The n-Widths of Rank $n + 1$ Kernels

Charles A. Micchelli
Mathematical Sciences Department, I.B.M., Thomas J. Watson Research Center, Yorktown Heights, New York 10598

Allan Pinkus
Department of Mathematics, Technion-Israel Institute of Technology, Haifa, Israel.

Let X and Y be two normed linear spaces of functions on [0,1] and let K: $Y \to X$ be an integral operator whose kernel $K(x,y)$ is of rank $n + 1$. Let $\mathfrak{A}$ denote the image under K of the unit ball in Y. This article is concerned with the exact determination of the n-width of $\mathfrak{A}$ in X. It is also shown that the n-width of $\mathfrak{A}$ in X is equal to the best rank n tensor product approximation of the kernel $K(x,y)$ in the appropriate norm.

1. Introduction

Let X be a normed linear space and $\mathfrak{A}$ a subset of X. The n-width of $\mathfrak{A}$ in X, in the sense of Kolmogorov, is defined as

$$d_n(\mathfrak{A}\,;X) = \inf_{Y_n} \sup_{x\in\mathfrak{A}} \inf_{y\in Y_n} \|x-y\| \tag{1.1}$$

where the infimum is taken with respect to all n-dimensional subspaces Y_n of X. Any subspace Y_n for which the infimum is attained is called optimal. Thus $d_n(\mathfrak{A}\,; X)$ measures the degree to which $\mathfrak{A}$ is approximable in X by subspaces of dimension n. Various other n-width concepts appear in the literature. For example, the linear n-width is often given by

$$\delta_n'(\mathfrak{A}\,;X) = \inf_{Y_n} \inf_{P_n:X\to Y_n} \sup_{x\in\mathfrak{A}} \|x-P_nx\| \tag{1.2}$$

where Y_n is an n-dimensional subspace and P_n is any linear map from X to Y_n. (Thus, δ_n' is the restriction from best approximations, as in d_n, to approximations linear in X.) There are many known examples where d_n and δ_n' do not agree.

The exact determination of d_n, δ_n', and other n-widths which abound in the literature (see e.g., Pietsch [7], Tichomirov [10]) is, of course, generally impossible to obtain. Much effort has nonetheless been devoted toward this goal, as well as to the related task of determining the asymptotic behavior

Journal of Integral Equations 1, 111–130 (1979)

0163-5549/79/02111-20$01.75

of these values as $n \to \infty$. This latter question is of importance not only in approximation theory, but also in the study of ideals of operators between Banach spaces (see, e.g., Pietsch [7], Triebel [11]).

We are here concerned with the exact determination of d_n and δ'_n, and the characterization of optimal subspaces. This has been successfully pursued in a few cases. Consider for example the case where $X = L^q[0, 1]$, and $\mathfrak{A}$ is the unit ball of the Sobolev space $W_p^r[0, 1]$, i.e., $\mathfrak{A} = \{f: f^{(r-1)}$ abs. cont. on $[0, 1]$, $\|f^{(r)}\|_{L^p[0,1]} \leqslant 1\}$. The n-widths d_n and δ'_n have been obtained in the case $r > 1$ only for $p = \infty$, arbitrary q; $q = 1$, arbitrary p; and $p = q = 2$, (see Kolmogorov [3], Tichomirov [9], and [5] and [4]). In fact, in these cases $d_n = \delta'_n$, and more importantly, it has been shown that spline functions of degree $r - 1$ with $n - r$ knots (depending on p and q) form an optimal subspace for d_n (and δ'_n). The other cases remain unsolved and, in fact, not even the asymptotic behavior of δ'_n is known for all $p, q \in [1, \infty]$.

This interesting result concerning the optimality of splines of degree $r - 1$ has its basis in the fact that $W_p^r[0, 1]$ is essentially the image of a unit ball under an integral map and in the total positivity structure of the kernel $(x - y)_+^{r-1}$. In fact, if $X = L^q[0, 1]$, and

$$\mathfrak{A} = \left\{Kh(x) = \int_0^1 K(x,y)h(y)\,dy: \|h\|_{L^p[0,1]} \leqslant 1\right\}, \tag{1.3}$$

where $K(x,y)$ is a totally positive kernel, then for p, q as above (i.e., $p = \infty$, arbitrary q; etc.), there exist $\{\eta_i\}_{i=1}^n$, $0 < \eta_1 < \cdots < \eta_n < 1$ (depending on p, q) such that the subspace spanned by $\{K(x, \eta_i)\}_{i=1}^n$ is optimal for the n-widths d_n and δ'_n. One would very much like to prove this property for all $p, q \in [1, \infty]$. n-widths of $\mathfrak{A}$ of the form (1.3) are one generalization of singular values of the kernel K in that for $p = q = 2$, $d_n(\mathfrak{A}) = [\lambda_{n+1}(K^T K)]^{1/2}$, (see [4]).

