Hausdorff moment problem via fractional moments

Pierluigi Novi Inverardi \(^a\), Giorgio Pontuale \(^b\),
Alberto Petri \(^b\), Aldo Tagliani \(^a,\)*

\(^a\) Faculty of Economics, Trento University, 38100 Trento, Italy
\(^b\) CNR, Istituto di Acustica “O.M. Corbino”, 00133 Roma, Italy

Abstract
We outline an efficient method for the reconstruction of a probability density
function from the knowledge of its infinite sequence of ordinary moments. The ap-
proximate density is obtained resorting to maximum entropy technique, under the
constraint of some fractional moments. The latter ones are obtained explicitly in terms
of the infinite sequence of given ordinary moments. It is proved that the approximate
density converges in entropy to underlying density, so that it turns out to be useful for
calculating expected values.

\(\text{Keywords: Entropy; Fractional moments; Hankel matrix; Maximum entropy; Moments}\)

1. Introduction

In applied sciences a variety of problems, formulated in terms of linear
boundary values or integral equations, leads to a Hausdorff moment problem.
Such a problem arises when a given sequence of real numbers may be repre-
sented as the moments around the origin of non-negative measure, defined on a
finite interval, typically \([0, 1]\). The underlying density \(f(x)\) is unknown, while its
moments \(\mu_j = \int_0^1 x^j f(x) \, dx, j = 0, 1, 2, \ldots, \mu_0 = 1\) are known. Next, through a
variety of techniques, for practical purposes \(f(x)\) is recovered by taking into

*Corresponding author.
E-mail address: ataglian@cs.unitn.it (A. Tagliani).

0096-3003/02/$ - see front matter \(\text{\textcopyright 2002 Elsevier Science Inc. All rights reserved.}\)
PII: S0096-3003(02)00391-0
account only a finite sequence \( \{\mu_j\}_{j=0}^M \). Such a process implies \( f(x) \) is well-characterized by first few moments. On the other hand, it is well known that the moment problem becomes ill-conditioned when the number of moments involved in the reconstruction increases [1,2]. In Hausdorff case, once fixed \((\mu_0, \ldots, \mu_{M-1})\), the moment \( \mu_M \) may assume values within the interval \([\mu_M^-, \mu_M^+]\), where [3]

\[
\mu_M^+ - \mu_M^- \leq 2^{-2(M-1)}.
\]

If one considers the approximating density \( f_M(x) = \exp(- \sum_{j=0}^M \lambda_j x^j) \) by entropy maximization, constrained by first \( M \) moments [4], then its entropy \( H[f_M] = - \int_0^1 f_M(x) \ln f_M(x) \, dx \) satisfies

\[
\lim_{\mu_M^+ - \mu_M^- \to 0} H[f_M] = -\infty.
\]

Such a relationship is satisfied by any other distribution constrained by same first \( M \) moments, since \( f_M(x) \) has maximum entropy. On the other hand \( f(x) \) and \( f_M(x) \) have same first \( M \) moments and as a consequence, as we illustrate in Section 3, the following relationship holds:

\[
I(f, f_M) = \int_0^1 f(x) \ln \frac{f(x)}{f_M(x)} \, dx = H[f_M] - H[f].
\]

Here \( H[f] \) is the entropy of \( f(x) \), while \( I(f, f_M) \) is the Kullback–Leibler distance between \( f(x) \) and \( f_M(x) \).

Eqs. (1.1)–(1.3) underline once more the ill-conditioning of moment problem.

The ill-conditioning may be even enlightened by considering the estimation of the parameters \( \lambda_j \) of \( f_M(x) \). The \( \lambda_j \) calculation leads to minimize a proper potential function \( \Gamma(\lambda_1, \ldots, \lambda_M) \) [4], with

\[
\min_{\lambda_1, \ldots, \lambda_M} \Gamma(\lambda_1, \ldots, \lambda_M) = \min_{\lambda_1, \ldots, \lambda_M} \left[ \ln \left( \int_0^1 \exp \left( - \sum_{j=1}^M \lambda_j x^j \right) \, dx \right) + \sum_{j=1}^M \lambda_j \mu_j \right].
\]

\( f_M(x) \) satisfies the constraints

\[
\mu_j = \int_0^1 x^j \exp \left( - \sum_{k=0}^M \lambda_k x^k \right) \, dx, \quad j = 0, \ldots, M.
\]

