Journal of the Korean
Statistical Society
Vol. 24, No. 1, 1995

Existence and Uniqueness of the Smoothest Density with Prescribed Moments †

Changkon Hong [1] and Choongrak Kim [2]

ABSTRACT

In this paper we will prove the existence and uniqueness of the smoothest density with prescribed moments. The space of functions considered is the Sobolev space $W_m^2[0,1]$ and the target functional to be minimized is the seminorm $\| f^{(m)} \|_{L^2}$, which measures the roughness of the function f.

KEYWORDS : Moment, Sobolev Space, Smoothing Spline.

1. INTRODUCTION

The ordinary moments of a probability measure μ on $[a,b]$ are given by

$$c_i = \int_a^b t^i \, d\mu(t), \qquad i = 0, 1, 2, \cdots .$$

The moments for finite discrete mass distributions have been used in the physical sciences for a long time. The modern theory of moments begins in the

† This research was supported by the Ministry of Education (BSRI-93-116)
[1] Department of Statistics, Pusan National University, Pusan, 609-735, Korea
[2] Department of Statistics, Pusan National University, Pusan, 609-735, Korea

late nineteenth century with the work of A.A. Markov (1898), Stieltjes (1884) and Tchebysheff (1874). Modern versions of this theory are in Kreĭn and Nudel′man (1977) and Akhiezer (1961). The method of moments in statistical theory began about the same time with the work of Karl Pearson (1894) and Edgeworth (1886, 1887). Let

$$\mathsf{M}_n = \{(c_1, \cdots, c_n) \mid \int_a^b d\mu(t) = 1\}.$$

denote the convex set of all possible first n moments from probability measures on $[a,b]$. It is well known that for every $\mathbf{c} = (c_1, \cdots, c_n) \in \mathsf{M}_n$ there exist finite discrete measures with these moments. For finite discrete measures μ on $[a,b]$ with mass $p_1, \cdots, p_k$ at distinct points $x_1, \cdots, x_k$ we define the index of the measure $I(\mu)$ through its support $x_1, \cdots, x_k$ by counting one for each $x_i \in (a,b)$ and 1/2 for each $x_i \in \{a,b\}$. In the course of their investigation into the problem of moments Markov and Stieltjes showed that $\mathbf{c} \in \partial\mathsf{M}_n$ if and only if $\mathbf{c}$ has a unique representing finite discrete measure of index $I(\mu) \leq n/2$ and that each $\mathbf{c}$ in the interior of M_n has two representing measures $\underline{\mu}$ and $\bar{\mu}$ of index $(n+1)/2$. In case $n = 2m-1$, $\underline{\mu}$ has index m and its support is the m zeros of the m-th orthogonal polynomial $p_m(x)$ defined with respect to $(c_1, \cdots, c_n)$.

The purpose of this paper is to show the existence and uniqueness of the smoothest density with given moments $c_1, \cdots, c_n$. Here 'the smoothest' means that it achieves the minimum roughness. Without loss of generality we will assume $[a,b] = [0,1]$. We formulate the problem as follows;

Problem (P1)

$$\begin{array}{lll} \textit{Minimize} & J(f) \quad \textit{on} \quad H & \\ \textit{subject to:} & L_i f = \int_0^1 t^i f(t)\, dt = c_i \ , & i = 0, \cdots, n \\ & \textit{and } f(t) \geq 0 \quad \forall t \in [0,1]. & \end{array}$$

where $c_0 = 1$, J is a functional which represents the roughness of f and H is a normed linear space of functions on [0,1].

In this paper, we only consider the penalty functional $J(f) = \int_0^1 (f^{(m)}(t))^2 dt$ for $m \geq 1$. The natural space H for this penalty functional will be Sobolev space W_m^2 defined as:

$$\begin{aligned} W_m^2 &:= W_m^2[0,1] \\ &= \{f \text{ on } [0,1] \mid f^{(i)} \textit{ is absolutely continuous}, \ i = 0, \cdots, m-1, \\ &\qquad \text{and} \quad f^{(m)} \in L^2[0,1] \}, \end{aligned}$$

with inner product $<,>$,

$$< f, g > = \sum_{i=0}^{m-1} f^{(i)}(1) g^{(i)}(1) + \int_0^1 f^{(m)}(t) g^{(m)}(t) \, dt.$$

(cf. Adams (1975)).