In Section 2 we consider the problem of the n-widths of rank $n + 1$ kernels, that is, $\mathfrak{A} = \{Kh: \|h\|_Y \leqslant 1\}$ (see (1.3)), where Y is some normed linear space independent of X. Due to the nature of $\mathfrak{A}$, it is convenient to consider δ_n in place of δ'_n, where

$$\delta_n(\mathfrak{A}; X) = \inf_{Y_n} \inf_{P_n: Y \to Y_n} \sup_{\|h\|_Y \leqslant 1} \|Kh - P_n h\|_X. \tag{1.4}$$

We show that the quantities d_n, δ_n, and the best rank n tensor product approximation to the kernel $K(x,y)$ in the appropriate norm (see Section 2), are all equal for any rank $n + 1$ kernel with suitable conditions on the nature of the norm. This latter problem is for $X = Y = L^2[0, 1]$ and for any real-valued kernel $K(x,y) \in L^2([0, 1] \times [0, 1])$, a classical result of Schmidt [8].

In Section 3 we prove that if, as in the previous example, p, $q \in [1, \infty]$, and $K(x,y)$ is both totally positive and of rank $n + 1$, then it is possible to choose an optimal subspace for the n-widths d_n and δ_n of the form $\{K(x, \eta_i)\}_{i=1}^n$, $0 < \eta_1 < \cdots < \eta_n < 1$.

Section 4 is concerned with various examples and extensions of the results of the previous sections. For example, if $X = l_q^m$, and $\mathfrak{A}$ is the unit ball in l_p^m, then d_n is known if $p \geqslant q$, or if $p = 1$, $q = 2$. However, if $m = n + 1$, then we are able to compute d_n exactly for all p, $q \in [1, \infty]$.

2. n-Widths of Rank $n + 1$ Kernels

In this section we consider several methods of approximating rank $n + 1$ kernels and show that they are all equivalent. In this, we follow the point of view of [5].

Let $\|\cdot\|_i$, $i = 1, 2$ be two norms defined on $X = C[0, 1]$ which have the following properties:

1. $\|\cdot\|_i$ are continuous with respect to the max-norm on X, i.e., there is a constant M such that $\|\cdot\|_i \leqslant M\|\cdot\|_\infty$, $i = 1, 2$.
2. $\|g\|_i = \|h\|_i$, if $g = |h|$.
3. If $\|h\|_{i'} = \sup\{|(g,h)|: \|g\|_i \leqslant 1, g \in X\}$, $((g,h) = \int_0^1 g(x)h(x)\,dx)$, is the conjugate norm, then $\|h\|_{i'} < \infty$ for $h \in X$, and

$$\|h\|_i = \sup\{|(g,h)|: \|g\|_{i'} \leqslant 1, g \in X\}$$

We define two norms on $C([0, 1] \times [0, 1])$ as follows:

Operator Norm. $\|K\|_{1,2} = \sup\{|(Kh,g)|: \|h\|_{1'} \leqslant 1, \quad \|g\|_{2'} \leqslant 1\}$ where $Kh(x) = \int_0^1 K(x,y)h(y)\,dy$.

Property 3 implies that $\|K\|_{1,2} = \sup\{\|Kh\|_2 : \|h\|_{1'} \leqslant 1\}$, which indicates that $\|\cdot\|_{1,2}$ is the operator norm of K as a mapping from $(X, \|\cdot\|_{1'})$ into $(X, \|\cdot\|_2)$.

Tensor Product Norm. $|K|_{1,2} = \|f\|_1$, where $f(y) = \|K(\cdot,y)\|_2$. Note that this definition is permissible since Property 1 implies that $f \in C[0, 1]$.

Lemma 2.1. $\|K\|_{1,2} \leqslant |K|_{1,2}$.

PROOF.

$$|(Kh,g)| = \left|\int_0^1 \left[\int_0^1 K(x,y)h(y)\,dy\right]g(x)\,dx\right|$$

$$\leqslant \int_0^1 \int_0^1 |K(x,y)|\,|h(y)|\,|g(x)|\,dy\,dx$$
$$= \int_0^1 \left[\int_0^1 |K(x,y)|\,|g(x)|\,dx \right] |h(y)|\,dy$$
$$\leqslant \int_0^1 \|K(\cdot,y)\|_2 \|g\|_{2'} |h(y)|\,dy$$
$$= \int_0^1 f(y)|h(y)|\,dy \|g\|_{2'}$$
$$\leqslant |K|_{1,2} \|h\|_{1'} \|g\|_{2'}.$$