Letting \( \mu = (\mu_0, \ldots, \mu_M) \) and \( \lambda = (\lambda_0, \ldots, \lambda_M) \), (1.5) may be written as the map

\[
\mu = \phi(\lambda)
\]

Then the corresponding Jacobian matrix, which is up to sign a Hankel matrix, has conditioning number \( \simeq (1 + \sqrt{2})^{4M} / \sqrt{M} \) [5]. All the previous remarks lead
56 to the conclusion that $f(x)$ may be efficiently recovered from moments if only
57 few moments are requested. In other terms, $f(x)$ may be recovered from mo-
58 ments if its information content is spread among first few moments.
59 In this paper we are looking for a way to overcome the above-quoted dif-
60 culties in recovering $f(x)$ from moments. First of all, we assume known the
61 infinite sequence of moments $\{\mu_j\}_{j=0}^{\infty}$. Then, from such a sequence we calculate
62 fractional moments

$$E(X^{\alpha}) := \int_0^1 x^{\alpha} f(x) \, dx = \sum_{n=0}^{\infty} b_n(\alpha) \mu_n, \quad \alpha_j > 0,$$

(1.7)

where the explicit analytic expression of $b_n(\alpha)$ is given by (2.5). Finally, from a
65 finite number of fractional moments $\{E(X^{\alpha_j})\}_{j=1}^{M}$ we recover $f_M(x) = \exp(-\sum_{j=1}^{M} \alpha_j x^{\alpha_j})$ by entropy maximization [4]. The exponents $\{\alpha_j\}_{j=1}^{M}$ are
67 chosen as follows:

$$\{\alpha_j\}_{j=1}^{M} : H[f_M] = \text{minimum}. \quad (1.8)$$

The choice of $\{\alpha_j\}_{j=1}^{M}$, according to (1.8), leads to a density $f_M(x)$ having
69 minimum distance from $f(x)$, as stressed by (1.3).

71 **Remark.** If the information content of $f(x)$ is sheared among first moments, so
72 that ME approximant $f_M(x)$ represents an accurate approximation of $f(x)$,
73 then fractional moments may be accurately calculated by replacing $f(x)$ with
74 $f_M(x)$. Then $f_M(x)$ converges in entropy and then in $L_1$-norm to $f(x)$ [6] and the
75 error obtained replacing $f(x)$ with $f_M(x)$

$$|E_f(X^{\alpha_j}) - E_{f_M}(X^{\alpha_j})| \leq \int_0^1 x^{\alpha_j} |f(x) - f_M(x)| \, dx$$

$$\leq \int_0^1 |f(x) - f_M(x)| \, dx \leq \sqrt{2[H[f_M] - H[f]]}$$

(1.9)

may be rendered arbitrarily small by increasing $M$ (inequalities in (1.9) are
78 proved in Section 3).

79 **2. Fractional moments from moments**

80 Let $X$ be a continuous random variable with density $f(x)$ on the support
81 $[0, 1]$, with moments of order $s$, centered in $c$, $c \in \mathbb{R}$

$$\mu_s(c) := \mathbb{E}[(X - c)^s] = \int_0^1 (x - c)^s f(x) \, dx, \quad s \in \mathbb{N}^* = \mathbb{N} \cup \{0\} \quad (2.1)$$
and moments from the origin $\mu_s := \mu_s(0)$ related to moments generically centered in $c$ through the relationship

$$
\mu_s = \sum_{h=0}^{s} \binom{s}{h} c^{s-h} \mu_h(c), \quad s \in \mathbb{N}^*.
$$

(2.2)

It is well known the relationship similar to (2.2) which permits to calculate the (fractional) moment of order $s \in \mathbb{R}^+$ (which replaces $x_j$ for notational convenience as in (1.7) and (3.2) involving all the central moments of a given distribution about the point $c$.

Firstly, by definition of non-central moment of order $s$, we can write

$$
E(X^s) = \int_0^1 x^s f(x) \, dx
$$

and then, by Taylor expansion of $x^s$ around $c$, where $c \in (0,1)$, we have

$$
x^s = \sum_{n=0}^{\infty} \binom{s}{n} \frac{(x-c)^n}{n!} = \sum_{n=0}^{\infty} \left[ \left( \frac{s}{n} \right) n! x^{s-n} \right] \frac{(x-c)^n}{n!} = \sum_{n=0}^{\infty} \binom{s}{n} c^{s-n} (x-c)^n,
$$

(2.3)

where $[k(x)]^{(n)}_{x=c}$ indicates the $n$th derivative of the function $k(x)$ w.r.t. $x$, evaluated at $c$.

