Journal of Approximation Theory **92**, 463–471 (1998)
Article No. AT973135

# A Minimax Problem Admitting the Equioscillation Characterization of Bernstein and Erdős*

Ying Guang Shi

*Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, P.O. Box 2719, Beijing, 100080, People's Republic of China*
E-mail: syg@lsec.cc.ac.cn

*Communicated by A. Pinkus*

Received May 23, 1996; accepted in revised form March 15, 1997

This paper shows that under certain conditions a solution of the minimax problem $\min_{a<x_1<\cdots<x_n<b} \max_{1\leqslant i\leqslant n+1} f_i(x_1, ..., x_n)$ admits the equioscillation characterizations of Bernstein and Erdős and has strong uniqueness. This problem includes as a particular example the optimal Lagrange interpolation problem.


## 1. INTRODUCTION

As we know, the equioscillation conditions of Bernstein [2] and Erdős [7] were originally conjectured to characterize optimal Lagrange interpolation. They have been confirmed not only for the original interpolation problem but also for many other optimal (extremal) problems [3, 5, 8–12]. It is natural to ask what kind of extremal problems admit the equioscillation characterizations of Bernstein and Erdős. The first aim of this paper is to investigate a general minimax problem, a solution of which admits characterizations of Bernstein and Erdős and which includes most of the above-mentioned examples. The second aim is to provide different proofs (based on Lagrange's method of multipliers for nonlinear programming with constraints). In addition, this paper also gives some new results, including an analogue of strong uniqueness for best uniform approximation by a Haar subspace [6, p. 81].

The minimax problem mentioned above is the following.

Let $-\infty<a<b<\infty$, $n\geqslant 1$, and

$$X:=\{\mathbf{x}=(x_1, x_2, ..., x_n): x_0:=a<x_1<x_2<\cdots<x_n<x_{n+1}:=b\}.$$

* The Project 19671082 supported by National Natural Science Foundation of China.



0021-9045/98 $25.00

Let $f_i(\mathbf{x}) \geqslant 0$, $i = 1, 2, ..., n+1$, be continuously differentiable functions on $X$ and

$$f(\mathbf{x}) := \max_{1 \leqslant i \leqslant n+1} f_i(\mathbf{x}).$$

We will investigate the minimax problem: Find a vector $\mathbf{y} \in X$ such that

$$f(\mathbf{y}) = \inf_{\mathbf{x} \in X} f(\mathbf{x}).$$

The main results of this paper are the following equioscillation characterizations of Bernstein and Erdős for such a minimax solution.

THEOREM 1. *Assume that the functions $f_i$ satisfy the conditions*

$$\lim_{\min_{0 \leqslant j \leqslant n}(x_{j+1} - x_j) \to 0} \max_{1 \leqslant i \leqslant n} |f_{i+1}(\mathbf{x}) - f_i(\mathbf{x})| = \infty \tag{1.1}$$

*and*

$$D_k(\mathbf{x}) := \det \left( \frac{\partial f_i(\mathbf{x})}{\partial x_j} \right)_{j=1,\ i=1,\ i \neq k}^{n\ \ n+1} \neq 0, \qquad \mathbf{x} \in X, \quad k = 1, 2, ..., n+1. \tag{1.2}$$

*Then the following statements are valid*:

(a) *there exists a unique vector $\mathbf{y} \in X$ such that*

$$f(\mathbf{y}) = \min_{\mathbf{x} \in X} f(\mathbf{x}); \tag{1.3}$$

(b) *Equation* (1.3) *holds if and only if*

$$f_1(\mathbf{y}) = f_2(\mathbf{y}) = \cdots = f_{n+1}(\mathbf{y}); \tag{1.4}$$

(c) *for any vector $\mathbf{x} \in X \backslash \{\mathbf{y}\}$*

$$\min_{1 \leqslant i \leqslant n+1} f_i(\mathbf{x}) < f(\mathbf{y}) < \max_{1 \leqslant i \leqslant n+1} f_i(\mathbf{x}); \tag{1.5}$$

(d) *for each vector $\mathbf{w} \in X$ there exists a constant $\gamma(\mathbf{w}) > 0$ such that for all $\mathbf{x} \in X$*

$$\max_{1 \leqslant i \leqslant n+1} [f_i(\mathbf{x}) - f_i(\mathbf{w})] \geqslant \gamma(\mathbf{w}) \, \|\mathbf{x} - \mathbf{w}\|, \tag{1.6}$$

