On a Best Estimator for the Class $\mathfrak{M}^r$ Using only Function Values

CHARLES A. MICCHELLI & ALLAN PINKUS

1. Introduction. Let $BV[0, 1]$ denote the set of all real-valued functions of bounded variation on $[0, 1]$. For $r \geqq 2$, we define the class of r-fold integrals of elements in $BV[0, 1]$ by

$$\mathfrak{M}^r = \mathfrak{M}^r[0, 1] = \left\{f : f(x) = \sum_{i=1}^{r} a_i x^{i-1} + \frac{1}{(r-1)!} \int_0^1 (x - t)_+^{r-1} \, d\lambda_f(t), \right.$$
$$\left. \lambda_f \in BV[0, 1], (a_1, \cdots, a_r) \in \mathbf{R}^r \right\}.$$

Fix $\mathbf{x} = (x_1, \cdots, x_n)$, $0 < x_1 < \cdots < x_n < 1$, $n \geqq r$, and let $\mathbf{f}$ denote the vector $(f(x_1), \cdots, f(x_n))$. In this paper, we are concerned with best methods of estimating a function $f \in \mathfrak{M}^r$, given its function values $\mathbf{f}$. Without any additional prior information on f, arbitrarily large errors in estimating f may be encountered. It is the purpose of this paper to describe an optimal estimator for f when a bound on the total variation $||\lambda_f||$ of λ_f on $[0, 1]$ is available.

The specific problem we will treat is formulated as follows:

Any mapping T (not necessarily linear) from $\mathbf{R}^n$ into $\mathfrak{M}^r$ determines an estimate $Sf = T\mathbf{f}$ for f. The (relative) error for this estimate S, given only that $f \in \mathfrak{M}^r$, is defined as

$$E(\mathbf{x}; S) = \sup_{f \in \mathfrak{M}^r} \frac{||f - Sf||_1}{||\lambda_f||},$$

where $||\cdot||_1 = L^1$-norm on $[0, 1]$. We call S_0 an optimal estimator provided that

$$E(\mathbf{x}) = \inf_S E(\mathbf{x}; S) = E(\mathbf{x}; S_0),$$

where the infimum is taken over all estimators S, as defined above.

Our main result depends on the following lemma which we prove in Section 2. Before stating this lemma, we require the following definitions.

Let $\mathcal{S}_{r,n} = \mathcal{S}_{r,n}(\mathbf{x})$ denote the class of spline functions of order r with the n knots $\mathbf{x} = (x_1, \cdots, x_n)$, that is,

$$\mathcal{S}_{r,n} = \left\{s : s(x) = \sum_{i=1}^{r} a_i x^{i-1} + \sum_{i=1}^{n} a_{i+r}(x - x_i)_+^{r-1}, (a_1, \cdots, a_{n+r}) \in \mathbf{R}^{n+r}\right\},$$

Indiana University Mathematics Journal ©, Vol. 26, No. 4 (1977)

and, in addition, for $n \geqq r$ define

$$\mathcal{G}_{r,n} = \left\{ g : g(x) = \int_0^1 (x-t)_+^{r-1} h_{\mathbf{x}}(t)\, dt + s(x), \right.$$

$$\left. s \in \mathcal{S}_{r,n}, \; g^{(i)}(0) = g^{(i)}(1) = 0, i = 0, 1, \cdots, r-1 \right\},$$

where $h_{\mathbf{x}}(t) = (-1)^j, x_j < t \leqq x_{j+1}, j = 0, 1, \cdots, n; x_0 = 0, x_{n+1} = 1$.

Lemma 1.1. *There exists a $g_{\mathbf{x}} \in \mathcal{G}_{r,n}$ which equioscillates $n - r + 1$ times on* $[0, 1]$, *i.e.,*

$$g_{\mathbf{x}}(\eta_i) = \sigma(-1)^i \, ||g_{\mathbf{x}}||_\infty, \qquad i = 1, \cdots, n - r + 1,$$

where $0 < \eta_1 < \cdots < \eta_{n-r+1} < 1$, $\sigma^2 = 1$, and $||\cdot||_\infty =$ sup norm on $[0, 1]$. Furthermore,

$$||g_{\mathbf{x}}||_\infty = \min_{g \in \mathcal{G}_{r,n}} ||g||_\infty,$$

and $g_{\mathbf{x}}$ has exactly $n - r$ zeros $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_{n-r})$ contained in $(0, 1)$.