Taking the expectation on both sides of the last equation in (2.3), we get the required relationship

$$
E(X^s) = \sum_{n=0}^{\infty} \binom{s}{n} c^{s-n} E[(X-c)^n] = \sum_{n=0}^{\infty} b_n \mu_n(c),
$$

(2.4)

where

$$
b_n = \binom{s}{n} c^{s-n}, \quad n \in \mathbb{N}^*.
$$

(2.5)

represents the coefficient of the integral $n$-order moment of $X$ centered at $c$.

The formulation of the $s$-order fractional moments as in (2.4) shows some numerical instabilities which depend on the structure of the relationship between $\mu_n(c)$ and $E(X^s)$; these instabilities are related to the value of the center $c$ and increase as the order of the central moments becomes high. In particular,

(a) the numerical error $\Delta E(X-c)^n$ due to the evaluation of $E(X-c)^n$ in terms of non-central integral moments $E(X^h)$, $h \leq n$, becomes bigger as $c$ and $n$ increase. In fact,
\[ |\Delta \mathcal{E}(X - c)^n| = \left| \sum_{h=0}^{n} \binom{n}{h} (-1)^h c^{n-h} \Delta \mathcal{E}(X^h) \right| \]
\[ \leq \sum_{h=0}^{n} \binom{n}{h} c^{n-h} |\Delta \mathcal{E}(X^h)| \]
\[ = \|\Delta \mathcal{E}(X^h)\|_{\infty} \sum_{h=0}^{n} \binom{n}{h} c^{n-h} \]
\[ = \|\Delta \mathcal{E}(X^h)\|_{\infty} (1 + c)^n \simeq \text{eps}(1 + c)^n, \]

where \(\text{eps}\) corresponds to the error machine.

(b) the numerical error \(\Delta \mathcal{E}(X^s)\) due to the evaluation of \(\mathcal{E}(X^s)\) involving the first \(M_{\text{max}}\) central moments \(\mathcal{E}(X - c)^n\), is given by

\[ |\Delta \mathcal{E}(X^s)| = \sum_{n=0}^{M_{\text{max}}} \binom{s}{n} c^{s-n} |\Delta \mathcal{E}(X - c)^n| \]
\[ \leq \|\Delta \mathcal{E}(X - c)^n\|_{\infty} c^s \max_{n} \left( \binom{s}{n} \sum_{n=0}^{M_{\text{max}}} \left( \frac{1}{c} \right)^n \right) \]
\[ = \|\Delta \mathcal{E}(X - c)^n\|_{\infty} c^s \max_{n} \left( \frac{1}{c} \right)^{M_{\text{max}}+1} - 1 \]

with

\[ \max_{n} \left( \binom{s}{n} \right) = \left( \binom{s}{\lfloor s/2 \rfloor} \right) \text{ if } [s] \text{ is even} \]

and

\[ \max_{n} \left( \binom{s}{n} \right) = \left( \binom{s}{\lfloor s/2 \rfloor + 1} \right) \text{ if } [s] \text{ is odd,} \]

where \([x]\) represents the integer part of \(x\). The product of first two factors of the right hand side of (2.7) is an increasing function of \(c\), whilst the last factor gives a function which decreases with \(c\).

Hence, taking in account both (a) and (b), a reasonable choice of \(c\) could be \(c = 1/2\). Further, rewriting the last inequality in (2.7) as

\[ |\Delta \mathcal{E}(X)^s| \leq \|\Delta \mathcal{E}(X - c)^n\|_{\infty} c^s \max_{n} \left( \frac{1}{c} \right)^{M_{\text{max}}+1} - 1 < \varepsilon \]
we can reconstruct the \(s\)-order fractional moment with a prefixed level of accuracy \(\varepsilon, \varepsilon > 0\), just involving a number of central moments equal to the value \(M_{\text{max}}\).

3. Recovering \(f(x)\) from fractional moments

Let \(X\) be a positive r.v. on \([0, 1]\) with density \(f(x)\), Shannon-entropy \(H[f] = -\int_0^1 f(x) \ln f(x) \, dx\) and moments \(\{\mu_j\}_{j=0}^{\infty}\), from which positive fractional moments \(E(X^{a_j}) = \sum_{n=0}^{\infty} b_n(x_j) \mu_n\) may be obtained, as in (2.4) and (2.5).