*where* $\|\mathbf{x}-\mathbf{w}\| = \max_{1 \leqslant j \leqslant n} |x_j - w_j|$; *in particular, there exists a constant* $\gamma^* > 0$ *such that for all* $\mathbf{x} \in X$

$$f(\mathbf{x}) \geqslant f(\mathbf{y}) + \gamma^* \|\mathbf{x}-\mathbf{y}\|. \tag{1.7}$$

We note that the actual verification of the assumptions (1.1) and (1.2) for the special case of Lagrange interpolation occupies a nontrivial part of [3].

We will also give other results, which are already proved in [3].

THEOREM 2. *Let* (1.1) *and* (1.2) *hold. Then*

(a) *the map* $F: X \to \mathbb{R}^n: \mathbf{x} \mapsto (f_{i+1}(\mathbf{x}) - f_i(\mathbf{x}))_{i=1}^n$ *is a homeomorphism of* $X$ *onto* $\mathbb{R}^n$;

(b) *if for* $\mathbf{u}, \mathbf{v} \in X$,

$$f_i(\mathbf{u}) \leqslant f_i(\mathbf{v}), \qquad i = 1, 2, ..., n+1,$$

*then* $\mathbf{u} = \mathbf{v}$;

(c) *for each fixed* $k$, $1 \leqslant k \leqslant n+1$, *the map* $F_k: X \to \mathbb{R}^n: \mathbf{x} \mapsto (f_i(\mathbf{x}))_{i \neq k}$ *is (globally) one-to-one.*

It is worth pointing out that Statements (a)–(d) of Theorem 1 are the analogues of the theorems of existence and uniqueness, alternation, de La Vallée Poussin, and strong uniqueness for best uniform approximation by a Haar subspace, respectively [6, pp. 72–81].

Our proofs are based on the following, in which

$$\nabla h(\mathbf{y}) = \left( \frac{\partial h(\mathbf{x})}{\partial x_1}, ..., \frac{\partial h(\mathbf{x})}{\partial x_n} \right) \Bigg|_{\mathbf{x}=\mathbf{y}}.$$

THEOREM A [1, Theorem 3.4]. *Assume that* $g, g_1, ..., g_m$ *are continuously differentiable on an open set* $S \subset \mathbb{R}^n$. *If* $\mathbf{y} \in S$ *is a solution of the problem to minimize* $g(\mathbf{x})$ *subject to* $\mathbf{x} \in S$ *and* $g_i(\mathbf{x}) \geqslant 0$, $i = 1, 2, ..., m$, *then there exists a vector* $\lambda = (\lambda_0, \lambda_1, ..., \lambda_m) \neq 0$, $\lambda \geqslant 0$, *such that*

$$\lambda_0 \nabla g(\mathbf{y}) - \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{y}) = 0$$

*and*

$$\lambda_i g_i(\mathbf{y}) = 0, \qquad i = 1, 2, ..., m.$$

We put some auxiliary lemmas in the next section and the proofs of the theorems in the last section.

## 2. AUXILIARY LEMMAS

LEMMA 1. *If* (1.1) *holds then there exists a vector* $\mathbf{y} \in X$ *satisfying* (1.3).

*Proof.* Since the $f_i$ are nonnegative, (1.1) implies

$$\lim_{\min_{0 \leqslant j \leqslant n}(x_{j+1} - x_j) \to 0} f(\mathbf{x}) = \infty. \tag{2.1}$$

Hence, the infimum may be taken over, instead of $X$, the compact subset of $X$

$$\{\mathbf{x} = (x_1, x_2, ..., x_n): \\ a + n\varepsilon \leqslant x_1 + (n-1)\,\varepsilon \leqslant \cdots \leqslant x_{n-1} + \varepsilon \leqslant x_n \leqslant b - \varepsilon\}$$

for some $\varepsilon > 0$ sufficiently small. ∎

LEMMA 2. *Let* (1.2) *hold. If* $\mathbf{y} \in X$ *satisfies* (1.3) *then* (1.4) *is valid.*

