



HAL
open science

Statistical minimax approach of the Hausdorff moment problem

Thanh Mai Pham Ngoc

► **To cite this version:**

Thanh Mai Pham Ngoc. Statistical minimax approach of the Hausdorff moment problem. 2007. hal-00145243v1

HAL Id: hal-00145243

<https://hal.science/hal-00145243v1>

Preprint submitted on 9 May 2007 (v1), last revised 22 Apr 2009 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Statistical minimax approach of the Hausdorff moment problem

Thanh Mai Pham Ngoc

Université Pierre et Marie Curie, Paris VI

Abstract: The purpose of this paper is to study the problem of estimating a compactly supported density of probability from noisy observations of its moments. In fact, we provide a statistical approach to the famous Hausdorff classical moment problem. We prove an upper bound and a lower bound on the rate of convergence of the mean squared error showing that the considered estimator attains minimax rate over the corresponding smoothness classes.

Key words and phrases: Minimax estimation, Problem of the moments, Legendre polynomials, Kullback-Leibler information, density estimation

MSC 2000 Subject Classification 62G07 44A60

1. Introduction

The classical moment problem can be stated as follows: it consists in getting some information about a distribution μ from the knowledge of its moments $\int x^k d\mu(x)$. It has been largely investigated in many mathematical topics, among others, in operator theory, mathematical physics, inverse spectral theory, probability theory, inverse problems, numerical analysis . . . We may cite the classical and pioneer books in the field (see Akhiezer (1965), Shohat and Tamarkin (1943)) which put emphasis on the existence aspect of the solution and its unicity. According to the support of the distribution of interest, one may refer to one of the three types of classical moment problems: the Hamburger moment problem whose support of μ is the whole real line, the Stieljes problem on $[0, +\infty)$ and finally the Hausdorff problem on a bounded interval. In this paper, we shall focus on the last issue.

The Hausdorff moment problem which dates back to 1921 (see Hausdorff (1921)) has been an object of great interest in the literature since then. For instance, the particular case when only a finite number of moments are known,

has aroused much attention in inverse problems (see Talenti (1987), Fasino and Inglese (1996), Inglese (1995), Tagliani (2002)). Moreover, more recently, Ang, Gorenflo, Le and Trong (2002) have presented the Hausdorff moment problem under the angle of ill posed problems, in a sense that solutions do not depend continuously on the given data. Nonetheless, until now, as far as we know, the statistical approach which consists in introducing some randomness in the noise has been very little put forward and rarely raised. However, in a slight different context, we can cite the work of Goldenshluger and Spokoiny (2004). In their paper, the authors tackled the problem of reconstructing a planar convex from noisy geometric moments observations.

We consider in this paper a statistical point of view of the Hausdorff moment problem. We aim at estimating an unknown probability density from noisy measurements of its moments on a symmetric bounded interval. Without loss of generality we may and will suppose that $[-a, a] = [-1, 1]$. The estimation procedure we use is based on the expansion of the unknown density through the basis of Legendre polynomials and an orthogonal series method. We establish an upper bound and a lower bound on the estimation accuracy of the procedure showing that it is optimal in a minimax sense. We show that the achieved rate is only of a logarithmic order. This fact has already been underlined by Goldenshluger and Spokoiny (2004). Indeed, they pointed out that in view of reconstructing a planar region, the upper bound was only in the order of logarithmic rate.

One might question this chronic slow rate which seems inherent to moment problems. In fact, the underlying problem lies in the non orthogonal nature of the monomials x^k . They actually hamper the convergence rate to be improved for bringing a small amount of information. This remark is highlighted in our proof of the upper bound. As for the proof of the lower bound, it confirms the severely ill-posed feature of the moment problem.

This paper is organized as follows: in section 2 we introduce the model and the estimator of the unknown probability density and we finally state the two theorems. Section 3 contains the proofs. The last section is an appendix in which we prove some useful inequalities about binomial coefficients.