2. EXISTENCE AND UNIQUENESS OF THE MINIMIZER

For the problem (P1) to make sense, there must exist at least one density $f \in W_m^2$ with the given moments $c_1, \cdots, c_n$. It can be shown that there exist infinitely many densities in W_m^2 with the given moments $\mathbf{c} = (c_1, \cdots, c_n)$ as long as $\mathbf{c}$ is an interior point of M_n. Let $\mathbf{c}$ be an interior point of M_n. Then there exist $n+1$ points $\mathbf{c}^{(0)}, \cdots, \mathbf{c}^{(n)}$ in M_n such that the convex hull $Co\{\mathbf{c}^{(0)}, \cdots, \mathbf{c}^{(n)}\}$ of $\mathbf{c}^{(0)}, \cdots, \mathbf{c}^{(n)}$ contains $\mathbf{c}$ as an interior point. For each $\mathbf{c}^{(i)}$, there exists a discrete probability measure $\sigma^{(i)}$ with finite support. We can then choose a sequence $\{\sigma_j^{(i)}\}_{j=1}^{\infty}$ of probability measures with densities in W_m^2 such that $\sigma_j^{(i)}$ converges weakly to $\sigma^{(i)}$ as $j \to \infty$. Let $\mathbf{c}_j^{(i)}$ be the first n moments of $\sigma_j^{(i)}$. Since $\mathbf{c}$ is an interior point of $Co\{\mathbf{c}^{(0)}, \cdots, \mathbf{c}^{(n)}\}$, we can choose j^* such that $\mathbf{c}$ is an interior point of $Co\{\mathbf{c}_j^{(0)}, \cdots, \mathbf{c}_j^{(n)}\}$ for all $j \geq j^*$. This proves the assertion.

The difficult part of problem (P1) is imposing the nonnegativity constraint on our density $f(t)$. The non-constrained problem where f is allowed negative and positive values is relatively easy and is related to problems in generalized spline smoothing. The classical problem is to fit a smooth curve $f \in W_m^2$ through "data points" $(t_i, y_i = f(t_i))$, $i = 1, \cdots, n$ such that $J(f)$ is minimized. Here the solution is the well known natural spline of order $2m-1$, that is, it is a spline $S(t)$ of order $2m-1$ with knots $t_1, \cdots, t_n$ and with the end point conditions $s^{(i)}(t) = 0$ for $i = m, \cdots, 2m-1$ and for all $t \leq t_1$, $t \geq t_n$. (cf. Wahba (1990) and Eubank (1988)). The evaluation linear functionals $L_i(f) = f(t_i)$, $i = 1, \cdots, n$ are continuous in the Hilbert

Space W_m^2. In our problem we simply change these to the linear functionals $L_i(f) = \int_a^b t^i f(t)\, dt$, $i = 0, \cdots, n$. The solution in either case involves the "representers" of the linear functionals and its existence and uniqueness are obvious. In the classical case the solution is a spline and in our case it is simply a polynomial of degree $\leq 2m + n$. The unconstrained solution to the problem of finding the smoothest density with moments $c_1, \cdots, c_n$ is the unique polynomial f of degree $\leq 2m + n$ satisfying the n moment conditions and the corresponding boundary conditions $f^{(k)}(0) = f^{(k)}(1) = 0$, $k = m, \cdots, 2m - 1$. As mentioned earlier the nonnegativity constraint makes the problem considerably more difficult and the existence and uniqueness are not obvious anymore. To prove the existence and uniqueness of the solution, we need some mathematical definitions and theorems.

2.1 Background Mathematics

Every optimization problem can be expressed as the following form.
Problem (P):
Given a normed linear space $(X, \| \|)$, an objective functional J defined on X and a constraint set $M \subset X$,

$$\textit{Minimize} \quad J(x) \quad \textit{on} \quad M.$$

Without proofs we state some propositions pertinent to the optimization problem (P).

Proposition 2.1. If $S \subset X$ is convex and closed, then S is weakly closed. If X is a reflexive Banach space and S is convex, closed and bounded, then S is weakly compact.

Proof. See Yoshida (1971). □

Proposition 2.2. Let D be a subset of X. A functional $J : X \to R$ is weakly lower semicontinuous in D if and only if the set $D_m = \{x \in D \mid J(X) \leq m\}$ is weakly closed for all real numbers m.