From [4], we know that the Shannon-entropy maximizing density function \(f_M(x)\), which has the same \(M\) fractional moments \(E(X^{a_j})\), of \(f(x)\), \(j = 0, \ldots, M\), is

\[
    f_M(x) = \exp \left( -\sum_{j=0}^{M} \lambda_j x^{a_j} \right). \tag{3.1}
\]

Here \((\lambda_0, \ldots, \lambda_M)\) are Lagrangian multipliers, which must be supplemented by the condition that first \(M\) fractional moments of \(f_M(x)\) coincide with \(E(X^{a_j})\), i.e.,

\[
    E(X^{a_j}) = \int_0^1 x^{a_j} f_M(x) \, dx, \quad j = 0, \ldots, M, \quad x_0 = 1. \tag{3.2}
\]

Shannon-entropy \(H[f_M]\) of \(f_M(x)\) is given as

\[
    H[f_M] = -\int_0^1 f_M(x) \ln f_M(x) \, dx = \sum_{j=0}^{M} \lambda_j E(X^{a_j}). \tag{3.3}
\]

Given two probability densities \(f(x)\) and \(f_M(x)\), there are two well-known measures of the distance between \(f(x)\) and \(f_M(x)\). Namely the divergence measure

\[
    I(f, f_M) = \int_0^1 f(x) \ln \frac{f(x)}{f_M(x)} \, dx
\]

and the variation measure \(V(f, f_M) = \int_0^1 |f_M(x) - f(x)| \, dx\). If \(f(x)\) and \(f_M(x)\) have the same fractional moments \(E(X^{a_j}), \quad j = 1, \ldots, M\), then

\[
    I(f, f_M) = H[f_M] - H[f] \tag{3.4}
\]

holds. In fact
\[ I(f, f_M) = \int_0^1 f(x) \ln \frac{f(x)}{f_M(x)} \, dx = -H[f] + \sum_{j=0}^{M} \lambda_j \int_0^1 x^{\alpha_j} f_M(x) \, dx \\
= -H[f] + \sum_{j=0}^{M} \lambda_j E(X^{\alpha_j}) = H[f_M] - H[f]. \]

In the literature several lower bounds for the divergence measure \( I \) based on \( V \) are available. We shall however use the following bound [7]:

\[ I \geq \frac{V^2}{2}. \] (3.5)

If \( g(x) \) denotes a bounded function, such that \( |g(x)| \leq K, K > 0 \), taking into account (3.4) and (3.5), we have

\[ |E_f(g) - E_{f_M}(g)| \leq \int_0^1 |g(x)| \cdot |f(x) - f_M(x)| \, dx \leq K \sqrt{2(H[f_M] - H[f])}. \] (3.6)

Eq. (3.6) suggests to us what fractional moments have to be chosen

\[ \{\alpha_j\}_{j=1}^M : H[f_M] = \text{minimum}. \] (3.7)

The use of fractional moments in the framework of ME underlines on the following two theoretical results. The first is a theorem [8, Theorem 2] which guarantees the existence of a probability density from the knowledge of an infinite sequence of fractional moments.

**Theorem 3.1** ([8, Theorem 2]). If \( X \) is a r.v. assuming values from a bounded interval \([0,1]\) and \( \{\alpha_j\}_{j=0}^\infty \) an infinite sequence of positive and distinct numbers satisfying \( \lim_{j \to \infty} \alpha_j = 0 \) and \( \sum_{j=0}^\infty \alpha_j = +\infty \), then the sequence of moments \( \{E(X^{\alpha_j})\}_{j=0}^\infty \) characterizes \( X \).

The second concerns the convergence in entropy of \( f_M(x) \), where entropy-convergence means \( \lim_{M \to \infty} H[f_M] = H[f] \). More precisely.

**Theorem 3.2.** If \( \{\alpha_j\}_{j=0}^M \) are equispaced within \([0,1]\), with \( \alpha_{M-j+1} = j/(M+1) \), \( j = 0, \ldots, M \) then ME approximant converges in entropy to \( f(x) \).