Let $S_{\mathbf{x}}$ denote the (linear) estimator which interpolates f at $\mathbf{x}$ by a spline function of order r with $n - r$ knots $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_{n-r})$, i.e., $S_{\mathbf{x}} f \in \mathcal{S}_{r,n-r}(\boldsymbol{\xi})$ and $(S_{\mathbf{x}} f)(x_i) = f(x_i)$, $i = 1, \cdots, n$. We shall prove (Lemma 2.3) that such an estimator exists. In addition, let $c(\mathbf{x}) = \sup \{||f||_1 : f(x_i) = 0, i = 1, \cdots, n, f \in \mathcal{B}_r\}$, where $\mathcal{B}_r = \{f : f \in \mathfrak{M}^r, ||\lambda_f|| \leqq 1\}$.

Theorem 1.1. *For fixed $\mathbf{x} = (x_1, \cdots, x_n)$, $0 < x_1 < \cdots < x_n < 1$, and $f \in \mathfrak{M}^r$, there exists a unique $S_{\mathbf{x}} f \in \mathcal{S}_{r,n-r}(\boldsymbol{\xi})$ such that*

$$(S_{\mathbf{x}} f)(x_i) = f(x_i), \qquad i = 1, \cdots, n,$$

and $S_{\mathbf{x}}$ is an optimal (linear) estimator for f. Furthermore,

$$E(\mathbf{x}; S_{\mathbf{x}}) = E(\mathbf{x}) = c(\mathbf{x}) = ||g_{\mathbf{x}}||_\infty .$$

This theorem will be proven in Section 2 in somewhat greater generality.

Finally, using results on the n-widths of $\mathcal{B}_r$ in $L^1[0, 1]$ obtained in [2], we are able to determine (in Section 3) a choice of $\mathbf{x}$ which minimizes $E(\mathbf{x})$.

The results on optimal estimation which we herein obtain may be viewed as the L^1-analogue of results in [3]. In that paper, an L^∞ bound on $f^{(r)}$ gave rise to a different optimal (spline) estimator for f.

2. An optimal estimator.

We shall make use of the notation

$$(2.1) \qquad k_i(x) = x^{i-1}, \qquad i = 1, \cdots, r, \quad \text{and} \quad K(x, t) = (x - t)_+^{r-1},$$

where $x_+^{r-1} = x^{r-1}$ for $x \geqq 0$, and zero otherwise.

Below we list properties of these functions required in the subsequent analysis.

I. $\{k_1(x), \cdots, k_r(x)\}$ forms a Chebyshev system on $[0, 1]$, that is,

$$K\begin{pmatrix}1, \cdots, r\\ y_1, \cdots, y_r\end{pmatrix} = \det_{i,j=1,\cdots,r} (k_i(y_j)) > 0,$$

for all $y_1, \cdots, y_r$, $0 \leqq y_1 < \cdots < y_r \leqq 1$.

II. For every $m \geqq 1$, and $s_1, \cdots, s_m$ with $0 < s_1 < \cdots < s_m < 1$, $\{k_1(x), \cdots, k_r(x), K(x, s_1), \cdots, K(x, s_m)\}$ forms a weak Chebyshev system of dimension $r + m$ on $[0, 1]$, and for some choice of $s_1, \cdots, s_{n-r}$, $0 < s_1 < \cdots < s_{n-r} < 1$, the linear map $f \to \mathbf{f} = (f(x_1), \cdots, f(x_n))$ is one-to-one on the space spanned by the functions $k_1(x), \cdots, k_r(x), K(x, s_1), \cdots, K(x, s_{n-r})$.

Our second property entails that the determinant

$$K\begin{pmatrix}1, \cdots, r, s_1, \cdots, s_m\\ t_1, \cdots, t_{r+m}\end{pmatrix} = \begin{vmatrix} k_1(t_1) & \cdots & k_1(t_{r+m})\\ \vdots & & \vdots\\ k_r(t_1) & \cdots & k_r(t_{r+m})\\ K(t_1, s_1) & \cdots & K(t_{r+m}, s_1)\\ \vdots & & \vdots\\ K(t_1, s_m) & \cdots & K(t_{r+m}, s_m)\end{vmatrix}$$

is nonnegative for all $0 < s_1 < \cdots < s_m < 1$ and $0 \leqq t_1 < \cdots < t_{r+m} \leqq 1$, the functions $k_1(x), \cdots, k_r(x), K(x, s_1), \cdots, K(x, s_m)$ are linearly independent for each choice of $0 < s_1 < \cdots < s_m < 1$, and there is some choice of $n - r$ points, $0 < s_1 < \cdots < s_{n-r} < 1$, for which

$$K\begin{pmatrix}1, \cdots, r, s_1, \cdots, s_{n-r}\\ x_1, \cdots, x_n\end{pmatrix}$$

is strictly positive.