Proof. See Tapia and Thompson (1990). □

Proposition 2.3. If the set M in problem (P) is weakly compact and the functional J is weakly lower semicontinuous on M, then problem (P) has at least one minimizer.

Proposition 2.4. Suppose $M \subset X$ is nonempty, convex and closed and the functional $J : M \to R$ is continuous and convex in M. Then J is weakly lower semicontinuous in M.

Definition. Let S be a subset of X and J be a functional on X. J is said to have the infinity property in S if $\{x_n\} \subset S$ and $\| x_n \| \to \infty$ as $n \to \infty$ implies $J(x_n) \to \infty$.

The following proposition can be proved using Proposition 2.3 and 2.4.

Proposition 2.5. Suppose X is a reflexive Banach space and M is a closed convex subset of X. If the objective functional J has the infinity property in M and it is continuous and convex in M, then problem (P) has at least one minimizer.

Proposition 2.6. Assume that $J : X \to R$ is twice Gâteaux differentiable in a convex subset S of X. Then

1. J is convex in S $\Leftrightarrow$ J'' is p.s.d. relative to S.
2. J is strictly convex in S $\Leftarrow$ J'' is p.d. relative to S.
3. J is uniformly convex in S with constant C $\Leftrightarrow$ J'' is u.p.d. relative to S with constant $2C$.

2.2. Existence and uniqueness

We now prove the existence and uniqueness of the minimizer for the problem (P1). From now on $J(f)$ will refer to $\int_0^1 (f^{(m)}(t))^2\,dt$ and S will refer to a subset of a Sobolev space. For given moments $\{c_1, \cdots, c_n\}$, let us define

$$S = \{f \in W_m^2 \mid \begin{array}{ll} \text{i)} & L_i f = c_i, i = 0, \cdots, n, \\ \text{ii)} & f(t) \geq 0, \ \ \forall t \in [0,1] \end{array} \},$$

where $c_0 = 1$ and L_i is the i-th moment functional.

Lemma 2.1. S is convex and closed.

Proof. Convexity is obvious. Since W_m^2 is a Reproducing Kernel Hilbert Space (R.K.H.S , cf. Aronszajn (1950)), the evaluation functional $t(f) = f(t)$ for fixed t is a continuous linear functional. It is sufficient to show that L_i is continuous. Because L_i is linear, it is continuous if and only if it is continuous at 0. Using integration by parts,

$$\begin{aligned} | L_i f | &= | \int_0^1 t^i f(t)\, dt | \\ &\leq | \sum_{j=0}^{m-1} (\prod_{k=0}^{j} (i+k+1))^{-1} f^{(j)}(1) | \\ &\quad + | \int_0^1 (\prod_{j=0}^{m-1} (i+k+1))^{-1} t^{m+i} f^{(m)}(t)\, dt | . \end{aligned}$$

By the Hölder inequality, we get

$$| L_i f | \leq D_1 (\mathbf{f}(1)^T \mathbf{f}(1))^{\frac{1}{2}} + D_2 \| f^{(m)} \|_{L^2}, \tag{2.1}$$

where

$$\begin{aligned} D_1 &= (\sum_{j=0}^{m-1} (\prod_{k=0}^{j} (i+k+1))^{-2})^{1/2}, \\ D_2 &= (\prod_{k=0}^{m-1} (i+k+1))^{-1} (2m+2i+1)^{-1/2}, \\ \mathbf{f} &= (f, f^{(1)}, \cdots, f^{(m-1)})^T. \end{aligned}$$

Choose $\{f_n\} \in W_m^2$ which converges to 0, i.e., such that $\| f_n \| \to 0$, as $n \to \infty$. Since $\| f_n \|^2 = \mathbf{f}_n(1)^T \mathbf{f}_n(1) + \| f_n^{(m)} \|_{L^2}^2$, inequality (??) shows that $\lim_{n\to\infty} \| f_n \| = 0$ implies $L_i(f_n) \to 0$ as $n \to \infty$, $\forall i$, and hence L_i is continuous at 0, $\forall i$. □

Lemma 2.2. J is continuous and convex. If $n \geq m-1$, then J is strictly convex on S.