**Proof.** See Appendix A. \( \square \)

We bound ourselves to point out that the choice of equispaced points \( \alpha_{M-j+1} = j/(M+1), j = 0, \ldots, M \) satisfies both conditions of Theorem 3.1, i.e.
\[ \lim_{M \to \infty} \gamma_M = 0 \quad \text{and} \quad \lim_{M \to \infty} \sum_{j=0}^{M} \gamma_j = \lim_{M \to \infty} \frac{1}{M+1} M (M+1) = +\infty. \]

As a consequence, if the choice of equispaced \( \gamma_{M-j+1} \) guarantees entropy-convergence then the choice (3.7) guarantees entropy-convergence too.

From a computational point of view, Lagrangian multipliers \( (\lambda_1, \ldots, \lambda_M) \) are obtained by (1.4), from which follows the normalizing constant \( \lambda_0 \), by imposing that the density integrates to 1. Then the optimal \( \{ \gamma_j \}_{j=1}^{M} \) exponents are obtained as

\[ \{ \gamma_j \}_{j=1}^{M} : \min_{\gamma_1, \ldots, \gamma_M} \left[ \min_{\lambda_1, \ldots, \lambda_M} \Gamma(\lambda_1, \ldots, \lambda_M) \right]. \quad (3.8) \]

4. Numerical results

We compare fractional and ordinary moments by choosing some probability densities on \([0, 1]\).

**Example 1.** Let
\[ f(x) = \frac{\pi}{2} \sin(\pi x) \]
with \( H[f] \approx -0.144729886 \). From \( f(x) \) we have ordinary moments satisfying the recursive relationship
\[ \mu_n = \frac{1}{2} - \frac{n(n-1)}{\pi^2} \mu_{n-2}, \quad n = 2, 3, \ldots, \quad \mu_0 = 1, \quad \mu_1 = \frac{1}{2}. \]

From \( \{ \mu_n \}_{n=0}^{\infty} \) we calculate \( E(X^2) = \sum_{n=0}^{\infty} b_n(\gamma_j)\mu_n \), as in (2.4) and (2.5).

From \( \{ E(X^2) \}_{j=0}^{M} \) we obtain ME approximant \( f_M(x) \) for increasing values of \( M \), where \( \{ \gamma_j \}_{j=1}^{M} \) satisfy (3.7).

In Table 1 are reported

(a) \( H[f_M] - H[f] = I(f, f_M) \) and exponents \( \{ \gamma_j \}_{j=1}^{M} \) satisfying (3.7), where \( H[f_M] \) is obtained using fractional moments.
(b) \( H[f_M] - H[f] = I(f, f_M) \), where \( H[f_M] \) is obtained using ordinary moments.

Inspection of Table 1 allows us to conclude that

- entropy decreasing is fast, so that practically 4–5 fractional moments determine \( f(x) \);
- on the converse a high number of ordinary moments are requested for a satisfactory characterization of \( f(x) \);
approximately 12 ordinary moments have an effect comparable to three fractional moments. $f(x)$ and $f_M(x)$, obtained by 4–5 fractional moments, are practically indistinguishable.

**Example 2.** This example is borrowed from [9]. Here the authors attempt to recover a non-negative decreasing differentiable function $f(x)$ from frequency moments $\omega_n$, with

$$\omega_n = \int_0^1 [f(x)]^n \, dx, \quad n = 1, 2, \ldots$$

The authors of [9] realize that other density reconstruction procedures, alternative to ordinary moments, would be desirable. We propose fractional moments density reconstruction procedure. Here

<table>
<thead>
<tr>
<th>$M$</th>
<th>${x_n}_{n=1}^M$</th>
<th>$H[f_M] - H[f]$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel a</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>13.4181</td>
<td>0.8716E-1</td>
</tr>
<tr>
<td>2</td>
<td>0.00289</td>
<td>0.2938E-2</td>
</tr>
<tr>
<td></td>
<td>4.69275</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.04680</td>
<td>0.3038E-3</td>
</tr>
<tr>
<td></td>
<td>1.84212</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.2143</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.00220</td>
<td>0.3276E-4</td>
</tr>
<tr>
<td></td>
<td>2.76784</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.7293</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20.5183</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.0024</td>
<td>0.1016E-4</td>
</tr>
<tr>
<td></td>
<td>2.7000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>20.500</td>
<td></td>
</tr>
<tr>
<td></td>
<td>25.200</td>
<td></td>
</tr>
<tr>
<td><strong>Panel b</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.9510E-2</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.2098E-2</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.7058E-3</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0.4442E-3</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.3357E-3</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>0.3288E-3</td>
</tr>
</tbody>
</table>
\[ f(x) = 2 \left[ \frac{1}{2} + \frac{1}{10} \ln \left( \frac{1}{Ax + B} - 1 \right) \right], \quad B = \frac{1}{1 + e^5}, \quad A \]