The fact that these properties hold for functions of the form (2.1) is a fundamental result from the theory of spline functions, see [4]. In fact, much more is known. The Schoenberg–Whitney Theorem [4], tells us exactly when the above determinants are positive. However, this information will not be required in our proofs. Rather, properties I and II will suffice for our purposes. We will now let $k_1(x), \cdots, k_r(x), K(x, t)$ *be any continuous functions which satisfy properties I and II*, and consider the problem of estimating the class

$$\mathfrak{M}^r = \Big\{ f : f(x) = \sum_{i=1}^{r} a_i k_i(x) + \int_0^1 K(x, t)\, d\lambda(t),$$
$$\lambda \in BV[0, 1], (a_1, \cdots, a_r) \in \mathbf{R}^r \Big\}.$$

using only the function values $\mathbf{f} = (f(x_1), \cdots, f(x_n))$.

We shall also have occasion to use the auxiliary kernel

$$J(x, t) = J_r(x, t; \mathbf{x}) = \frac{K\begin{pmatrix} 1, \cdots, r, t \\ x_1, \cdots, x_r, x \end{pmatrix}}{K\begin{pmatrix} 1, \cdots, r \\ x_1, \cdots, x_r \end{pmatrix}}.$$

For fixed $\mathbf{x} = (x_1, \cdots, x_n)$, $x_0 = 0 < x_1 < \cdots < x_n < x_{n+1} = 1$, let $h_\mathbf{x}(x) = (-1)^j$, $x_j < x \leqq x_{j+1}$, $j = 0, 1, \cdots, n$, and set

$$\mathcal{G} = \left\{ g : g(t) = \int_0^1 J(x, t)h_\mathbf{x}(x)\, dx - \sum_{j=1}^{n-r} c_j J(x_{r+j}, t), (c_1, \cdots, c_{n-r}) \in \mathbf{R}^{n-r} \right\}.$$

Lemma 2.1. *Every $g \in \mathcal{G}$ has at most $n - r$ distinct zeros in* $(0, 1)$.

Proof. Suppose to the contrary that there exists a $g \in \mathcal{G}$ with $n - r + 1$ zeros $\{\zeta_i\}_{i=1}^{n-r+1}$ in $(0, 1)$, $0 < \zeta_1 < \cdots < \zeta_{n-r+1} < 1$. Since the functions $\{k_1(x), \cdots, k_r(x), K(x, \zeta_1), \cdots, K(x, \zeta_{n-r+1})\}$ form a weak Chebyshev system, there exists a non-trivial function

$$u(x) = \sum_{j=1}^{r} a_j k_j(x) + \sum_{j=1}^{n-r+1} b_j K(x, \zeta_j)$$

which exhibits (weak) sign changes at $x_1, \cdots, x_n$, *i.e.*, $(-1)^j u(x) \geqq 0$, $x_j \leqq x \leqq x_{j+1}$, $j = 0, 1, \cdots, n$. Since

$$u(x) = \frac{1}{K\begin{pmatrix} 1, \cdots, r \\ x_1, \cdots, x_r \end{pmatrix}} \begin{vmatrix} k_1(x_1) & \cdots & k_1(x_r) & k_1(x) \\ \vdots & & \vdots & \vdots \\ k_r(x_1) & \cdots & k_r(x_r) & k_r(x) \\ u(x_1) & \cdots & u(x_r) & u(x) \end{vmatrix}$$

$$= \sum_{j=1}^{n-r+1} b_j J(x, \zeta_j),$$

we have

$$\int_0^1 |u(x)|\, dx = \sum_{j=0}^{n} (-1)^j \int_{x_j}^{x_{j+1}} u(x)\, dx - \sum_{j=1}^{n-r} c_j u(x_{r+j})$$

$$= \sum_{l=1}^{n-r+1} b_l \left[\sum_{j=0}^{n} (-1)^j \int_{x_j}^{x_{j+1}} J(x, \zeta_l)\, dx - \sum_{j=1}^{n-r} c_j J(x_{r+j}, \zeta_l) \right]$$

$$= \sum_{l=1}^{n-r+1} b_l g(\zeta_l) = 0.$$

This contradiction proves the lemma.