Proof. Continuity of J is obvious. The first and second Gâteaux derivatives are given by

$$J'(f)(g) = 2\int_0^1 f^{(m)}(t)\, g^{(m)}(t)\, dt,$$
$$J''(f)(g,g) = 2\int_0^1 (g^{(m)}(t))^2\, dt.$$

Since the inequality $J''(f)(g,g) \geq 0$ obviously holds for all $g \in T(S,f) = \{h \in W_m^2 \mid \exists\lambda \ \textit{such that}\ f + \lambda h \in S\}$, by Proposition 2.3 J is convex. Suppose $J''(f)(g,g) = 0$ for some $g \in T(S,f)$, then $g^{(m)}(t) = 0, a.e.$. This implies that $g \in \Pi_{m-1}[0,1]$, where $\Pi_{m-1}[0,1]$ is the space of polynomials of degree $\leq m-1$. Since $g \in T(S,f)$,

$$L_i g = 0, \quad \forall i = 0, \cdots, n.$$

Let $g(t) = \sum_{j=0}^{m-1} a_j t^j$. Then we get

$$\mathbf{B}\,\mathbf{a} = \mathbf{0}, \tag{2.2}$$

where

$$\mathbf{B} = (b_{ij}), \qquad i = 0,\cdots,n, \ \ j = 0,\cdots,m-1,$$
$$b_{ij} = \int_0^1 t^{i+j}\, dt,$$
$$\mathbf{a} = (a_0,\cdots,a_{m-1})^T.$$

If $n \geq m-1$, then $\mathbf{B}$ has full column rank and Equation (2.2) has trivial solution as unique solution, this in turn implies $g \equiv 0$. Therefore when $n \geq m-1$ J is strictly convex. □

Lemma 2.3. If $n \geq m-1$, then J has the infinity property in S.

Proof. Since $\| f \|^2 = \mathbf{f}(1)^T\mathbf{f}(1) + J(f)$, it is sufficient to show that $\mathbf{f}(1)^T\mathbf{f}(1)$ is dominated by $J(f)$. Remember that every element f in S satisfies the moment constraints, i.e., $(L_0 f, \cdots, L_n f) = \mathbf{c}$. Now

$$c_i = L_i f = \mathbf{e}_i^T\mathbf{f}(1) + \int_0^1 h_i(t) f^{(m)}(t)\, dt, \tag{2.3}$$

where

$$e_{ij} = (-1)^j \prod_{k=0}^{j} (i+k+1)^{-1}, \quad i = 0, \cdots, n, \quad j = 0, \cdots, m-1,$$

$$h_i(t) = (-1)^m \prod_{j=1}^{m-1} (i+k+1)^{-1} t^{m+i}, \quad i = 0, \cdots, n.$$

Equation (2.3) yields

$$\mathbf{E}\,\mathbf{f}(1) = \mathbf{c} - \int_0^1 \mathbf{h}(t) f^{(m)}(t)\, dt,$$

where $\mathbf{E} = (\mathbf{e}_0, \cdots, \mathbf{e}_n)^T$ and $\mathbf{h} = (h_0, \cdots, h_n)$. Since $\mathbf{E}$ is of full column rank when $n \geq m-1$,

$$\mathbf{f}(1) = (\mathbf{E}^T\mathbf{E})^{-1}\,\mathbf{E}\,\mathbf{c} - (\mathbf{E}^T\mathbf{E})^{-1}\mathbf{E}^T \int_0^1 \mathbf{h}(t) f^{(m)}(t)\, dt \tag{2.4}$$

$$= \mathbf{d} + \int_0^1 \mathbf{g}(t) f^{(m)}(t)\, dt, \tag{2.5}$$

where $\mathbf{d} = (\mathbf{E}^T\mathbf{E})^{-1}\mathbf{E}\,\mathbf{c}$ and $\mathbf{g}(t) = -(\mathbf{E}^T\mathbf{E})^{-1}\mathbf{E}\,\mathbf{h}(t)$. By using (2.4) and the Hölder inequality, we get

$$\begin{aligned}
\mathbf{f}(1)^T\mathbf{f}(1) &= \mathbf{d}^T\mathbf{d} + \int_0^1 2\mathbf{d}^T\mathbf{g}(t) f^{(m)}(t)\, dt \\
&\quad + \left(\int_0^1 \mathbf{g}(t) f^{(m)}(t)\, dt\right)^T \left(\int_0^1 \mathbf{g}(t) f^{(m)}(t)\, dt\right) \\
&\leq \mathbf{d}^T\mathbf{d} + \| 2\mathbf{d}^T\mathbf{g} \|_{L^2} \| f^{(m)} \|_{L^2} \\
&\quad + \sum_{i=0}^{n} \left(\int_0^1 g_i(t) f^{(m)}(t)\, dt\right)^2 \\
&\leq \mathbf{d}^T\mathbf{d} + \| 2\mathbf{d}^T\mathbf{g} \|_{L^2} J(f)^{1/2} + \sum_{j=0}^{n} \| g_i \|_{L^2}^2 J(f).
\end{aligned}$$