2. Estimation from moments

Let μ_k be the moments of the unknown probability density f given by:

$$\mu_k = \int_{-1}^1 x^k f(x) dx \quad k = 0, 1, \dots,$$

One gets the following sequence of noisy observations:

$$y_k = \mu_k + \varepsilon \xi_k \quad k = 0, 1, \dots, \quad (4.1)$$

where $0 < \varepsilon < 1$ and ξ_k are i.i.d standard Gaussian random variables. The objective is to estimate the density f given these noisy moments observations. Being supported on the finite interval $[-1, 1]$, the measure having density f is unique (see Feller (1968, Chap.7)), so that the statistical problem of recovering f from (4.1) is relevant.

The smoothness of the probability density f is assessed in terms of its Legendre-Fourier coefficients. In fact, we assume that f belongs to the Sobolev ellipsoid \mathcal{F}_r of order $2r$ defined by:

$$\mathcal{F}_r = \{f \in L^2[-1, 1] : f \geq 0, \int_{-1}^1 f(x) dx = 1, \sum_k k^{2r} |\theta_k|^2 < \infty\}$$

where $\theta_k = \int_{-1}^1 f(x) P_k(x) dx$ and P_k denotes the normalized Legendre polynomial of degree k .

The use of the Legendre polynomials in the Hausdorff classical moment problem in order to approximate the unknown measure is quite natural (see Ang, Gorenflo, Le and Trong (2002), Bertero, De Mol and Pike (1985)) as they directly result from the Gram-Schmidt orthonormalization of the family $\{x^k\}$, $k = 0, 1, \dots$. Consequently, the procedure which consists in representing the density f to be estimated in the basis of Legendre polynomials appears naturally.

Let us define now the estimator of f . This latter is induced by an orthogonal series method through the Legendre polynomials.

Any function in $L^2[-1, 1]$ has an expansion:

$$f(x) = \sum_{k=0}^{\infty} \theta_k P_k(x) \quad \text{with} \quad \theta_k = \int_{-1}^1 f(x) P_k(x) dx.$$

The problem of estimating f reduces to estimation of the sequence $\{\theta_k\}_{k=1}^{+\infty}$ for Legendre polynomials form a complete orthogonal function system in $L^2[-1, 1]$.

Denote $\beta_{n,j}$ the coefficients of the Legendre polynomial of degree n :

$$P_n(x) = \sum_{j=0}^n \beta_{n,j} x^j$$

which yields

$$\theta_k = \int_{-1}^1 f(x) P_k(x) dx = \sum_{j=0}^k \beta_{k,j} \int_{-1}^1 f(x) x^j dx = \sum_{j=0}^k \beta_{k,j} \mu_j.$$

This leads us to consider the following estimator of θ :

$$\hat{\theta}_k = \sum_{j=0}^k \beta_{k,j} y_j$$

and hence the estimator \hat{f}_N of f :

$$\hat{f}_N(x) = \sum_{k=0}^N \sum_{j=0}^k \beta_{k,j} y_j P_k(x) \equiv \sum_{k=0}^N \hat{f}_{N,k}(x),$$

where y_j is given by (4.1) and N is an integer to be properly selected later.

The mean integrated square error of the estimator \hat{f}_N is:

$$\mathbb{E}_f \|\hat{f}_N - f\|^2,$$

where \mathbb{E}_f denotes the expectation w.r.t the distribution of the data in the model (4.1) and for a function $g \in L^2[-1, 1]$,

$$\|g\| = \left(\int_{-1}^1 g^2(x) dx \right)^{1/2}.$$

In this paper we shall consider the problem of estimating f using the mean integrated square risk in the model (4.1).

We state now the two results of the paper. The first theorem establishes an upper bound.

Theorem 1. For $\alpha > 0$, define the integer $N = \lfloor \alpha \log(1/\varepsilon) \rfloor$. Then we have

$$\sup_{f \in \mathcal{F}_r} \mathbb{E}_f \|\hat{f}_N - f\|^2 \leq C[\log(1/\varepsilon)]^{2r},$$

where C is an absolute positive constant and $\lfloor \cdot \rfloor$ denotes the floor function.

The second theorem provides a lower bound.