The lemma is proven. □

Let

$$\mathfrak{K}_{1'} = \{Kh\colon \|h\|_{1'} \leqslant 1\}, \text{ and}$$
$$\mathfrak{K}_{2'}^T = \{K^T g\colon \|g\|_{2'} \leqslant 1\}, \text{ where}$$
$$K^T = \text{transpose of } K, \text{ and set}$$
$$E_{1,2}^n(K) = \inf_{u_i, v_i \in X} \left| K - \sum_{i=1}^n u_i \otimes v_i \right|_{1,2},$$

where $(u \otimes v)(x,y) = u(x)v(y)$. Further, let X_i denote the space $X = C[0\ 1]$, with the norm $\|\cdot\|_i$. Then,

Lemma 2.2.

$$d_n(\mathfrak{K}_{1'}; X_2) \leqslant \delta_n(\mathfrak{K}_{1'}; X_2) \leqslant E_{1,2}^n(K),$$
$$\textit{and } d_n(\mathfrak{K}_{2'}^T; X_1) \leqslant \delta_n(\mathfrak{K}_{2'}^T; X_1) \leqslant E_{1,2}^n(K).$$

PROOF. The left-hand side inequalities follow from the definitions of d_n and δ_n. To obtain the right-hand side inequalities, we invoke Lemma 2.1 so that for any $\{u_i\}_{i=1}^n, \{v_i\}_{i=1}^n \subseteq X$,

$$\left| \left(\left(K - \sum_{i=1}^n u_i \otimes v_i \right) h, g \right) \right| \leqslant \left| K - \sum_{i=1}^n u_i \otimes v_i \right|_{1,2} \|h\|_{1'} \|g\|_{2'}.$$

Thus,

$$\left\| Kh - \sum_{i=1}^n u_i(v_i, h) \right\|_2 \leqslant \left| K - \sum_{i=1}^n u_i \otimes v_i \right|_{1,2} \|h\|_{1'}$$

and

$$\left\| K^T g - \sum_{i=1}^{n} (u_i, g) v_i \right\|_1 \leqslant \left| K - \sum_{i=1}^{n} u_i \otimes v_i \right|_{1,2} \|g\|_{2'}.$$

These inequalities easily imply the assertion of the lemma since $\sum_{i=1}^{n} u_i(v_i, h) = P_n h$ and $\sum_{i=1}^{n} (u_i, g)v_i = Q_n g$, are linear maps of X into $[u_1, \ldots, u_n] \equiv \operatorname{span}\{u_1 \cdots, u_n\}$ and $[v_1, \ldots, v_n]$, respectively. The lemma is proven.

□

The kernel K is said to be of rank $n + 1$ if

$$K = \sum_{i=1}^{n+1} \phi_i \otimes \psi_i,$$

where $\{\phi_1, \ldots, \phi_{n+1}\}$ and $\{\psi_1, \ldots, \psi_{n+1}\}$ are each linearly independent subsets of X. The main result of this section is the following:

Theorem 2.1. *If K is a rank $n + 1$ kernel, then $E_{1,2}^n(K) = \delta_n(\mathfrak{K}_{1'}; X_2) = \delta_n(\mathfrak{K}_{2'}^T; X_1) = d_n(\mathfrak{K}_{1'}; X_2) = d_n(\mathfrak{K}_{2'}^T; X_1)$ and their common value is*

$$\delta = \min_{\sum_{i=1}^{n+1} a_i b_i = 1} \left\| \sum_{i=1}^{n+1} a_i \phi_i \right\|_2 \left\| \sum_{i=1}^{n+1} b_i \psi_i \right\|_1. \tag{2.1}$$

The following two lemmas are used in the proof of the theorem.

Lemma 2.3. *Let $\phi_1, \ldots, \phi_{n+1}$ be linearly independent functions in X. Given any numbers $(a_1, \ldots, a_{n+1})$, not all zero, there exist measures $d\alpha_1, \ldots, d\alpha_n$ such that*

$$a_i = (-1)^{i+1} \det_{\substack{j=1,\ldots,n \\ k=1,\ldots,n+1 \\ k \neq i}} \left\| \int_0^1 \phi_k(x)\, d\alpha_j(x) \right\|. \tag{2.2}$$

PROOF. First observe that from the linear independence of the $\{\phi_i\}_{i=1}^{n+1}$, there exist $\{x_i\}_{i=1}^{n+1}$, $0 \leqslant x_1 < \cdots < x_{n+1} \leqslant 1$, such that the matrix $C = \|C_{ij}\|$, $C_{ij} = \phi_j(x_i)$, is nonsingular. Since $a \neq 0$, the vector $c = Ca$ is nonzero. Construct any nonsingular $(n + 1) \times (n + 1)$ matrix B such that $Bc = e_{n+1}$, where $(e_{n+1})_i = \delta_{i,n+1}$. Let

$$d\alpha_i = \lambda \sum_{j=1}^{n+1} B_{ij} d\mu_{x_j}, \qquad i = 1, \ldots, n,$$

where $d\mu_x$ represents point evaluation at x, and λ is some nonzero constant to be chosen later. Now,

$$\int_0^1 \phi_j(x)d\alpha_i(x) = \sum_{k=1}^{n+1} B_{ik}\phi_j(x_k)$$
$$= \sum_{k=1}^{n+1} B_{ik}C_{kj}$$
$$= (BC)_{ij}.$$