Lemma 2.2. *There exists a function $g_\mathbf{x} \in \mathcal{G}$ which equioscillates at $n - r + 1$ points $0 \leqq \eta_1 < \cdots < \eta_{n-r+1} \leqq 1$, that is,*

$$g_\mathbf{x}(\eta_i) = \sigma(-1)^i \|g_\mathbf{x}\|_\infty, \quad \sigma^2 = 1, \quad i = 1, \cdots, n - r + 1 \tag{2.2}$$

and

$$||g_x||_\infty = \min_{g \in \mathcal{G}} ||g||_\infty .$$

Proof. Let $v_i(t) = J(x_{r+i}, t), i = 1, \cdots, n-r$. Then according to Sylvester's determinant identity,

$$\det_{i,j=1,\cdots,n-r} (v_i(t_j)) = \frac{K\begin{pmatrix} 1, \cdots, r, t_1, \cdots, t_{n-r} \\ x_1, \cdots, x_n \end{pmatrix}}{K\begin{pmatrix} 1, \cdots, r \\ x_1, \cdots, x_r \end{pmatrix}} .$$

Hence properties I and II imply that $\{v_1(t), \cdots, v_{n-r}(t)\}$ is a weak Chebyshev subspace of dimension $n - r$. Thus, by a result of Jones and Karlovitz [1], the function $\int_0^1 J(x, t)h_x(x)\,dx$ has a best Chebyshev approximation from the subspace spanned by $v_1(t), \cdots, v_{n-r}(t)$ which equioscillates at at least $n - r + 1$ points. The lemma is proven.

Furthermore, the method of proof of Jones and Karlovitz [1] ("smoothing" a weak Chebyshev system into a Chebyshev system) may easily be applied to prove the existence of constants $\lambda_1, \cdots, \lambda_{n-r+1}$,

$$\sum_{j=1}^{n-r+1} |\lambda_j| = 1, \qquad \lambda_j(-1)^j \geqq 0, \qquad j = 1, \cdots, n-r+1,$$

such that

$$\sum_{j=1}^{n-r+1} \lambda_j v_i(\eta_j) = \sum_{j=1}^{n-r+1} \lambda_j J(x_{r+i}, \eta_j) = 0, \qquad i = 1, \cdots, n-r \tag{2.3}$$

and

$$\left|\sum_{j=1}^{n-r+1} \lambda_j g_x(\eta_j)\right| = ||g_x||_\infty .$$

Define

$$f_x(x) = \sum_{j=1}^{n-r+1} \lambda_j J(x, \eta_j). \tag{2.4}$$

Thus,

$$f_x(x_j) = 0, \qquad j = 1, \cdots, n, \tag{2.5}$$

and

$$||\lambda_{f_x}|| = \sum_{j=1}^{n-r+1} |\lambda_j| = 1.$$

The usefulness of f_x is in the fact that it solves the dual maximum problem associated with the best approximation of $\int_0^1 J(x, t)h_x(x)\,dx$. Combining (2.2), (2.3), (2.4) and the definition of g_x gives

$$\begin{aligned}(h_{\mathbf{x}}, f_{\mathbf{x}}) &= \int_0^1 h_{\mathbf{x}}(x) f_{\mathbf{x}}(x)\,dx \\ &= \sum_{i=1}^{n-r+1} \lambda_i \int_0^1 h_{\mathbf{x}}(x) J(x, \eta_i)\,dx \\ &= \sum_{i=1}^{n-r+1} \lambda_i \left[\int_0^1 h_{\mathbf{x}}(x) J(x, \eta_i)\,dx - \sum_{l=1}^{n-r} d_l J(x_{r+l}, \eta_i) \right] \\ &= \sum_{i=1}^{n-r+1} \lambda_i g_{\mathbf{x}}(\eta_i) \\ &= \pm \|g_{\mathbf{x}}\|_\infty .\end{aligned}$$