with \( H[f] \approx -0.06118227 \) \((f(x))\), compared to [9], contains the normalizing constant 2. From \( f(x) \) we have ordinary moments \( \mu_n \) through numerical procedure. From \{\( \mu_n \)\}_{n=0}^{\infty} we calculate \( E(X^n) = \sum_{n=0}^{\infty} b_n(z_j)\mu_n \), as in (2.4) and (2.5). From \{\( E(X^n) \)\}^{M}_{j=0} we obtain ME approximant \( f_M(x) \) for increasing values of \( M \), where \{\( z_j \)\}^{M}_{j=1} satisfy (3.7).

In Table 2 are reported

(a) \( H[f_M] - H[f] = I(f, f_M) \) and exponents \{\( z_j \)\}^{M}_{j=1} satisfying (3.7), where \( H[f_M] \) is obtained using fractional moments.

(b) \( H[f_M] - H[f] = I(f, f_M) \), where \( H[f_M] \) is obtained using ordinary moments.

Inspection of Table 2 allows us to conclude that

• entropy decreasing is fast, so that practically four fractional moments determine \( f(x) \);

<table>
<thead>
<tr>
<th>Panel a</th>
<th>( z_j )^{M}_{j=1}</th>
<th>( H[f_M] - H[f] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.56280</td>
<td>0.6278E–2</td>
</tr>
<tr>
<td>2</td>
<td>0.52500</td>
<td>0.3152E–2</td>
</tr>
<tr>
<td>3</td>
<td>1.05000</td>
<td>0.1169E–2</td>
</tr>
<tr>
<td></td>
<td>3.00000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.87500</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.44062</td>
<td>0.1025E–3</td>
</tr>
<tr>
<td></td>
<td>7.65470</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.5262</td>
<td></td>
</tr>
<tr>
<td></td>
<td>63.9093</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel b</th>
<th>( z_j )^{M}_{j=1}</th>
<th>( H[f_M] - H[f] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td>0.5718E–2</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.1776E–2</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.1320E–2</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0.6744E–3</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.3509E–3</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>0.2648E–3</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>0.1914E–3</td>
</tr>
</tbody>
</table>
• a high number of ordinary moments are requested for a satisfactory charac-
• approximately 14 ordinary moments have an effect comparable to four frac-
terization of \( f(x) \);
tional moments.
\( f(x) \) and \( f_M(x) \), obtained by four fractional moments, are practically in-
distinguishable. As a consequence we argue four fractional moments are
equivalent to eight frequency moments (as in [9]). The latter ones, indeed,
provide an approximant \( f_M(x) \) practically indistinguishable from \( f(x) \) (see Fig.
1 of [9]).

5. Conclusions

In this paper we have faced up to the Hausdorff moment problem and we
have solved it using a low number of fractional moments, calculated explicitly
in terms of given ordinary moments. The approximating density, constrained
by few fractional moments, has been obtained by maximum-entropy method.
Fractional moments have been chosen minimizing the entropy of the ap-
proximating density. The strategy of the paper in recovering a given density
function consists in accelerating the convergence by a proper choice of frac-
tional moments obtaining an approximating density, using moments of low
order as (1.1) suggests.

Appendix A. Entropy convergence

A.1. Some background

Let us consider a sequence of equispaced points \( x_j = j/(M + 1) \),
\( j = 0, \ldots, M \) and

\[
\mu_j =: E(X^{x_j}) = \int_0^1 t^{x_j} f_M(t) \, dt, \quad j = 0, \ldots, M \tag{A.1}
\]

with \( f_M(t) = \exp(- \sum_{j=0}^M \lambda_j t^{x_j}) \). With a simple change of variable \( x = t^{1/(M+1)} \),
from (A.1) we have

\[
\mu_j = E(X^{x_j}) = \int_0^1 x^j \exp \left[ - (\lambda_0 - \ln(M + 1)) - \sum_{j=1}^M \lambda_j x^j + M \ln x \right] \, dx, \quad j = 0, \ldots, M \tag{A.2}
\]
which is a reduced Hausdorff moment problem for each fixed $M$ value and a
determinate Hausdorff moment problem when $M \to \infty$. Referring to (A.2) the
following symmetric definite positive Hankel matrices are considered