Since $BCa = e_{n+1}$, the lemma follows by Cramer's rule, the nonsingularity of BC, and the appropriate choice of λ. □

REMARK 2.1. Note that (2.2) is equivalent to

$$\underset{\substack{i=1,\ldots,n\\ j=1,\ldots,n+1}}{\operatorname{rank}} \left\| \int_0^1 \phi_j(x)d\alpha_i(x) \right\| = n$$

$$\int_0^1 \left(\sum_{j=1}^{n+1} a_j\phi_j(x) \right) d\alpha_i(x) = 0, \qquad i = 1, \ldots, n. \tag{2.3}$$

For ease of exposition, we introduce the following notation.

Definition 2.1. Let $d\alpha_1, \ldots, d\alpha_n$, $d\beta_1, \ldots, d\beta_n$ be measures, $K \in C([0,1] \times [0,1])$, and $u_i \in C[0,1]$, $i = 1, \ldots, n+1$. We set

$$K\begin{pmatrix} d\alpha_1, \ldots, d\alpha_n \\ d\beta_1, \ldots, d\beta_n \end{pmatrix} = \det_{i,j=1,\ldots,n} \left\| \int_0^1 \int_0^1 K(x,y)\, d\alpha_i(x)\, d\beta_j(y) \right\|.$$

Also,

$$K\begin{pmatrix} d\mu_x, d\alpha_1, \ldots, d\alpha_n \\ d\mu_y, d\beta_1, \ldots, d\beta_n \end{pmatrix}$$

will often be denoted by

$$K\begin{pmatrix} x, d\alpha_1, \ldots, d\alpha_n \\ y, d\beta_1, \ldots, d\beta_n \end{pmatrix}.$$

Further,

$$U\begin{pmatrix} x, d\alpha_1, \ldots, d\alpha_n \\ 1, \ldots, n+1 \end{pmatrix} = \det_{\substack{k=0,1,\ldots,n\\ j=1,\ldots,n+1}} \left\| \int_0^1 u_j(x)\, d\alpha_k(x) \right\| \text{ where } d\alpha_0 = d\mu_x,$$

and

$$U\begin{pmatrix} d\alpha_1, \ldots, d\alpha_n \\ 1, \ldots, \hat{\imath}, \ldots, n+1 \end{pmatrix} = \det_{\substack{k=1,\ldots,n\\ j=1,\ldots,n+1\\ j\neq i}} \left\| \int_0^1 u_j(x)\, d\alpha_k(x) \right\|.$$

The reason for the above notation is partially contained in the following lemma.

Lemma 2.4. *If* $K \doteq \sum_{i=1}^{n+1} \phi_i \otimes \psi_i$, *then*

$$K\begin{pmatrix} x, d\alpha_1, \ldots, d\alpha_n \\ y, d\beta_1, \ldots, d\beta_n \end{pmatrix} = \phi\begin{pmatrix} x, d\alpha_1, \ldots, d\alpha_n \\ 1, \ldots, n+1 \end{pmatrix}\psi\begin{pmatrix} y, d\beta_1, \ldots, d\beta_n \\ 1, \ldots, n+1 \end{pmatrix}$$

and

$$K\begin{pmatrix} d\alpha_1, \ldots, d\alpha_n \\ d\beta_1, \ldots, d\beta_n \end{pmatrix}$$
$$= \sum_{i=1}^{n+1} \phi\begin{pmatrix} d\alpha_1, \ldots, d\alpha_n \\ 1, \ldots, \hat{\imath}, \ldots, n+1 \end{pmatrix}\psi\begin{pmatrix} d\beta_1, \ldots, d\beta_n \\ 1, \ldots, \hat{\imath}, \ldots, n+1 \end{pmatrix}.$$

PROOF. The first equality is a direct result of the fact that $\det(AB) = (\det A)(\det B)$. The second equality follows from an application of the Cauchy–Binet formula, (see Karlin [2, p. 1]).

PROOF OF THEOREM 2.1. We first prove the upper bound. Note that

$$E_{1,2}^n(K) = \inf_{u_i, v_i \in X} \left| K - \sum_{i=1}^{n} u_i \otimes v_i \right|_{1,2}$$
$$= \inf_{\substack{u_i, v_j \in X \\ c_{ij} \in R}} \left| K - \sum_{i,j=1}^{n} c_{ij}(u_i \otimes v_j) \right|_{1,2}.$$

