PERGAMON

Applied Mathematics Letters 16 (2003) 519-524

Applied Mathematics Letters

www.elsevier.nl/locate/aml

Maximum Entropy Solutions and Moment Problem in Unbounded Domains

A. TAGLIANI
Faculty of Economics, Trento University
38100 Trento, Italy

(*Received and accepted April 2002*)

Abstract—The classical Stieltjes and Hamburger moment problems in the maximum entropy approach have been reconsidered. Incorrect conditions of existence, previously appeared in literature, have been reviewed and improved. © 2003 Elsevier Science Ltd. All rights reserved.

Keywords—Entropy, Hankel matrix, Hausdorff, Hamburger, Stieltjes moment problems, Moment space.

1. INTRODUCTION

In past decades, maximum entropy (ME) distributions, having assigned moments, have been widely used in applied sciences to recover a discrete (or absolutely continuous) distribution on the basis of partial information. Much effort had been devoted to provide an answer to the main problems underlying the correct use of such distributions, like existence, convergence, and stability. On the basis of the attained results, the ME distributions have been involved in the numerical inversion of integral transforms, as Laplace and Mellin.

Results widely accepted in literature have been recently questioned in [1]. The above paper, concerning the conditions of existence of Stieltjes and Hamburger moment problems, gives different results compared to the ones provided in several previously published papers [2–4]. The discrepancy between Junk [1], Tagliani [2,3], and Frontini [4] consists in the evaluation of the admissible values of the highest-order moment employed, in order to guarantee the existence of the maximum entropy solution.

The purpose of the present note is:

(1) to prove that the conclusions in [1], as well as in [2–4], are only partially correct and consequently,
(2) to state the correct existence conditions underlying Stieltjes and symmetric Hamburger moment problems joining the results of the both papers.

2. THE PROBLEM IN STILTJES CASE

Our attention will be mainly addressed to the Stieltjes moment problem. The symmetric Hamburger case is similar.

 Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX
PII: S0893-9659(03)00030-2

Let $\mu = (1 = \mu_0, \mu_1, \dots, \mu_M)$ be a vector of given moments. The reduced Stieltjes moment problem consists in finding a probability density function $f(x)$ so that

$$\mu_i = \int_0^\infty x^i f(x)\, dx, \qquad i = 0, \dots, M. \tag{2.1}$$

Problem (2.1) is indeterminate and we call D^M the set of solutions such that

$$\mu(f) =: \int_0^\infty x^i f(x)\, dx, \qquad i = 0, \dots, M, \quad \forall\, f \in D^M.$$

A common way to regularize the problem is the maximum entropy principle (see [5]) in which a solution of (2.1) is singled out as minimizer of the strictly concave entropy functional

$$H[f] = -\int_0^\infty f(x) \ln f(x)\, dx$$

under constraints (2.1). The approximate density takes the analytical form (see [5])

$$f_M(x) = \exp\left(-\sum_{j=0}^{M} \lambda_j x^j\right) =: \exp_\lambda \tag{2.2}$$

satisfying

$$\mu_i = \int_0^\infty x^i f_M(x)\, dx, \qquad i = 0, \dots, M, \tag{2.3}$$

where $\lambda = (\lambda_0, \dots, \lambda_M)$ is the vector of Lagrange multipliers. The required integrability of $\exp_\lambda$ restricts the multipliers vector to the set

$$\Lambda = \left\{\lambda \in \mathbb{R}^{M+1} : \exp_\lambda \in L^1[0, +\infty)\right\}.$$

For $\lambda \in \Lambda$, the moments of $\exp_\lambda$ in any order are well defined so that the collection of integrable exponential densities

$$E^M =: \{\exp_\lambda : \lambda \in \Lambda\}$$

is a subset of D^M. In general $\mu(D^M)$ (the interior of moment space) will include strictly $\mu(E^M)$ (the moment space relative to the ME densities). Consequently, there are admissible moment vectors $\mu \in \mu(D^M)$ for which the moment problem (2.1) is solvable, but the ME problem (2.3) has no solution. This is the main result in [1], namely, for $M \geq 4$ then $\mu(E^M) \subset \mu(D^M)$ holds, in contrast to the result $\mu(E^M) = \mu(D^M)$ in [2,3]. More precisely, the procedure in [1] to obtain $\mu(E^M) \subset \mu(D^M)$ is as follows.

- Pick any $\lambda \in \Lambda \cap \partial\Lambda$. (In what follows int$(\Lambda)$ indicates the interior of a set Λ, while $\partial\Lambda$ is its boundary. $\lambda \in \Lambda \cap \partial\Lambda$ implies that highest component $\lambda_M = 0$.)
- Calculate the moment vector $\mu_\lambda = \mu(\exp_\lambda)$.
- Add any positive number to the highest component $\mu = \mu_\lambda + \epsilon e_{M+1}$, $\epsilon > 0$ and e_{M+1} the canonical unit vector $\in \mathbb{R}^{M+1}$.