*Proof.* We introduce a new argument $\xi$ and a new function $f_0(\xi, \mathbf{x}) = \xi$. Let us consider the programming problem:

$$\min_{\xi \in \mathbb{R},\, \mathbf{x} \in X} f_0(\xi, \mathbf{x}) \tag{2.2}$$

subject to

$$f_0(\xi, \mathbf{x}) - f_i(\mathbf{x}) \geqslant 0, \qquad i = 1, 2, ..., n+1.$$

It is not hard to see that if (1.3) holds then $(f(\mathbf{y}), \mathbf{y})$ is a solution of the above-mentioned problem. Applying Theorem A one can find a vector $\lambda = (\lambda_0, \lambda_1, ..., \lambda_{n+1}) \neq 0$, $\lambda \geqslant 0$, such that

$$\lambda_0 \nabla f_0(f(\mathbf{y}), \mathbf{y}) - \sum_{i=1}^{n+1} \lambda_i \nabla[f_0(f(\mathbf{y}), \mathbf{y}) - f_i(\mathbf{y})] = 0, \tag{2.3}$$

$$\lambda_i[f_0(f(\mathbf{y}), \mathbf{y}) - f_i(\mathbf{y})] = 0, \qquad i = 1, 2, ..., n+1. \tag{2.4}$$

Equation (2.3) becomes

$$\lambda_0 - \sum_{i=1}^{n+1} \lambda_i = 0$$

and (2.4) becomes

$$\sum_{i=1}^{n+1} \lambda_i \frac{\partial f_i(\mathbf{y})}{\partial x_j} = 0, \qquad j = 1, 2, ..., n. \tag{2.5}$$

Equation (2.5) by (1.2) implies, remembering $\lambda \geqslant 0$,

$$\lambda_i > 0, \qquad i = 1, 2, ..., n+1, \tag{2.6}$$

and hence (1.4) follows from (2.4). ▌

LEMMA 3. *Let* (1.1) *and* (1.2) *hold. Then*

$$D(\mathbf{x}) := \det \left( \frac{\partial [f_{i+1}(\mathbf{x}) - f_i(\mathbf{x})]}{\partial x_j} \right)_{i,\, j=1}^{n} \neq 0, \qquad \mathbf{x} \in X. \tag{2.7}$$

*Proof.* Although (2.7) is verified in the proof of [3, Lemma 3] only for the specific case of the optimal Lagrange interpolation, it remains true for the present general minimax problem. But, for completeness, we now give another proof. There must exist a nonzero vector $(\lambda_1, (\mathbf{x}), ..., \lambda_{n+1}(\mathbf{x}))$ with at least one positive component that satisfies

$$\sum_{i=1}^{n+1} \lambda_i(\mathbf{x}) \frac{\partial f_i(\mathbf{x})}{\partial x_j} = 0, \qquad j = 1, 2, ..., n. \tag{2.8}$$

Then by (1.2) we can conclude that $\lambda_i(\mathbf{x}) \neq 0$ for all $i$. Comparing (2.8) with (2.5) and (2.6) according to continuity of $\partial f_i(\mathbf{x})/\partial x_j$ and connectivity of $X$ we can further conclude $\lambda_i(\mathbf{x}) > 0$ for all $i$. Thus for a fixed $\mathbf{x} \in X$ the system of homogeneous linear equations with unknowns $\lambda_1, ..., \lambda_{n+1}$

$$\sum_{i=1}^{n+1} \lambda_i = 0,$$

$$\sum_{i=1}^{n+1} \lambda_i \frac{\partial f_i(\mathbf{x})}{\partial x_j} = 0, \qquad j = 1, 2, ..., n$$

has only the trivial solution, for otherwise $\lambda_i \lambda_k < 0$ would occur for some indices $i$ and $k$, a contradiction. Therefore, the determinant of its coefficient matrix is nonzero:

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \dfrac{\partial f_1(\mathbf{x})}{\partial x_1} & \dfrac{\partial f_2(\mathbf{x})}{\partial x_1} & \cdots & \dfrac{\partial f_{n+1}(\mathbf{x})}{\partial x_1} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial f_1(\mathbf{x})}{\partial x_n} & \dfrac{\partial f_2(\mathbf{x})}{\partial x_n} & \cdots & \dfrac{\partial f_{n+1}(\mathbf{x})}{\partial x_n} \end{vmatrix} \neq 0.$$

Subtracting each column from its successor, proceeding from right to left, and expanding the resulting determinant along the first row just gives (2.7). ∎

## 3. PROOFS OF THEOREMS

3.1. *Proof of Theorem* 2. (a) Equation (2.7) implies that the map $F$ is a local homeomorphism (the definition of which can be found, say, in [13, p. 103]). It, coupled with (1.1), means that $F$ is a homeomorphism by a well-known result (see [3, 4]).