Theorem 2. We have

$$\inf_{\hat{f}} \sup_{f \in \mathcal{F}_r} \mathbb{E}_f \|\hat{f} - f\|^2 \geq c[\log(1/\varepsilon)]^{2r},$$

where c is a positive constant which depends only on r and the infimum is taken over all estimators \hat{f} .

3. Proofs

3.1 Proof of Theorem 1. By the usual MISE decomposition which involves a variance term and a bias term, we get

$$\mathbb{E}_f \|\hat{f}_N - f\|^2 = \mathbb{E}_f \sum_{k=0}^N (\hat{f}_{N,k} - \theta_k)^2 + \sum_{k \geq N+1} \theta_k^2$$

but

$$\begin{aligned} \mathbb{E}_f \sum_{k=0}^N (\hat{f}_{N,k} - \theta_k)^2 &= \mathbb{E}_f \sum_{k=0}^N \left(\sum_{j=0}^k \beta_{k,j} (y_j - \mu_j) \right)^2 \\ &= \varepsilon^2 \mathbb{E} \sum_{k=0}^N \left(\sum_{j=0}^k \beta_{k,j} \xi_j \right)^2 \end{aligned}$$

and since $\xi_j \stackrel{iid}{\sim} N(0, 1)$, it follows that

$$\begin{aligned} \mathbb{E}_f \|\hat{f}_N - f\|^2 &= \varepsilon^2 \sum_{k=0}^N \sum_{j=0}^k \beta_{k,j}^2 + \sum_{k \geq N+1} \theta_k^2 \\ &= V_N + B_N^2 \end{aligned}$$

We first deal with the variance term V_N . To this end, we have to upper bound the

sum of the squared coefficients of the normalized Legendre polynomial of degree k . Set $\sigma_k^2 = \sum_{j=0}^k \beta_{k,j}^2$. An explicit form of $P_k(x)$ is given by (see Abramowitz and Stegun (1970)):

$$P_k(x) = \left(\frac{2k+1}{2}\right)^{1/2} \frac{1}{2^k} \sum_{j=0}^{\lfloor k/2 \rfloor} (-1)^k \binom{k}{j} \binom{2k-2j}{k} x^{k-2j},$$

where $\lfloor \cdot \rfloor$ denotes the integer part and $\binom{k}{j}$ denotes the binomial coefficient, $\binom{k}{j} = \frac{k!}{(k-j)!j!}$. This involves

$$\begin{aligned} \sigma_k^2 &= \frac{2k+1}{2} \frac{1}{4^k} \sum_{j=0}^{\lfloor k/2 \rfloor} \left\{ \binom{k}{j} \binom{2k-2j}{k} \right\}^2 \\ &\leq \frac{2k+1}{2} \frac{1}{4^k} \left\{ \binom{2k}{k} \right\}^2 \sum_{j=0}^{\lfloor k/2 \rfloor} \left\{ \binom{k}{j} \right\}^2 \\ &\leq \frac{2k+1}{2} \frac{1}{4^k} \left\{ \binom{2k}{k} \right\}^2 (2^k)^2. \end{aligned}$$

By using now the fact that $\left\{ \binom{2k}{k} \right\}^2 \leq \frac{4^{2k}}{\sqrt{k}}$ (see lemma 3 from Appendix) we have

$$\sigma_k^2 \leq \frac{2k+1}{2} \frac{4^{2k}}{\sqrt{k}}$$

which yields

$$V_N \leq C\varepsilon^2 N^{1/2} 4^{2N},$$

where $C > 0$ denotes an absolute positive constant.

Now, it remains to upper bound the bias term B_N^2 . As the density f belongs to the Sobolev ellipsoid \mathcal{F}_r

$$B_N^2 \leq N^{-2r}.$$

Finally we get

$$\mathbb{E}_f \|\hat{f}_N - f\|^2 \leq C\varepsilon^2 N^{1/2} 4^{2N} + N^{-2r}$$

from which the desired result follows.