Then μ is an admissible vector (namely, there exists a positive density $f(x)$ so that $\mu(f) = \mu$), but the ME problem with constraints μ has no solution.

The following are the opposing results in [1–4], respectively.

THEOREM 1.

$$\mu\left(D^M\right) \backslash \mu\left(E^M\right) = \left\{\mu : \mu > \mu\left(\exp_{\bar\lambda}\right),\ \bar\lambda \in \Lambda \cap \partial\Lambda\right\}.$$

In particular, the ME problem is solvable if and only if $\mu \in \mu(D^M)$ satisfies $\mu \not> \mu(\exp_{\bar\lambda})$ for all $\bar\lambda \in \Lambda \cap \partial\Lambda$.

THEOREM 2. *When $M \geq 4$, the ME problem is solvable if and only if $\mu \in \mu(D^M) \equiv \mu(E^M)$.*

The order relation

$$(u_0, \ldots, u_M) \geq (v_0, \ldots, v_M) \iff u_0 = v_0, \ldots, u_{M-1} = v_{M-1}, \qquad u_M \geq v_M$$

and the symmetric definite positive Hankel matrices

$$\Delta_{2k} = \|\mu_{i+j}\|_{i,j=0}^{k}, \qquad \Delta_{2k+1} = \|\mu_{i+j+1}\|_{i,j=0}^{k} \tag{2.4}$$

are introduced.

Let us fix $(\mu_0, \ldots, \mu_{i-1}, \mu_{i+1}, \ldots, \mu_M)$, with $i = 0, \ldots, M$, while only μ_i varies continuously. From (2.3) and (2.4), we have

$$\Delta_{2M} \cdot \begin{bmatrix} \dfrac{d\lambda_0}{d\mu_i} \\ \vdots \\ \dfrac{d\lambda_M}{d\mu_i} \end{bmatrix} = -e_{i+1} \tag{2.5}$$

whilst from (2.5), with $i = M$, we have

$$0 < \left[\frac{d\lambda_0}{d\mu_M}, \ldots, \frac{d\lambda_M}{d\mu_M}\right] \cdot \Delta_{2M} \cdot \begin{bmatrix} \dfrac{d\lambda_0}{d\mu_M} \\ \vdots \\ \dfrac{d\lambda_M}{d\mu_M} \end{bmatrix} = -\left[\frac{d\lambda_0}{d\mu_M}, \ldots, \frac{d\lambda_M}{d\mu_M}\right] \cdot e_{M+1} = -\frac{d\lambda_M}{d\mu_M}. \tag{2.6}$$

3. THE EXISTENCE CONDITIONS IN STIELTJES CASE

Without loss of generality, we may assume $\mu_1 = 1$, so that in the moment space $\mu(E^M)$ or $\mu(D^M)$ we include only $(\mu_2, \ldots, \mu_M)$, while μ_1 may be disregarded. Before facing the case $M = 4$ we review the existence conditions for ME solutions when $M = 2$ or $M = 3$ moments are assigned [1,2].

$M = 2$. $\mu(E^M) = \{\mu_2 : 1 < \mu_2 \leq 2\}$.

$M = 3$. The admissible values of (μ_2, μ_3) ($\equiv \mu(E^M)$) are shown in Figure 1 (region [a]). The moment space $\mu(D^M)$ is provided by regions [a] $\cup$ [b].

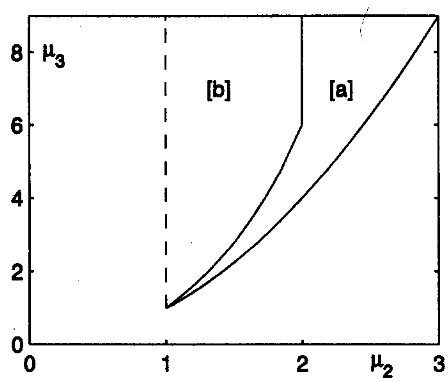

Figure 1. $M = 3$. [a]$\equiv \mu(E^M)$, [a] $\cup$ [b]$\equiv \mu(D^M)$. Lower boundary of [a] is given by $|\Delta_M| = 0$.

3.1. The Existence Conditions when $M = 4$

(i) Domain of μ_2

The admissible values of μ_2 stem from the case $M = 3$, putting $\lambda_4 = 0$. According to Figure 1, then the ME density exists for $\mu_2 > 1$.

(ii) Domain of μ_3

The following two cases have to be distinguished.

CASE 1. $\mu_2 > 2$. The admissible values of μ_3 stem from the case $M = 3$, putting $\lambda_4 = 0$. Then μ_3 does not admit an upper bound.