(b) Statement (b) is given by the proof of [3, Theorem 2]. But we now provide a simpler proof using a somewhat different approach. If $f_i(\mathbf{u}) = f_i(\mathbf{v})$ for all $i$ then $\mathbf{u} = \mathbf{v}$ by Statement (a). Now suppose that $f_k(\mathbf{u}) < f_k(\mathbf{v})$ for some $k$, $1 \leqslant k \leqslant n+1$. Let us consider the programming problem:

$$\min_{\mathbf{x} \in X} f_k(\mathbf{x}) \tag{3.1}$$

subject to

$$f_i(\mathbf{x}) \leqslant f_i(\mathbf{v}) - \alpha, \qquad i \in N := \{1, ..., k-1, k+1, ..., n+1\}. \tag{3.2}$$

Denote by $A$ the set of $\alpha$ for which the feasible set of this programming problem is nonempty. Since the feasible set for $\alpha \leqslant 0$ contains $\mathbf{u}$, we have $(-\infty, 0] \subset A$. Meanwhile (1.2) implies that $A$ is open. Thus $\bar{\alpha} := \sup_{\alpha \in A} \alpha > 0$, $\bar{\alpha} \notin A$, and by (2.1) for each $\alpha < \bar{\alpha}$ one can find a solution $\mathbf{w}_\alpha \in X$. By Theorem A there exists a vector $\lambda = (\lambda_1, \lambda_2, ..., \lambda_{n+1}) \neq 0$, $\lambda \geqslant 0$, such that

$$\lambda_k \nabla f_k(\mathbf{w}_\alpha) - \sum_{i \in N} \lambda_i \nabla [f_i(\mathbf{v}) - \alpha - f_i(\mathbf{x})]|_{\mathbf{x} = \mathbf{w}_\alpha} = 0, \tag{3.3}$$

$$\lambda_i [f_i(\mathbf{v}) - \alpha - f_i(\mathbf{w}_\alpha)] = 0, \qquad i \in N. \tag{3.4}$$

Since each $f_i(\mathbf{v})$ is just a constant, (3.3) yields

$$\sum_{i=1}^{n+1} \lambda_i \frac{\partial f_i(\mathbf{w}_\alpha)}{\partial x_j} = 0, \qquad j = 1, 2, ..., n, \tag{3.5}$$

which by (1.2) implies $\lambda_i > 0$ for all $i$ (since $\lambda \neq 0$ and $\lambda \geqslant 0$). Then by (3.4) we get

$$f_i(\mathbf{w}_\alpha) = f_i(\mathbf{v}) - \alpha, \qquad i \in N. \tag{3.6}$$

By (1.2), it follows that $\mathbf{w}_\alpha$ is uniquely determined by (3.6), hence, by the Implicit Function Theorem, $\mathbf{w}_\alpha$ is a continuous function of $\alpha$. Meanwhile, since $\mathbf{w}_\alpha$ is a solution of the programming problem (3.1) subject to (3.2) and the feasible set corresponding to $\alpha_1$ contains the one corresponding to $\alpha_2$ if $\alpha_1 < \alpha_2$, we can conclude that $f_k(\mathbf{w}_\alpha)$ is increasing on $[0, \bar{\alpha})$ from $f_k(\mathbf{w}_0)$ to $\infty$ (because $\sup_{\alpha < \bar{\alpha}} f_k(\mathbf{w}_\alpha) < \infty$ can be shown to imply that $\bar{\alpha} \in A$, contradicting the previous conclusion $\bar{\alpha} \notin A$). Clearly, from the definition

$$f_k(\mathbf{w}_0) \leqslant f_k(\mathbf{u}) < f_k(\mathbf{v}). \tag{3.7}$$

Therefore there is an $\alpha > 0$ such that

$$f_k(\mathbf{w}_\alpha) = f_k(\mathbf{v}) - \alpha, \tag{3.8}$$

which, coupled with (3.6), implies

$$f_{i+1}(\mathbf{w}_\alpha) - f_i(\mathbf{w}_\alpha) = f_{i+1}(\mathbf{v}) - f_i(\mathbf{v}), \qquad i = 1, 2, ..., n.$$

By statement (a) we have $\mathbf{w}_\alpha = \mathbf{v}$, contradicting (3.8).