Furthermore, by the well-known duality formula for the distance of a function to a subspace,

$$\begin{aligned}|(h_{\mathbf{x}}, f_{\mathbf{x}})| &= \|g_{\mathbf{x}}\|_\infty \\ &= \max_{\|\lambda\| \le 1} \int_0^1 \left(\int_0^1 J(x, t) h_{\mathbf{x}}(x)\,dx \right) d\lambda(t)\end{aligned} \tag{2.6}$$

where

$$\int_0^1 J(x_i, t)\,d\lambda(t) = 0, \qquad i = r+1, \cdots, n.$$

Let us note that specializing Lemmas 2.1 and 2.2 to the choice (2.1) yields Lemma 1.1.

Lemma 2.3. *The function $g_{\mathbf{x}}$ constructed in Lemma 2.2 has exactly $n - r$ distinct zeros in $(0, 1)$ at $0 < \xi_1 < \cdots < \xi_{n-r} < 1$, and*

$$K\begin{pmatrix} 1, \cdots, r, \xi_1, \cdots, \xi_{n-r} \\ x_1, \cdots, x_n \end{pmatrix} > 0. \tag{2.7}$$

Proof. From Lemma 2.1, $g_{\mathbf{x}}$ has exactly $n - r$ distinct zeros in $(0, 1)$. Assume that (2.7) is false. Hence there exists a non-trivial function

$$s(t) = \sum_{i=1}^{n-r} c_i J(x_{r+i}, t)$$

which vanishes at $\xi_1, \cdots, \xi_{n-r}$. Since $J(x_{r+1}, t), \cdots, J(x_n, t)$ are linearly independent (property II), there is a $z \in (0, 1)\backslash\{\xi_1, \cdots, \xi_{n-r}\}$ with $s(z) \neq 0$. Thus we may choose a constant c such that the function $g_{\mathbf{x}}(t) - cs(t)$ vanishes $n - r + 1$ times at $z, \xi_1, \cdots, \xi_{n-r}$. However $g_{\mathbf{x}} - cs \in \mathcal{G}$, and this conclusion contradicts Lemma 2.1. The proof is complete.

Now, from Lemma 2.3, we obtain that for every vector $\mathbf{y} = (y_1, \cdots, y_n) \in \mathbf{R}^n$ there exists a unique function $T\mathbf{y}$ in

$$\mathcal{S}_{r,n-r}(\xi) = \left\{ s \,.\, s(x) = \sum_{i=1}^{r} a_i k_i(x) + \sum_{j=1}^{n-r} b_j K(x, \xi_j) \right\}$$

where $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_{n-r})$, as constructed in Lemma 2.3, such that $(T\mathbf{y})(x_i) = y_i$, $i = 1, \cdots, n$. Set $S_\mathbf{x} f = T\mathbf{f}$. Thus

$$(2.8) \qquad f(x) - (S_\mathbf{x} f)(x) = \int_0^1 \frac{K\begin{pmatrix} 1, \cdots, r, \xi_1, \cdots, \xi_{n-r}, t \\ x_1, \cdots, x_n, x \end{pmatrix}}{K\begin{pmatrix} 1, \cdots, r, \xi_1, \cdots, \xi_{n-r} \\ x_1, \cdots, x_n \end{pmatrix}} d\lambda_f(t).$$

We are now prepared to prove our main theorem.

Theorem 2.1. *$S_\mathbf{x}$ is an optimal (linear) estimator, and*

$$E(\mathbf{x}; S_\mathbf{x}) = c(\mathbf{x}) = \|g_\mathbf{x}\|_\infty = \|f_\mathbf{x}\|_1 = E(\mathbf{x}).$$

Proof. Recall that $c(\mathbf{x}) = \sup \{\|f\|_1 : f(x_i) = 0, i = 1, \cdots, n, f \in \mathcal{B}_r\}$, where $\mathcal{B}_r = \{f : f \in \mathfrak{M}^r, \|\lambda_f\| \leqq 1\}$, and note that $f \in \mathcal{B}_r$, $f(x_i) = 0$, $i = 1, \cdots, n$, if and only if $f(x) = \int_0^1 J(x, t)\, d\lambda_f(t)$, $\|\lambda_f\| \leqq 1$, and $f(x_j) = 0$, $j = r + 1, \cdots, n$. Thus,

$$c(\mathbf{x}) \geqq \sup_{f \in \mathcal{B}_r} \{(h_\mathbf{x}, f) : f(x_j) = 0, j = 1, \cdots, n\}$$

$$= \max_{\|\lambda\| \leq 1} \left\{\int_0^1 \left(\int_0^1 J(x, t) h_\mathbf{x}(x)\, dx\right) d\lambda(t) : \int_0^1 J(x_i, t)\, d\lambda(t) = 0, \right.$$

$$\left. i = r + 1, \cdots, n\right\}$$

$$= \|g_\mathbf{x}\|_\infty,$$

as a result of (2.6).