Let $d\alpha_1, \ldots, d\alpha_n, d\beta_1, \ldots, d\beta_n$ be any $2n$ measures for which $K\binom{d\alpha_1, \ldots, d\alpha_n}{d\beta_1, \ldots, d\beta_n} \neq 0$. Then,

$$E_{1,2}^n(K) \leqslant \frac{\left| K\begin{pmatrix} x, d\alpha_1, \ldots, d\alpha_n \\ y, d\beta_1, \ldots, d\beta_n \end{pmatrix} \right|}{\left| K\begin{pmatrix} d\alpha_1, \ldots, d\alpha_n \\ d\beta_1, \ldots, d\beta_n \end{pmatrix} \right|} 1, 2$$

by the choice of $u_i(x) = \int_0^1 K(x,y)\, d\beta_i(y)$, $v_i(y) = \int_0^1 K(x,y)\, d\alpha_i(x)$, and c_{ij} appropriately taken. From Lemma 2.4, we obtain

$$E_{1,2}^n(K) \leqslant \frac{\left| \phi\begin{pmatrix} x, d\alpha_1, \ldots, d\alpha_n \\ 1, \ldots, n+1 \end{pmatrix}\psi\begin{pmatrix} y, d\beta_1, \ldots, d\beta_n \\ 1, \ldots, n+1 \end{pmatrix} \right|_{1,2}}{\left| \sum_{i=1}^{n+1} \phi\begin{pmatrix} d\alpha_1, \ldots, d\alpha_n \\ 1, \ldots, \hat{\imath}, \ldots n+1 \end{pmatrix}\psi\begin{pmatrix} d\beta_1, \ldots, d\beta_n \\ 1, \ldots, \hat{\imath}, \ldots n+1 \end{pmatrix} \right|}$$

$$= \frac{\left\|\phi\begin{pmatrix} x, d\alpha_1, \ldots, d\alpha_n \\ 1, \ldots, n+1 \end{pmatrix}\right\|_2 \left\|\psi\begin{pmatrix} y, d\beta_1, \ldots, d\beta_n \\ 1, \ldots, n+1 \end{pmatrix}\right\|_1}{\left|\sum_{i=1}^{n+1} \phi\begin{pmatrix} d\alpha_1, \ldots, d\alpha_n \\ 1, \ldots, \hat{\imath}, \ldots n+1 \end{pmatrix} \psi\begin{pmatrix} d\beta_1, \ldots, d\beta_n \\ 1, \ldots, \hat{\imath}, \ldots, n+1 \end{pmatrix}\right|}.$$

Let

$$a_i = (-1)^{i+1}\phi\begin{pmatrix} d\alpha_1, \ldots, d\alpha_n \\ 1, \ldots, \hat{\imath}, \ldots, n+1 \end{pmatrix}, \text{ and}$$

$$b_i = (-1)^{i+1}\psi\begin{pmatrix} d\beta_1, \ldots, d\beta_n \\ 1, \ldots, \hat{\imath}, \ldots, n+1 \end{pmatrix}.$$

The above inequality then reduces to

$$E_{1,2}^n(K) \leqslant \frac{\left\|\sum_{i=1}^{n+1} a_i\phi_i\right\|_2 \left\|\sum_{i=1}^{n+1} b_i\psi_i\right\|_1}{\left|\sum_{i=1}^{n+1} a_i b_i\right|},$$

for any $\{a_i\}_{i=1}^{n+1}$, $\{b_i\}_{i=1}^{n+1}$ obtained as above for which $\sum_{i=1}^{n+1} a_i b_i = K\binom{d\alpha_1, \ldots, d\alpha_n}{d\beta_1, \ldots, d\beta_n} \neq 0$. An application of Lemma 2.3 yields $E_{1,2}^n(K) \leqslant \delta$.

To prove the lower bound, we first note that $\mathfrak{K}_{1'}$, is an $n+1$ dimensional subset of the $n+1$ dimensional subspace $[\phi_1, \ldots, \phi_{n+1}]$ of X_2. Since $\mathfrak{K}_{1'}$ is convex, centrally symmetric, and closed, it follows from Lemma 5 and Theorems 5 and 6 of Brown [1], that $d_n(\mathfrak{K}_{1'}; x_2) = \inf\{\|g\|_2 : g \in \text{boundary of } \mathfrak{K}_{1'}\}$. Now, $g = Kh = \sum_{i=1}^{n+1} a_i\phi_i$, where $a_i = \int_0^1 \psi_i(y)h(y)\,dy$, $i = 1, \ldots, n+1$. Since $(a_1, \ldots, a_{n+1}) \to \sum_{i=1}^{n+1} a_i\phi_i$ is a homeomorphism, the boundary of $\mathfrak{K}_{1'}$ consists of all such g where $(a_1, \ldots, a_{n+1})$ lies on the boundary of the moment space

$$\mathfrak{M} = \left\{(c_1, \ldots, c_{n+1}) : c_i = \int_0^1 \psi_i(y)h(y)\,dy, \|h\|_{1'} \leqslant 1\right\}.$$