\[
\begin{align*}
\Delta_0 &= \mu_0, \\
\Delta_2 &= \begin{bmatrix} 
\mu_0 & \mu_1 \\
\mu_1 & \mu_2 
\end{bmatrix}, \ldots, \\
\Delta_{2M} &= \begin{bmatrix} 
\mu_0 & \cdots & \mu_M \\
\vdots & \ddots & \vdots \\
\mu_M & \cdots & \mu_{2M} 
\end{bmatrix}
\end{align*}
\] (A.3)

whose $(i, j)$th entry $i, j = 0, 1, \ldots$ holds

\[
\mu_{i+j} = \int_0^1 x^{i+j} f_M(x) \, dx,
\]

where

\[
f_M(x) = \exp \left[ - (\lambda_0 - \ln(M + 1)) - \sum_{j=1}^{M} \lambda_j x^j + M \ln x \right].
\]

Hausdorff moment problem is determinate and the underlying distribution has
a continuous distribution function $F(x)$, with density $f(x)$. Then the mass $\rho(x)$ which can be concentrated at any real point $x$ is equal to zero ([10,
Corollary 2.8]). In particular, at $x = 0$ we have

\[
0 = \rho(0) = \lim_{i \to \infty} \rho_i^{(0)} =: \begin{bmatrix} 
|\Delta_2| \\
\mu_2 & \cdots & \mu_{i+1} \\
\vdots & \ddots & \vdots \\
\mu_{i+1} & \cdots & \mu_{2i} 
\end{bmatrix} = \lim_{i \to \infty} (\mu_0 - \mu_0^{(i)}),
\] (A.4)

where $\rho_i^{(0)}$ indicates the largest mass which can be concentrated at a given point
$x = 0$ by any solution of a reduced moment problem of order $\geq i$ and $\mu_0^{(i)}$
indicates the minimum value of $\mu_0$ once assigned the first $2i$ moments.

Let us fix $\{\mu_0, \ldots, \mu_{i-1}, \mu_{i+1}, \ldots, \mu_M\}$ while only $\mu_i, i = 0, \ldots, M$ varies con-
tinuously. From (A.2) we have

\[
\begin{bmatrix} 
\frac{d\lambda_0}{d\mu_i} \\
\vdots \\
\frac{d\lambda_M}{d\mu_i} 
\end{bmatrix} = -e_{i+1},
\] (A.5)

where $e_{i+1}$ is the canonical unit vector $\in \mathbb{R}^{M+1}$, from which
The following theorem holds.

**Theorem A.1.** If \( \alpha_j = j/(M + 1), j = 0, \ldots, M \) and \( f_M(x) = \exp\left(-\sum_{j=0}^{M} \lambda_j x^j\right) \) then

\[
\lim_{M \to \infty} H[f_M] = - \int_0^1 f_M(x) \ln f_M(x) \, dx = H[f] = - \int_0^1 f(x) \ln f(x) \, dx.
\]  

(A.7)

**Proof.** From (A.1) and (A.7) we have

\[
H[f_M] = \sum_{j=0}^{M} \lambda_j \mu_j.
\]  

(A.8)

Let us consider (A.8). When only \( \mu_0 \) varies continuously, taking into account (A.3)–(A.6) and (A.8) we have

\[
\frac{d}{d\mu_0} H[f_M] = \sum_{j=0}^{M} \mu_j \frac{d\lambda_j}{d\mu_0} + \lambda_0 = \lambda_0 - 1,
\]

\[
\frac{d^2}{d\mu_0^2} H[f_M] = \frac{d\lambda_0}{d\mu_0} + \begin{vmatrix} \mu_2 & \cdots & \mu_{M+1} \\ \vdots & \ddots & \vdots \\ \mu_{M+1} & \cdots & \mu_{2M} \end{vmatrix}
= - \frac{1}{\mu_0 - \mu_0^{(M)}} < 0.
\]

Thus \( H[f_M] \) is a concave differentiable function of \( \mu_0 \). When \( \mu_0 \to \mu_0^{(M)} \) then \( H[f_M] \to -\infty \), whilst at \( \mu_0 \) it holds \( H[f_M] > H[f] \), \( f_M(x) \) being the maximum entropy density once assigned \((\mu_0, \ldots, \mu_M)\). Besides, when \( M \to \infty \) then \( \mu_0^{(M)} \to \mu_0 \). So the theorem is proved. \( \square \)
References