3.2 Proof of Theorem 2. In order to prove the lower bound of theorem 2, we first see that the model (4.1) is equivalent to an heteroscedastic model. Indeed, if

we multiply both sides of the model (4.1) by the coefficients $\beta_{k,j}$ of the Legendre polynomials we get the following model:

$$\tilde{y}_k = \theta_k + \varepsilon \sigma_k \xi_k \quad (3.1)$$

where $\sigma_k^2 = \sum_{j=0}^k \beta_{k,j}^2$, $\tilde{y}_k = \sum_{j=0}^k \beta_{k,j} y_j$, $\theta_k = \sum_{j=0}^k \beta_{k,j} \mu_j = \int_{-1}^1 f(x) P_k(x) dx$ and ξ_k are i.i.d standard Gaussian random variables.

In consequence, the model (3.1) is an heteroscedastic gaussian sequence space model through the basis of Legendre polynomials. From now on, we shall consider the model (3.1).

Before going any further, we can make a remark at this stage concerning the model (3.1). Indeed, this latter has clear similarities with standard ill-posed problems (see Cavalier, Golubev, Lepski and Tsybakov (2004), Golubev and Khasminskii (1999)). Depending on the asymptotic behavior of the intensity noise σ_n^2 one may characterize the nature of the problem. Here, in our case, $\sigma_n^2 \geq \frac{1}{4} 4^n$ (see lemma 4 from Appendix) and hence tends to infinity exponentially. We may say that we are dealing with a severely ill-posed problem with log-rates (see Cavalier, Golubev, Lepski and Tsybakov (2004)).

The proof of the lower bound essentially leans on the following particular version of Fano's lemma (see Birgé and Massart (2001)). It uses the Kullback-Leibler divergence $K(p_1, p_0)$ between two probability densities p_1 and p_0 defined by :

$$K(p_1, p_0) = \begin{cases} \int_{\mathbb{R}} \log\left(\frac{p_1(x)}{p_0(x)}\right) p_1(x) dx & \text{if } \mathbb{P}_1 \ll \mathbb{P}_0 \\ +\infty & \text{otherwise} \end{cases}$$

Lemma 1. *Let η be a strictly positive real number and \mathcal{C} be a finite set of probability densities $\{f_0, \dots, f_M\}$ on \mathbb{R} with $|\mathcal{C}| \geq 6$ such that :*

- (i) $\|f_i - f_j\| \geq \eta > 0$, $\forall 0 \leq i < j \leq M$.
- (ii) $\mathbb{P}_j \ll \mathbb{P}_0$, $\forall j = 1, \dots, M$, and

$$K(f_j, f_0) \leq H < \log M$$

then for any estimator \hat{f} and any nondecreasing function ℓ

$$\sup_{f \in \mathcal{C}} \mathbb{E}_f [\ell(\|\hat{f} - f\|)] \geq \ell\left(\frac{\eta}{2}\right) \left[1 - \left(\frac{2}{3} \vee \frac{H}{\log M}\right)\right].$$

We define \mathcal{E} as a set of functions of the following type

$$\mathcal{E} = \left\{ f_\delta \in \mathcal{F}_r : f_\delta = \frac{1}{2} \mathbf{1}_{[-1,1]} \left(1 + \frac{c_0}{m^{(4r+3)/2}} \sum_{k=m}^{2m-1} \delta_k k^{(2r+2)/2} P_k \right), \delta = (\delta_m, \dots, \delta_{2m-1}) \in \Delta = \{0, 1\}^m \right\}$$

We check that the functions f_δ defined above are probability densities. Clearly, since $\int_{-1}^1 P_k(x) dx = 0$, for any $\delta \in \{0, 1\}^m$, f_δ satisfies $\int_{-1}^1 f_\delta(x) dx = 1$. Furthermore, it is well known that $|P_k(x)| \leq 1$, $\forall x \in [-1, 1]$ (see Abramowitz and Stegun (1970)) so $f_\delta \geq 0$.