Since a $\in$ boundary of $\mathfrak{M}$, there exists a $b \in R^{n+1}$ such that

$$|(b, c)| \leqslant (b, a) \text{ for all } c \in \mathfrak{M}.$$

Hence $\|\sum_{i=1}^{n+1} b_i\psi_i\|_1 \leqslant |\sum_{i=1}^{n+1} a_i b_i|$, and thus

$$\|g\|_2 = \left\|\sum_{i=1}^{n+1} a_i\phi_i\right\|_2 \geqslant \frac{\left\|\sum_{i=1}^{n+1} a_i\phi_i\right\|_2 \left\|\sum_{i=1}^{n+1} b_i\psi_i\right\|_1}{\left|\sum_{i=1}^{n+1} a_i b_i\right|} \geqslant \delta.$$

Consequently, we have shown that $d_n(\mathfrak{K}_{1'};X_2) \geqslant \delta$. Similarly, $d_n(\mathfrak{K}_{2'}^T;X_1) = \inf\{\|\sum_{i=1}^{n+1} b_i\psi_i\|_1 : (b_1,\ldots,b_{n+1}) \in \partial\mathfrak{N}\}$ where $\mathfrak{N} = \{(c_1,\ldots,c_{n+1}) : c_1 = \int_0^1 \phi_i(x)h(x)\,dx, \|h\|_{2'} \leqslant 1\}$. As above, it follows that $d_n(\mathfrak{K}_{2'}^T;X_1) \geqslant \delta$. We now apply Lemma 2.2.

It remains to be proven that the value δ is, in fact, attained. Suppose that $\{a_i^k\}$, $\{b_i^k\}$ is a minimizing sequence. We may renormalize these sequences such that $\sum_{i=1}^{n+1} a_i^k b_i^k = 1 = \sum_{i=1}^{n+1} |b_i^k|$. Hence $b_i^{k'} \to b_i$, through a subsequence, and therefore $\lim_{k'\to\infty} \|\sum_{i=1}^{n+1} a_i^{k'} \phi_i\|_2 = \delta/\|\sum_{i=1}^{n+1} b_i\psi_i\|_1$. Consequently, some subsequence of $\{a_i^k\}$ also converges. The proof of the theorem is complete. □

Several useful facts emerge from the proof of the theorem. Suppose $\delta = \|\overline{\phi}\|_2\|\overline{\psi}\|_1$, where $\overline{\phi} = \sum_{i=1}^{n+1} \overline{a}_i\phi_i$, $\overline{\psi} = \sum_{i=1}^{n+1} \overline{b}_i\psi_i$, and $\sum_{i=1}^{n+1} \overline{a}_i\overline{b}_i = 1$. Let $d\overline{\alpha}_i$, $d\overline{\beta}_i$, $i = 1, \ldots, n$ be any $2n$ measures such that (2.2) (or (2.3)) holds for $\overline{a}_i$ and $\overline{b}_i$, $i = 1, \ldots, n+1$, respectively. Set $\overline{u}_i(x) = \int_0^1 K(x,y)\,d\overline{\beta}_i(y)$, and $\overline{v}_i(y) = \int_0^1 K(x,y)\,d\overline{\alpha}_i(x)$, $i = 1, \ldots, n$. Since $1 = \sum_{i=1}^{n+1} \overline{a}_i\overline{b}_i = K\binom{d\overline{\alpha}_1,\ldots,d\overline{\alpha}_n}{d\overline{\beta}_1,\ldots,d\overline{\beta}_n}$, we may define two projections $P_{\overline{\alpha}}$ and $P_{\overline{\beta}}$ as follows:

1. $P_{\overline{\alpha}}: X \to [\overline{u}_1,\ldots,\overline{u}_n]$ and satisfies

$$\int_0^1 [Kh(x) - P_{\overline{\alpha}}h(x)]\,d\overline{\alpha}_i(x) = 0,\ i = 1, \ldots, n,\ h \in X.$$

2. $P_{\overline{\beta}}: X \to [\overline{v}_1,\ldots,\overline{v}_n]$ and satisfies $\int_0^1 [K^Tg(y) - P_{\overline{\beta}}g(y)]\,d\overline{\beta}_i(y) = 0$, $i = 1, \ldots, n$, $g \in X$. Then,

1. $[\overline{u}_1,\ldots,\overline{u}_n]$ is an optimal subspace for $d_n(\mathfrak{K}_{1'};X_2)$ and $\delta_n(\mathfrak{K}_{1'};X_2)$, and $P_{\overline{\alpha}}$ (2.4)

furnishes an optimal linear method of approximation.

2. $[\overline{v}_1,\ldots,\overline{v}_n]$ is an optimal subspace for $d_n(\mathfrak{K}_{2'}^T;X_1)$ and $\delta_n(\mathfrak{K}_{2'}^T;X_1)$, and P_{β} (2.5)

furnishes an optimal linear method of approximation.