So we get an upper bound for $\| f \|^2$,

$$\| f \|^2 \leq \mathbf{d}^T\mathbf{d} + \| 2\mathbf{d}^T\mathbf{g} \|_{L^2} J(f)^{1/2} + (1+\gamma) J(f),$$

where $\gamma = \sum_{i=0}^{n} \| g_i \|_{L^2}^2$, as was to be shown. □

Now we state and prove the existence and uniqueness of the minimizer for problem (P1).

Theorem 2.1. (Existence and uniqueness of the solution of problem (P1)) If $n \geq m-1$, then problem (P1) has a solution and it is unique.

Proof. Since W_m^2 is a Hilbert space, it is a reflexive Banach Space. Therefore, the existence of the solution follows from Lemmas 2.1, 2.2 and 2.3 and Proposition 2.5. By Lemma 2.2 J is strictly convex provided $n \geq m-1$. Suppose $f_1 \neq f_2$ and both assumes the minimum, i.e., $J(f_1) = \inf_{f \in S} J(f) = J(f_2)$. Let $f_\alpha = \alpha f_1 + (1-\alpha) f_2$, where $\alpha \in (0,1)$. Then by strict convexity of J,

$$\begin{aligned} J(f_\alpha) &< \alpha J(f_1) + (1-\alpha) J(f_2) \\ &= J(f_1). \end{aligned}$$

This contradicts that f_1 is a minimizer for problem (P1). □

Remark. When $m = 2$, the existence of the minimizer for problem (P1) is guaranteed even when the condition $n \geq m-1$ is violated. For the value $n = 0$, any nonnegative linear function f on [0,1] which integrates to one is a minimizer.

REFERENCES

(1) Adams, R. A. (1975). *Sobolev Spaces*, Academic Press, New York.

(2) Akhiezer, N. I. (1961). *The Classical Moment Problem and Some Related Topics in Analysis*, Fizmatgiz, Moscow; English Transl., Hafner, New York. MR 27 #4028; 32 #1518.

(3) Aronszajn, N. (1950). Theory of reproducing kernels. *American Mathematical Society:Transactions*, **68**, 337-404.

(4) Edgeworth, F. Y. (1886). The law of error and the elimination of chance. *Philosophical Magazine, 5th series*, **21**, 308-324.

(5) Edgeworth, F. Y. (1887). The emphirical proof of the law of error. *Philosophical Magazine, 5th series*, **24**, 330-342.

(6) Eubank, R. L. (1988). *Spline Smoothing and Nonparametric Regression*, M. Dekker, New York.

(7) Kreĭn, M. G. and Nudel'man, A. A. (1977). *The Markov Problem and Extremal Problems*, Translations of Mathematical Monographs, vol 50. American Mathematical Society, Providence, Rhode Island.

(8) Markov, A. A. (1898). Sur les valeurs extrèmes des intégrales et l'interpolation. *Mémoires de l'Académie des Sciences de St.-Petersburg, classe physico-mathématique*, **(8) 6 (5)**; 69.

(9) Pearson, K. (1894). Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London (A)*, **185**, 71–110. *Page references are to the reprint in Pearson*, **1956**, 1–40.

(10) Stieltjes, T. J. (1884). Quelques recherches sur les quadratures dites mécaniques. *Annales Scientifiques de l'Ecold Normale Supérieure*, **1 (3)**, 409–426.

(11) Tapia, J. R. and Thompson, R. A. (1990). *Nonparametric Function Estimation, Modeling, and Simulation*. SIAM.

(12) Tchebycheff, P. (1874). Sur les valeurs limits des intégrales. *Journal des Mathématiques pures et eppliquées* **19 (2)** ; 157–160. Collected Papers 2, No 10, 181-185.

(13) Wahba, G. (1990). *Spline Models for Observational Data*, CBMS 59. SIAM.

(14) Yoshida, K. (1971). *Functional Analysis, 3rd edition*, Springer–Verlag, Berlin–New York.