CASE 2. $\mu_2 \leq 2$. First, we consider the auxiliary density

$$f_4^{(1)}(x) = \exp\left(-\lambda_0 - \lambda_1 x - \lambda_2 x^2 - \lambda_4 x^4\right) \tag{3.1}$$

whose moments $(\mu_0, \dots, \mu_3)$ are assigned. Let (μ_0, μ_1, μ_2) be fixed, while μ_3 varies continuously. $\lambda_0, \lambda_1, \lambda_2, \lambda_4$, as well μ_4 are functions of μ_3. Differentiating both sides of (2.3), with $f_4(x)$ replaced by $f_4^{(1)}(x)$, we have

$$\begin{bmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_4 & 0 \\ \mu_1 & \mu_2 & \mu_3 & \mu_5 & 0 \\ \mu_2 & \mu_3 & \mu_4 & \mu_6 & 0 \\ \mu_3 & \mu_4 & \mu_5 & \mu_7 & 0 \\ \mu_4 & \mu_5 & \mu_6 & \mu_8 & 1 \end{bmatrix} \cdot \begin{bmatrix} \dfrac{d\lambda_0}{d\mu_3} \\[2ex] \dfrac{d\lambda_1}{d\mu_3} \\[2ex] \dfrac{d\lambda_2}{d\mu_3} \\[2ex] \dfrac{d\lambda_4}{d\mu_3} \\[2ex] \dfrac{d\mu_4}{d\mu_3} \end{bmatrix} = - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}. \tag{3.2}$$

Thus, λ_4 is a monotonic function. The character of monotonicity does not vary by varying μ_2, as λ_4 represents a family of disjointed curves. Equation (3.2) admits the solution $\lambda_4 = 0$, and therefore, λ_4 is monotonic decreasing for each value of μ_2, from which the relationship

$$\begin{vmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_4 \\ \mu_1 & \mu_2 & \mu_3 & \mu_5 \\ \mu_2 & \mu_3 & \mu_4 & \mu_6 \\ \mu_3 & \mu_4 & \mu_5 & \mu_7 \end{vmatrix} > 0. \tag{3.3}$$

Let us consider (3.1). It is easy to prove that the domain of the admissible values of (μ_2, μ_3) is given by region [a] of Figure 1. (Indeed, the upper boundary of region [a], when $1 < \mu_2 < 2$, is obtained putting $\lambda_3 = 0$, which is equivalent to $\lambda_4 = 0$, when (3.1) is considered.)

Let $f_4(x)$ be given by (2.2), having $(\mu_0, \dots, \mu_3)$ assigned. For each $(\mu_2, \mu_3) \in [a]$ and $\lambda_3 = 0$, then $f_4(x)$ exists. Let $(\mu_0, \dots, \mu_3)$ be fixed, while λ_3 varies continuously, assuming negative values only, starting from $\lambda_3 = 0$. Then $\lambda_0, \lambda_1, \lambda_2, \lambda_4$, as well μ_4 are functions of λ_3. Differentiating both sides of (2.3), we have

$$\begin{bmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_4 & 0 \\ \mu_1 & \mu_2 & \mu_3 & \mu_5 & 0 \\ \mu_2 & \mu_3 & \mu_4 & \mu_6 & 0 \\ \mu_3 & \mu_4 & \mu_5 & \mu_7 & 0 \\ \mu_4 & \mu_5 & \mu_6 & \mu_8 & 1 \end{bmatrix} \cdot \begin{bmatrix} \dfrac{d\lambda_0}{d\lambda_3} \\[2ex] \dfrac{d\lambda_1}{d\lambda_3} \\[2ex] \dfrac{d\lambda_2}{d\lambda_3} \\[2ex] \dfrac{d\lambda_4}{d\lambda_3} \\[2ex] \dfrac{d\mu_4}{d\lambda_3} \end{bmatrix} = - \begin{bmatrix} \mu_3 \\ \mu_4 \\ \mu_5 \\ \mu_6 \\ \mu_7 \end{bmatrix}. \tag{3.4}$$

Taking into account (3.4) and (3.3), λ_4 is monotonic decreasing. The solution $\lambda_4 = 0$ is not allowed, being $\lambda_3 < 0$. For each $(\mu_2, \mu_3) \in [a]$ and each $\lambda_3 = \lambda_3^* < 0$, then the auxiliary function

$$f_4^{(2)}(x) = \exp\left(-\lambda_0 - \lambda_1 x - \lambda_2 x^2 - \lambda_3^* x^3 - \lambda_4 x^4\right) \tag{3.5}$$

exists. Now we consider (3.5). Let μ_1 and μ_2 be fixed, whilst μ_3 varies continuously. The multipliers $\lambda_0, \lambda_1, \lambda_2, \lambda_4$, as well μ_4, are functions of μ_3. Differentiating both sides of (2.3), with $f_4(x)$ replaced by $f_4^{(2)}(x)$, we have (3.2). The determinant