(c) If $F_k(\mathbf{u}) = F_k(\mathbf{v})$ then either $f_i(\mathbf{u}) \leqslant f_i(\mathbf{v})$ for all $i$ or $f_i(\mathbf{u}) \geqslant f_i(\mathbf{v})$ for all $i$. Hence $\mathbf{u} = \mathbf{v}$ by statement (b). ∎

3.2. *Proof of Theorem* 1. Statements (a)–(c) are given by Lemmas 1 and 2 as well as Theorem 2. The remainder of the proof is devoted to showing Statement (d). Clearly, in the case when $\mathbf{w} = \mathbf{y}$, (1.6) by (1.4) becomes (1.7), where $\gamma^* = \gamma(\mathbf{y})$. Now let us show (1.6).

The conclusion is trivial for $\mathbf{x} = \mathbf{w}$. Now let $\mathbf{x} \in X \backslash \{\mathbf{w}\}$. In this case put

$$d_i(\mathbf{x}) := \frac{f_i(\mathbf{x}) - f_i(\mathbf{w})}{\|\mathbf{x} - \mathbf{w}\|}, \qquad i = 1, 2, ..., n+1$$

and

$$d(\mathbf{x}) := \max_{1 \leqslant i \leqslant n+1} d_i(\mathbf{x}).$$

It suffices to show $d(\mathbf{x}) \geqslant \gamma > 0$. Suppose to the contrary that there would be a sequence $\mathbf{x}^{(m)} \in X \backslash \{\mathbf{w}\}$ such that $d(\mathbf{x}^{(m)}) \to 0$ as $m \to \infty$. We claim that $\mathbf{x}^{(m)} \to \mathbf{w}$ as $m \to \infty$. In fact, let $\mathbf{x}^{(m_k)}$ be an arbitrary convergent subsequence of $\mathbf{x}^{(m)}$: $\mathbf{x}^{(m_k)} \to \mathbf{z}$ $(k \to \infty)$. From the definition we have

$$f_i(\mathbf{x}^{(m_k)}) - f_i(\mathbf{w}) \leqslant d(\mathbf{x}^{(m_k)}) \, \|\mathbf{x}^{(m_k)} - \mathbf{w}\|, \qquad i = 1, 2, ..., n+1.$$

Then by the assumption that $d(\mathbf{x}^{(m)}) \to 0$ as $m \to \infty$, $f_i(\mathbf{x}^{(m_k)})$ must be bounded. By virtue of (2.1) we conclude $\mathbf{z} \in X$ and hence

$$f_i(\mathbf{z}) - f_i(\mathbf{w}) \leqslant 0, \qquad i = 1, 2, ..., n+1.$$

Applying Theorem 2 gives $\mathbf{z} = \mathbf{w}$. This proves that $\mathbf{x}^{(m)} \to \mathbf{w}$ as $m \to \infty$. Using the differential expression for multivariate functions we have

$$f_i(\mathbf{x}^{(m)}) - f_i(\mathbf{w}) = \nabla f_i(\mathbf{w})(\mathbf{x}^{(m)} - \mathbf{w}) + o(\|\mathbf{x}^{(m)} - \mathbf{w}\|), \qquad i = 1, 2, ..., n+1,$$

or equivalently,

$$d_i(\mathbf{x}^{(m)}) = \frac{\nabla f_i(\mathbf{w})(\mathbf{x}^{(m)} - \mathbf{w})}{\|\mathbf{x}^{(m)} - \mathbf{w}\|} + o(1), \qquad i = 1, 2, ..., n+1. \tag{3.9}$$