To prove $E(\mathbf{x}) \geqq c(\mathbf{x})$, we parallel the argument used in [3]. For any estimator S obtained from a mapping $T : \mathbf{R}^n \to \mathfrak{M}^r$, and for $f \in \mathcal{B}_r$, $f(x_i) = 0$, $i = 1, \cdots, n$, we have

$$E(\mathbf{x}; S) \geqq \|f - T\mathbf{0}\|_1, \quad \text{and} \quad E(\mathbf{x}; S) \geqq \|f + T\mathbf{0}\|_1 .$$

Thus,

$$E(\mathbf{x}; S) \geqq \tfrac{1}{2}[\|f - T\mathbf{0}\|_1 + \|f + T\mathbf{0}\|_1]$$

$$\geqq \|f\|_1 ,$$

and therefore $E(\mathbf{x}) \geqq c(\mathbf{x})$. Hence

$$(2.9) \qquad \|g_\mathbf{x}\|_\infty \leqq c(\mathbf{x}) \leqq E(\mathbf{x}) \leqq E(\mathbf{x}; S_\mathbf{x}).$$

To prove that equality holds throughout, we consider $\|f - S_\mathbf{x} f\|_1$ with $\|\lambda_f\| \leqq 1$. Now, utilizing (2.8) and property II,

$$||f - S_{\mathbf{x}}f||_1 = \max_{\substack{h \in L^\infty[0,1] \\ ||h||_\infty \leq 1}} (h, f - S_{\mathbf{x}}f)$$

$$\leqq \max_{\substack{h \in L^\infty[0,1] \\ ||h||_\infty \leq 1}} \max_{0 \leq t \leq 1} \left| \int_0^1 \frac{K\begin{pmatrix} 1, \cdots, r, \xi_1, \cdots, \xi_{n-r}, t \\ x_1, \cdots, x_n, x \end{pmatrix}}{K\begin{pmatrix} 1, \cdots, r, \xi_1, \cdots, \xi_{n-r} \\ x_1, \cdots, x_n \end{pmatrix}} h(x)\, dx \right|$$

$$= \max_{0 \leq t \leq 1} \left| \int_0^1 \frac{K\begin{pmatrix} 1, \cdots, r, \xi_1, \cdots, \xi_{n-r}, t \\ x_1, \cdots, x_n, x \end{pmatrix}}{K\begin{pmatrix} 1, \cdots, r, \xi_1, \cdots, \xi_{n-r} \\ x_1, \cdots, x_n \end{pmatrix}} h_{\mathbf{x}}(x)\, dx \right| .$$

The function

$$R(t) = g_{\mathbf{x}}(t) - \int_0^1 \frac{K\begin{pmatrix} 1, \cdots, r, \xi_1, \cdots, \xi_{n-r}, t \\ x_1, \cdots, x_n, x \end{pmatrix}}{K\begin{pmatrix} 1, \cdots, r, \xi_1, \cdots, \xi_{n-r} \\ x_1, \cdots, x_n \end{pmatrix}} h_{\mathbf{x}}(x)\, dx$$

vanishes at $\xi_1, \cdots, \xi_{n-r}$ and is in $\mathcal{S}_{r,n-r}(\boldsymbol{\xi})$. Thus, from (2.7), $R(t) \equiv 0$, and $||f - S_{\mathbf{x}}f||_1 \leqq ||g_{\mathbf{x}}||_\infty \, ||\lambda_f||$ for $f \in \mathfrak{M}^r$. Therefore $||g_{\mathbf{x}}||_\infty \geqq E(\mathbf{x}; S_{\mathbf{x}})$, and from (2.9), we obtain

$$|(h_{\mathbf{x}}, f_{\mathbf{x}})| = ||g_{\mathbf{x}}||_\infty = E(\mathbf{x}) = c(\mathbf{x}) = E(\mathbf{x}; S_{\mathbf{x}}).$$

Since $||f_{\mathbf{x}}||_1 \leqq c(\mathbf{x}) = |(h_{\mathbf{x}}, f_{\mathbf{x}})| \leqq ||f_{\mathbf{x}}||_1$, the theorem is proven.