Secondly, we verify that f_δ belongs to \mathcal{F}_r . In this aim, we have to calculate the Legendre-Fourier coefficients associated with the density f_δ :

$$\begin{aligned} \theta_{\delta l} &= \int_{-1}^1 f_\delta(x) P_l(x) dx \\ &= \begin{cases} \frac{c_0}{m^{(4r+3)/2}} \cdot l^{(2r+2)/2} \cdot \delta_l & \text{if } l \in [m, 2m-1] \\ 0 & \text{else} \end{cases} \end{aligned} \quad (3.2)$$

hence

$$\begin{aligned} \sum_{k=0}^{+\infty} k^{2r} \theta_{\delta k}^2 &= \frac{c_0^2}{m^{4r+3}} \sum_{k=m}^{2m-1} k^{2r} k^{2r+2} \delta_k^2 \\ &\leq \frac{c_0^2}{m^{4r+3}} \sum_{k=m}^{2m-1} k^{4r+2} \delta_k^2 \\ &\leq \frac{c_0^2 (2m)^{4r+2}}{m^{4r+3}} \sum_{k=m}^{2m-1} \delta_k^2 \leq c_0^2 2^{4r+2} < \infty, \end{aligned}$$

since $\delta_k \in \{0, 1\}$.

We set $\delta^{(0)} = (0, \dots, 0)$ and $f_{\delta^{(0)}} \equiv f_0$. The Legendre-Fourier coefficients of f_0 are null:

$$\theta_{0l} = 0 \quad \forall l \in \mathbb{N}. \quad (3.3)$$

We now exhibit a suitable subset of densities \mathcal{C} . We only take into consideration a subset of $M+1$ densities of \mathcal{E} :

$$\mathcal{C} = \{f_{\delta^{(0)}}, \dots, f_{\delta^{(M)}}\}$$

where $\{\delta^{(1)}, \dots, \delta^{(M)}\}$ is a subset of $\{0, 1\}^m$.

We are now going to apply lemma 1. We first check the condition (i), accordingly, we have to assess the distance $\|f_{\delta^{(i)}} - f_{\delta^{(j)}}\|^2$. By the orthogonality of the system $\{P_k\}_k$ and thanks to Parseval equality we get, for $0 \leq i < j \leq M$,

$$\begin{aligned} \|f_{\delta^{(i)}} - f_{\delta^{(j)}}\|^2 &= \frac{c_0^2}{m^{4r+3}} \sum_{k=m}^{2m-1} k^{2r+2} (\delta_k^{(i)} - \delta_k^{(j)})^2 \\ &\geq \frac{c_0^2}{m^{4r+3}} \cdot m^{2r+2} \sum_{k=m}^{2m-1} (\delta_k^{(i)} - \delta_k^{(j)})^2 \\ &\geq \frac{c_0^2}{m^{2r+1}} \sum_{k=m}^{2m-1} (\delta_k^{(i)} - \delta_k^{(j)})^2 \\ &= \frac{c_0^2}{m^{2r+1}} \rho(\delta^{(i)}, \delta^{(j)}), \end{aligned}$$

where $\rho(\cdot, \cdot)$ is the *Hamming distance*. We are going to resort to the Varshamov-Gilbert bound which is stated in the following lemma:

Lemma 2. (Varshamov-Gilbert bound, 1962). *Fix $m \geq 8$. Then there exists a subset $\{\delta^{(0)}, \dots, \delta^{(M)}\}$ of Δ such that $M \geq 2^{m/8}$ and*

$$\rho(\delta^{(j)}, \delta^{(k)}) \geq \frac{m}{8}, \quad \forall 0 \leq j < k \leq M.$$

Moreover we can always take $\delta^{(0)} = (0, \dots, 0)$.

For a proof of this lemma see for instance Tsybakov (2004), p 89.

