$$\frac{N^{1-p/k}}{C_{k,p}} \le \mathsf{E}\inf_\pi \sum_{i\le N} d(X_i, Y_{\pi(i)})^p \le C_{k,p} N^{1-p/k}\,, \tag{3}$$

where $C_{k,p}$ is independent of N. It is the case $p=k$, which gives rise to interesting scaling effects. To explain the heuristic, we assume now that the X_i have a common distribution μ. Let us assume that in a certain box A of side a, the probability μ is nearly uniform, in a sense that its density ρ with respect to Lebesgue's measure varies on that square by at most a factor 2, say $b\le\rho\le 2b$. Assuming $n = Nba^k \ge N^{1/100}$, there are about n points X_i that belong to A. The typical distance between a point X_i and the closest point Y_j is about $a/n^{1/k}$. We then expect that, for any permutation π, we have

$$\sum_{X_i\in A} d(X_i, Y_{\pi(i)})^k \ge \frac{1}{C_k} n \times a^k/n = \frac{a^k}{C_k}\,. \tag{4}$$

The fundamental fact here is that this quantity is independent of b, so that in a sense the points in A do not overall contribute like $\mu(A)$, but rather like $\lambda_k(A)$, the k-dimensional volume of A. For this reason, we should heuristically have

$$\sum_{i\le N} d(X_i, Y_{\pi(i)})^k \ge \frac{U}{C_k} \tag{5}$$

where U is the area of the union of the squares A as described above. In the case where μ is the standard Gaussian measure on $\mathbb{R}^k$, it turns out that the union of such squares contains a sphere of radius $\sqrt{\log N}/C$, which is volume $(\log N)^{k/2}/C$, and we have completed the heuristic proof of the following.

Proposition 1.1. *Assume $k\ge 3$. If the sequence $(X_i)_{i\le N}$ is independently distributed according to the standard Gaussian measure μ and $(Y_i)_{i\le N}$ is an independent copy of this sequence, then*

$$\mathsf{E}\inf_\pi \sum_{i\le N} d(X_i, Y_{\pi(i)})^k \ge \frac{(\log N)^{k/2}}{C_k}\,. \tag{6}$$

It is extremely easy to turn the heuristic argument into a rigorous proof. There is every reason to believe that the bound (6) can be reversed, and we outline below a variation of the standard “transportation method” that should prove that, but we have not performed the computations to check that it works.

The situation is more subtle when $k=2$. There is another type of more global fluctuations in a random sample. These fluctuations create the extra factor $\log N$ in (2). These fluctuations are also present inside each box A, and created an extra factor $\log n \simeq \log N$. The lower bound in (5) has to be replaced by $(U\log N)/C$, and this is how one reaches the lower bound in (1).

In the next section, we make the previous ideas precise in order to prove the lower bound in (1). The upper bound in (1) was proved by Michel Ledoux [3] (by adapting the methods of [2]), who introduced the problem. We will give a completely elementary proof of this upper bound, using only (2).

It should be noted that our proofs depend heavily on the fact that the tails of the Gaussian distribution decrease fast. More precisely, calling $\rho(x)$ the density of this distribution with respect to Lebesgue's measure, the regions where $\rho(x)\ge 1/N$ and $\rho(x)\ge N^{-1/100}$ have areas of the same order. It seems certain that our result can be extended to the case of a distribution μ with density proportional to $\exp(-d(0,x)^\alpha)$ with respect to Lebesgue's measure (where $\alpha>0$) (replacing of course the factor $(\log N)^2$ by $(\log N)^{1+2/\alpha}$). But what happens is the case where μ has a density proportional to $(1+d(0,x))^{-\alpha}$ is far less clear.