Note that specializing Theorem 2.1 to the choice (2.1) yields Theorem 1.1.

3. A best choice for $\mathbf{x} = (x_1, \cdots, x_n)$. In [2], we prove the existence of an $\mathbf{x}^* = (x_1^*, \cdots, x_n^*)$, $0 < x_1^* < \cdots < x_n^* < 1$, such that

$$g_{\mathbf{x}^*}(t) = \sum_{j=0}^{n} (-1)^j \int_{x_j^*}^{x_{j+1}^*} J_r(x, t; \mathbf{x}^*)\, dx$$

equioscillates at $n - r + 1$ points in [0, 1], and

$$\sum_{j=0}^{n} (-1)^j \int_{x_j^*}^{x_{j+1}^*} k_i(x)\, dx = 0, \qquad i = 1, \cdots, r,$$

where

$$J_r(x, t; \mathbf{x}^*) = \frac{K\begin{pmatrix} 1, \cdots, r, t \\ x_1^*, \cdots, x_r^*, x \end{pmatrix}}{K\begin{pmatrix} 1, \cdots, r \\ x_1^*, \cdots, x_r^* \end{pmatrix}} .$$

Consequently, we may express $g_{\mathbf{x}^*}$ in the form

$$g_{\mathbf{x}^*}(t) = \sum_{i=0}^{n} (-1)^i \int_{x_i^*}^{x_{i+1}^*} K(x, t)\, dx.$$

For this choice of $\mathbf{x}^*$, we also show (in [2]) that

$$||g_{\mathbf{x}^*}||_\infty = d_n(\mathcal{B}_r),$$

where $d_n(\mathcal{B}_r)$ is the n-width of $\mathcal{B}_r$ in $L^1[0, 1]$, defined by

$$d_n(\mathcal{B}_r) = \inf_{X_n} \sup_{f \varepsilon \mathcal{B}_r} \inf_{g \varepsilon X_n} ||f - g||_1 ,$$

where X_n is any n-dimensional subspace of $L^1[0, 1]$.

Theorem 2.1.

$$\inf_{\mathbf{x}} E(\mathbf{x}) = E(\mathbf{x}^*) = d_n(\mathcal{B}_r).$$

Proof.

$$\begin{aligned} E(\mathbf{x}) &= \sup_{f \varepsilon \mathcal{B}_r} ||f - S_{\mathbf{x}} f||_1 , \\ &\geqq d_n(\mathcal{B}_r) = ||g_{\mathbf{x}^*}||_\infty \\ &= \sup_{f \varepsilon \mathcal{B}_r} ||f - S_{\mathbf{x}^*} f||_1 = E(\mathbf{x}^*). \end{aligned}$$

REFERENCES

1. R. C. JONES & L. A. KARLOVITZ, *Equioscillation under nonuniqueness in the approximation of continuous functions*, J. Approx. Theory **3** (1970), 138–145.
2. C. A. MICCHELLI & A. PINKUS, *Total positivity and the exact n-width of certain sets in L^1*, Pacific J. of Math. (to appear).
3. C. A. MICCHELLI, T. J. RIVLIN & S. WINOGRAD, *Optimal recovery of smooth functions*, Numer. Math. **26** (1976), 191–200.
4. I. J. SCHOENBERG & A. WHITNEY, *On Pólya frequency functions III. The positivity of translation determinants with an application to the interpolation problem by spline curves*, Trans. Amer. Math. Soc. **74** (1953), 246–259.

Acknowledgement: We are indebted to Professor C. de Boor for many helpful discussions concerning the results of this paper.

This work was sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and the National Science Foundation under Grant No. MCS 75-17385.

Received April 12, 1976.

MICCHELLI: IBM, T. J. WATSON RESEARCH CENTER, YORKTOWN HEIGHTS, NY 10598
PINKUS: MATH. RES. CENTER, UNIVERSITY OF WISCONSIN, MADISON, WI 53706