Hence

$$\|f_{\delta^{(i)}} - f_{\delta^{(j)}}\|^2 \geq (c_0^2)/(8m^{2r}).$$

We now check the condition (ii) in lemma 1. It is well known that for the Kullback-Leibler divergence in the case of an heteroscedastic gaussian sequence space model we have

$$K(f_\delta, f_0) = \frac{1}{\varepsilon^2} \sum_{l=1}^{\infty} \frac{|\theta_{\delta l} - \theta_{0l}|^2}{\sigma_l^2}. \quad (3.4)$$

Hence, by virtue of (3.2), (3.3) and (3.4), the Kullback-Leibler divergence between the two probability densities f_0 and f_δ for all $\delta \in \mathcal{C}$ satisfies

$$\begin{aligned} K(f_\delta, f_0) &= \frac{1}{\varepsilon^2} \frac{c_0^2}{m^{4r+3}} \sum_{l=m}^{2m-1} \frac{l^{2r+2} \delta_l^2}{\sigma_l^2} \\ &\leq \frac{c_0^2 2^{2r+2}}{\varepsilon^2} \frac{m^{2r+2}}{m^{4r+3}} \sum_{l=m}^{2m-1} \frac{\delta_l^2}{\sigma_l^2} \end{aligned}$$

but thanks to lemma 4 (see Appendix) we have

$$\frac{1}{\sigma_l^2} \leq 4^{l-1}$$

which implies

$$K(f_\delta, f_0) \leq \frac{c_0^2 2^{2r+4}}{\varepsilon^2} \frac{1}{m^{2r+1} 4^m} \sum_{l=m}^{2m-1} \delta_l^2 \leq \frac{c_0^2 2^{2r+4} m}{\varepsilon^2 4^m}$$

One chooses $m = \frac{1}{\log 4} \log(\frac{1}{\varepsilon^2})$ so that

$$K(f_\delta, f_0) \leq c_0^2 2^{2r+4} m$$

and since $m \leq 8 \log M / \log 2$ (see lemma 2)

$$K(f_\delta, f_0) \leq \frac{c_0^2 2^{2r+7}}{\log 2} \log M$$

Eventually one can choose c_0 small enough to have $c \equiv \frac{c_0^2 2^{2r+7}}{\log 2} < 1$.

We are now in position to apply lemma 1 with $\ell(x) = x^2$. We derive that, whatever the estimator \hat{f} ,

$$\sup_{f \in \mathcal{C}} \mathbb{E}_f [\|\hat{f} - f\|^2] \geq \frac{c_0^2}{32m^{2r}} \left[1 - \left(\frac{2}{3} \vee \frac{c}{\log M} \right) \right] \geq \frac{c_0^2}{96m^{2r}},$$

which gives the desired result.

4. Appendix

Lemma 3. For all $n \geq 1$ we have:

$$\binom{2n}{n} \leq \frac{4^n}{n^{1/4}} \quad (4.1)$$

Proof. Let us prove (4.1) by recursion on n . The inequality is clearly true for $n = 1$.

Suppose (4.1) true for a certain $n \geq 1$.

$$\binom{2(n+1)}{n+1} = \binom{2n}{n} \frac{2(2n+1)}{n+1} \leq \frac{4^n}{n^{1/4}} \frac{2(2n+1)}{n+1},$$

by recursion hypothesis. It remains to prove that

$$\frac{4^n}{n^{1/4}} \frac{2(2n+1)}{n+1} \leq \frac{4^{n+1}}{(n+1)^{1/4}}. \quad (4.2)$$

$$\begin{aligned} (4.2) &\iff \frac{2(2n+1)}{n^{1/4}(n+1)} \leq \frac{4}{(n+1)^{1/4}} \\ &\iff \frac{n+1}{n} \left(\frac{2n+1}{n+1}\right)^4 \leq 2^4 \\ &\iff \left(n + \frac{1}{2}\right)^{1/4} \leq n(n+1)^3, \end{aligned}$$

which is true because we have $(n + \frac{1}{2})^{1/4} \leq (n + \frac{1}{2})^3(n+1)$ and $(n + \frac{1}{2})^3 \leq n(n+1)^2$ since $\frac{1}{8} \leq n^2/2 + n/4$. This completes the proof.

Lemma 4. For all $n \geq 1$, we have:

$$\sigma_n^2 \geq 4^{n-1} \quad (4.3)$$

where σ_n is defined in (3.1).