Assume without loss of generality (passing to a subsequence if necessary) that all the following limits exist:

$$\mu_i := \lim_{n \to \infty} d_i(\mathbf{x}^{(m)}), \qquad i = 1, 2, ..., n+1,$$

$$v_j := \lim_{m \to \infty} \frac{x_j^{(m)} - w_j}{\|\mathbf{x}^{(m)} - \mathbf{w}\|}, \qquad j = 1, 2, ..., n.$$

Then (3.9) yields

$$\mu_i = \sum_{j=1}^{n} v_j \frac{\partial f_i(\mathbf{w})}{\partial x_j}, \qquad i = 1, 2, ..., n+1. \tag{3.10}$$

We claim that

$$\mu_i = 0, \qquad i = 1, 2, ..., n+1. \tag{3.11}$$

In fact, multiplication of (2.8) with $v_j$ and summation (after replacing $\mathbf{x}$ by $\mathbf{w}$) yield

$$\sum_{i=1}^{n+1} \lambda_i(\mathbf{w}) \sum_{j=1}^{n} v_j \frac{\partial f_i(\mathbf{w})}{\partial x_j} = 0.$$

Substituting (3.10) into the above equation gives

$$\sum_{i=1}^{n+1} \lambda_i(\mathbf{w})\, \mu_i = 0. \tag{3.12}$$

But from the definition $d_i(\mathbf{x}^{(m)}) \leqslant d(\mathbf{x}^{(m)})$, so $\mu_i \leqslant \lim_{m \to \infty} d(\mathbf{x}^{(m)}) = 0$ for all $i$. Remembering $\lambda_i(\mathbf{w}) > 0$ for all $i$, (3.12) implies (3.11). Therefore (3.10) becomes

$$\sum_{j=1}^{n} v_j \frac{\partial f_i(\mathbf{w})}{\partial x_j} = 0, \qquad i = 1, 2, ..., n+1.$$

Since $(v_1, v_2, ..., v_n) \neq 0$, we conclude that $D_k(\mathbf{w}) = 0$ for all $k$. This contradicts (1.2) and proves (1.6). ▮

## ACKNOWLEDGMENT

The author thanks the referees for pointing out the mistakes in earlier versions of the manuscript and for many valuable suggestions and comments.

## REFERENCES

1. M. Avriel, "Nonlinear Programming: Analysis and Methods," Prentice Hall, 1979.
2. S. Bernstein, Sur la limitation des valeurs d'un polynôme $P_n(x)$ de degré $n$ sur tout un segment par ses valeurs en $n+1$ points du segment, *Izv. Akad. Nauk SSSR* **7** (1931), 1025–1050.
3. C. de Boor and A. Pinkus, Proof of the conjectures of Bernstein and Erdős concerning the optimal nodes for polynomial interpolation, *J. Approx. Theory* **24** (1978), 289–303.
4. F. E. Browder, Covering spaces, fibre spaces, and local homeomorphisms, *Duke Math. J.* **21** (1954), 329–336.
5. L. Brutman and A. Pinkus, On the Erdős conjecture concerning minimal norm interpolation on the unit circle, *SIAM J. Numer. Anal.* **17** (1980), 373–375.
6. E. W. Cheney, "Introduction to Approximation Theory," McGraw–Hill, New York, 1966.
7. P. Erdős, Some remarks on polynomials, *Bull. Amer. Math. Soc.* (*N.S.*) **53** (1947), 1169–1176.
8. T. A. Kilgore, A characterization of the Lagrange interpolating projection with minimal Tchebycheff norm, *J. Approx. Theory* **24** (1978), 273–288.
9. T. A. Kilgore, Optimal interpolation with incomplete polynomials, *J. Approx. Theory* **41** (1984), 279–290.
10. T. A. Kilgore, Optimal interpolation with polynomials having fixed roots, *J. Approx. Theory* **49** (1987), 378–389.
11. T. A. Kilgore, Optimal interpolation with exponentially weighted polynomials on an unbounded interval, *Acta Math. Hungar.* **57**, Nos. 1, 2 (1991), 85–90.
12. Y. G. Shi, Optimal Lebesgue function type sum, *Acta Math. Hungar.* **77** (1997), 287–300.
13. J. Szabados and P. Vértesi, "Interpolation of Functions," World Scientific, Singapore, 1990.