3. $E_{1,2}^n(K) = \left|K - \sum_{i,j=1}^{n} \overline{c}_{ij}(\overline{u}_i \otimes \overline{v}_j)\right|_{1,2}$ where $\overline{u}_i$, $\overline{v}_j$, $i,j = 1, \ldots, n$, (2.6)

are as above and $\overline{c}_{ij}$ are appropriately chosen numbers.

3. Measures of Smallest Support

Our goal in this section is to consider a class of kernels and to prove that for these given kernels there exist optimal subspaces for $d_n(\mathfrak{K}_{1'};X_2)$ and $d_n(\mathfrak{K}_{2'}^T;X_1)$, (and also of course for δ_n), which are of the form $[K(x,\eta_1),\ldots,K(x,\eta_n)]$ and $[K(\xi_1,y)\cdots,K(\xi_n,y)]$, respectively, where $0<\eta_1<\cdots<\eta_n<1$ and $0<\xi_1<\cdots<\xi_n<1$. In other words, we seek conditions on the kernel for which, in the notation of the previous section,

$$d\bar{\alpha}_i = d\mu_{\xi_i} \quad , i=1,\ldots,n \tag{3.1}$$

$$d\bar{\beta}_i = d\mu_{\eta_i} \quad , i=1,\ldots,n. \tag{3.2}$$

On the basis of (2.3), if this is possible, it would be necessary that

$$\begin{aligned} \operatorname{rank}\|\phi_i(\xi_j)\| &= n, \qquad i=1,\ldots,n+1; j=1,\ldots,n. \\ \bar{\phi}(\xi_j) &= 0, \qquad j=1,\ldots,n \end{aligned} \tag{3.3}$$

and analogously,

$$\begin{aligned} \operatorname{rank}\|\psi_i(\eta_j)\| &= n, \qquad i=1,\ldots,n+1; j=1,\ldots,n. \\ \bar{\psi}(\eta_j) &= 0, \qquad j=1,\ldots,n, \end{aligned} \tag{3.4}$$

where as previously,

$$\min_{\sum_{i=1}^{n+1} a_i b_i = 1} \left\|\sum_{i=1}^{n+1} a_i\phi_i\right\|_2 \left\|\sum_{i=1}^{n+1} b_i\psi_i\right\|_1 = \|\bar{\phi}\|_2\|\bar{\psi}\|_1 = \delta.$$

Before stating our result, the following definition is necessary.

Definition 3.1. A set $\{u_1,\ldots,u_n\}$ of linearly independent functions on X is called a *weak Chebyshev system* if for all $0 \leqslant x_1 < \cdots < x_n \leqslant 1$,

$$\det_{i,j=1,\ldots,n} \|u_i(x_j)\| = U\begin{pmatrix} 1,\ldots,n \\ x_1,\ldots,x_n \end{pmatrix} \geqslant 0.$$

If strict inequality prevails above, then $\{u_1,\ldots,u_n\}$ is called a *Chebyshev system.*

In this section $\|\cdot\|_p$ shall denote the usual $L^p[0,1]$ norm, $1 \leqslant p \leqslant \infty$. We now state the result.

Theorem 3.1. *Let* $\{\phi_1,\ldots,\phi_{i-1},\phi_{i+1},\ldots,\phi_{n+1}\}$, $\{\psi_1,\ldots,\psi_{i-1},\psi_{i+1},\ldots,\psi_{n+1}\}$, $i=1,\ldots,n+1$, $\{\phi_i\}_{i=1}^{n+1}$, *and* $\{\psi_i\}_{i=1}^{n+1}$ *all be Chebyshev systems on* $[0,1]$. *Let* p, $q \in [1,\infty]$. *If* $\{\bar{a}_i\}_{i=1}^{n+1}$ *and* $\{\bar{b}_i\}_{i=1}^{n+1}$, *are such that* $\sum_{i=1}^{n+1}\bar{a}_i\bar{b}_i = 1$, *and*

$$\delta = \left\| \sum_{i=1}^{n+1} \bar{a}_i \phi_i \right\|_q \left\| \sum_{i=1}^{n+1} \bar{b}_i \psi_i \right\|_p ,$$

then $\sum_{i=1}^{n+1} \bar{a}_i \phi_i$ *and* $\sum_{i=1}^{n+1} \bar{b}_i \psi_i$ *each have exactly n distinct zeros in* $(0, 1)$.

Recall that δ is always the value of the minimum problem of Theorem 2.1.

Before embarking on the proof of Theorem 3.1, we record its consequence.