$$\begin{vmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_4 \\ \mu_1 & \mu_2 & \mu_3 & \mu_5 \\ \mu_2 & \mu_3 & \mu_4 & \mu_6 \\ \mu_4 & \mu_5 & \mu_6 & \mu_8 \end{vmatrix}$$

is positive, being the principal minor of the definite positive matrix $P\Delta_8 P$, where P is a permutation matrix which exchanges last row and column with the previous one. Then λ_4 is monotonic decreasing, with $\lambda_4 > 0$, the particular solution $\lambda_4 = 0$ being not admissible because of the condition $\lambda_3 = \lambda_3^* < 0$. Therefore, μ_3 does not admit an upper bound.

Then the domain of the admissible values of (μ_2, μ_3) coincides with $\mu(D^3)$ (Figure 1, [a] $\cup$ [b]). (It might be in contrast with Junk [1], where the domain of the admissible values of (μ_2, μ_3) seems coincident with region [a].)

(iii) Domain of μ_4

Fixed $(\mu_2, \mu_3) \in [a] \cup [b]$, whilst μ_4 varies continuously we obtain (2.5). Then λ_4 is monotonic decreasing. The following two cases have to be distinguished.

(i) $(\mu_2, \mu_3) \in [b]$. The particular solution $\lambda_4 = 0$ is not allowed. Indeed, we are led back to the case $M = 3$, but in presence of the couple (μ_2, μ_3) not belonging to the domain of the admissible values $\mu(E^3)$ (such a result is in accordance with Theorem 2, but in contrast with Theorem 1).

(ii) $(\mu_2, \mu_3) \in [a]$. In such a case, $\lambda_4 = 0$ is allowed. The corresponding value $\mu_4 = \bar{\mu}_4 = \int_0^\infty x^4 f_3(x)\, dx \equiv \int_0^\infty x^4 f_4(x, \lambda_4 = 0)\, dx$ represents the upper bound of μ_4 for the ME solution (such a result is in accordance with Theorem 1, but in contrast with Theorem 2).

Following Theorem 1, the moment vectors $\mu \in \mu(D^M)$ for which the ME problem is not solvable are found only if $(\mu_2, \mu_3) \in [a]$. When $(\mu_2, \mu_3) \in [b]$, then $\mu(E^M) \equiv \mu(D^M)$ holds.

The above constructive procedure enables us to extend the existence conditions to the general case $M \geq 3$. Such results are summarized through the following theorem which improves both Theorems 1 and 2.

THEOREM 3. *Let $M \geq 3$. The domain of the admissible values of $(\mu_2, \ldots, \mu_M)$ which guarantees the existence of ME solution is as follows.*

1. *The moments $(\mu_2, \ldots, \mu_{M-1})$ do not admit any upper bound. Their values are provided by $|\Delta_2| > 0, \ldots, |\Delta_{M-1}| > 0$, respectively.*
2. *If $(\mu_2, \ldots, \mu_{M-1}) \in \mu(E^{M-1})$, then $\mu_M \leq \bar{\mu}_M = \int_0^\infty x^M f_{M-1}(x)\, dx$, while its lower bound is given by $|\Delta_M| > 0$.*
3. *If $(\mu_2, \ldots, \mu_{M-1}) \in \mu(D^{M-1}) \backslash \mu(E^{M-1})$, then μ_M does not admit an upper bound. Its values are given by $|\Delta_M| > 0$.*

From Theorem 3, the moment space $\mu(E^M)$, with $M \geq 3$, is obtained only numerically. Then for practical purposes, the use of ME distributions is quite cumbersome. Once given the vector $(\mu_0, \ldots, \mu_M)$, the existence of $f_M(x)$ is based only on the numerical evidence.

The procedure and then the results in the symmetric Hamburger case is similar to the Stieltjes one.

REFERENCES

1. M. Junk, Maximum entropy for reduced moment problems, *Mathematical Models and Methods in Applied Sciences* **10**, 1001–1025, (2000).
2. A. Tagliani, Maximum entropy in the problem of moments: An analytical study, *Statistica* **anno LII**, 533–547, (1992).
3. A. Tagliani, Maximum entropy in the Hamburger moments problem, *J. Math. Physics* **35**, 5087–5096, (1994).
4. M. Frontini and A. Tagliani, Maximum entropy in finite Stieltjes and Hamburger moment problem, *J. Math. Physics* **35**, 6748–6756, (1994).
5. H.K. Kesavan and J.N. Kapur, *Entropy Optimization Principles with Applications*, Academic Press, (1992).
6. A. Tagliani, On the application of maximum entropy to the moments problem, *J. Math. Physics* **34**, 326–337, (1993).