Proof. Firstly, let us recall the value of the noise intensity σ_n^2 :

$$\begin{aligned} \sigma_n^2 &= \frac{2n+1}{2} \frac{1}{4^n} \sum_{j=0}^{\lfloor n/2 \rfloor} \left\{ \binom{n}{j} \binom{2n-2j}{n} \right\}^2 \\ &\geq \frac{n}{4^n} \binom{2n}{n}^2. \end{aligned}$$

And so, in order to prove (4.3) it remains to prove that

$$\binom{2n}{n} \geq \frac{4^n}{2\sqrt{n}} \quad n \geq 1.$$

We again use a recursion on n .

The inequality (4.3) is clear for $n = 1$. We suppose the property true for a certain $n \geq 1$ and we shall prove it at the rank $(n + 1)$.

$$\begin{aligned} \binom{2(n+1)}{n+1} &= \binom{2n}{n} \frac{2(2n+1)}{n+1} \\ &\geq \frac{4^n}{2\sqrt{n}} \frac{2(2n+1)}{n+1} \\ &> \frac{4^{n+1}}{2\sqrt{n+1}} \end{aligned} \tag{4.4}$$

the inequality (4.4) is true because it is equivalent to $4n^2 + 4n + 1 > 4n^2 + 4n$ what we always have.

References

- Abramowitz, M. and Stegun, I. A. (1970). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical tables*. Dover Publications, New York.
- Akhiezer, N. I. (1965). *The classical Moment Problem and Some Related Questions in Analysis*. Oliver and Boyd, Edinburgh.
- Ang, D. D., Gorenflo, R., Le, V. K. and Trong, D. D. (2002). Moment Theory and Some inverse Problems in Potential Theory and Heat Conduction. *Lectures Notes in Mathematics* **1792**, Springer, Berlin.
- Birgé, L. and Massart, P. (2001). Gaussian model selection. *J. Eur. Math. Soc.*, **3**, 203-268.
- Bertero, M., De Mol, C., and Pike, E. R. (1985). Linear inverse problems with discrete data. I: General formulation and singular system analysis. *Inverse Problems*, **1**, 301-330.

- Cavalier, L., Golubev, Y., Lepski, O., and Tsybakov, A. B. (2004). Block thresholding and sharp adaptative estimation in severely ill-posed inverse problems. *Theory Probab. Appl.*, **48**(3), 426-446.
- Fasino, D. and Inglese, G. (1996). Recovering a probability density from a finite number of moments and local a priori information. *Rend. Instituto. Mat. Univ. Trieste*, **28**(1), 184-200.
- Feller, W. (1968). *An introduction to Probability Theory and its Applications. Vol II. 2nd ed.* Wiley Series in Probability and Mathematical Statistics, New York.
- Goldenshluger, A. and Spokoiny, V. (2004). On the shape-from-moments problem and recovering edges from noisy Radon data. *Probab. Theory Relat. Fields*, **128**, 123-140.
- Golubev, Y. and Khasminskii, R. Z. (1999). A statistical approach to some inverse problems for partial differential equations. *Probl. Inf. Transm.*, **35**(2), 51-66.
- Hausdorff, F. (1921). Summationsmethoden und Momentfolgen, I. *Mathematische Zeitschrift*, **9**, 74-109.
- Inglese, G. (1995). Christoffel functions and finite moments problems. *Inverse problems*, **11**, 949-960.
- Shohat, J. A. and Tamarkin, J. D. (1943). *The problem of moments.* American Mathematical Society, New York.
- Tagliani, A. (2002). Entropy estimate of probability densities having assigned moments: Hausdorff case. *Appl. Mah. Lett.*, **15**(3), 309-314.
- Talenti, G. (1987). Recovering a function from a finite number of moments. *Inverse Problems*, **3**, 501-517.
- Tsybakov, A. B. (2004). *Introduction à l'estimation non-paramétrique.* Springer-Verlag. Berlin.

Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, Université Paris
6, case 188, 4, Pl. Jussieu, F-75252 Paris Cedex 5, France.

E-mail: phamngoc@ccr.jussieu.fr