Theorem 3.2. *Let* $K = \sum_{i=1}^{n+1} \phi_i \otimes \psi_i$, *where* $\{\phi_i\}_{i=1}^{n+1}$ *and* $\{\psi_i\}_{i=1}^{n+1}$ *satisfy the conditions of Theorem 3.1. The Kolmogorov and linear n-widths of* $\mathfrak{K}_{p'} = \{Kh\colon \|h\|_{p'} \leqslant 1\}$, *where* $(1/p) + (1/p') = 1$, *as a subset of* L^q *are* δ. *Moreover, if* $\sum_{i=1}^{n+1} \bar{a}_i \phi_i(\xi_j) = 0, j = 1, \ldots, n, 0 < \xi_1 < \cdots < \xi_n < 1$, *and* $\sum_{i=1}^{n+1} \bar{b}_i \psi_i(\eta_j) = 0, j = 1, \ldots, n, 0 < \eta_1 < \cdots < \eta_n < 1$, *then the set* $\{K(x, \eta_1), \ldots, K(x, \eta_n)\}$ *spans an optimal subspace for the n-widths, and interpolation at* $\xi_1, \ldots, \xi_n$ *is an optimal linear method of approximating* $\mathfrak{K}_{p'}$.

There is of course, on the basis of Theorems 2.1 and 3.1, a companion result for $\mathfrak{K}_{q'}^T$ as a subset of L^p. We leave its precise formulation to the reader.

The proof of Theorem 3.1 necessitates the following proposition. In all that follows, we assume that as in Theorem 3.1, $\{\phi_1, \ldots, \phi_{i-1}, \phi_{i+1}, \ldots, \phi_{n+1}\}$, $\{\psi_1, \ldots, \psi_{i-1}, \psi_{i+1}, \ldots, \psi_{n+1}\}$, $i = 1, \ldots, n+1$, $\{\phi_i\}_{i=1}^{n+1}$, and $\{\psi_i\}_{i=1}^{n+1}$ are Chebyshev systems on $[0, 1]$.

Proposition 3.1. *Let* $\{d_i\}_{i=1}^{n+1}$ *be any real nonnegative numbers such that* $\sum_{i=1}^{n+1} d_i > 0$. *Then for fixed* $q \in [1, \infty]$, *there is a unique* $\bar{\phi}(x) = \sum_{i=1}^{n+1} \bar{a}_i \phi_i(x)$ *which minimizes* $\|\sum_{i=1}^{n+1} a_i \phi_i\|_q$ *subject to the constraint* $\sum_{i=1}^{n+1} |a_i| d_i = 1$. *Moreover,* $\bar{\phi}$ *has exactly n distinct zeros (sign changes) in* $(0, 1)$, *and the* $\bar{a}_j$ *alternate in sign.*

PROOF. Assume that $1 \leqslant q < \infty$. The case $q = \infty$ will be handled separately.

Let $\phi^*(x) = \sum_{i=1}^{n+1} a_i^* \phi_i(x)$ be any function which minimizes $\|\sum_{i=1}^{n+1} a_i \phi_i\|_q$, with $\sum_{i=1}^{n+1} |a_i^*| d_i = 1$.

We first wish to prove

$$\begin{aligned} &\int_0^1 |\phi^*(x)|^{q-1} [\operatorname{sgn} \phi^*(x)] \phi_k(x)\, dx \\ &= \left[\int_0^1 |\phi^*(x)|^q\, dx \right] [\operatorname{sgn} a_k^*]\, d_k, \quad k = 1, \ldots, n+1. \end{aligned} \tag{3.5}$$

The problem in the proof of (3.5) is that $|\bar{a}_k|$ is not differentiable at $\bar{a}_k = 0$. This, of course, is irrelevant if $d_k = 0$, in which case the orthogonality

$$\int_0^1 |\phi^*(x)|^{q-1}[\operatorname{sgn} \phi^*(x)]\phi_k(x)\,dx = 0$$

is easily proven since there is then no restriction whatsoever on a_k. If $d_k > 0$ and $a_k^* \neq 0$, then formula (3.5) is easily shown to hold (by the Lagrange multiplier method, for example). In fact however, one cannot have $a_k^* = 0$ if $d_k > 0$. To see this note that if $d_k > 0$ and $a_k^* = 0$, then taking left and right limits, it is necessary that

$$\pm \int_0^1 |\phi^*(x)|^{q-1}[\operatorname{sgn} \phi^*(x)]\phi_k(x)\,dx \geqslant \left[\int_0^1 |\phi^*(x)|^q\,dx\right] d_k,$$

i.e., both inequalities hold. Since $d_k[\int_0^1 |\bar{\phi}(x)|^q\,dx] > 0$, this is impossible. Thus (3.5) holds for all $k = 1, \ldots, n + 1$.

Now, consider $\min \|\sum_{i=1}^{n+1} a_i\phi_i\|_q$ subject to the constraint $\sum_{i=1}^{n+1} a_i(-1)^i d_i = 1$. Let $\bar{\phi}(x) = \sum_{i=1}^{n+1} \bar{a}_i\phi_i(x)$ be the unique solution to this problem, with $\sum_{i=1}^{n+1} \bar{a}_i(-1)^i d_i = 1$. By the method of Lagrange multipliers it follows that

$$\int_0^1 |\bar{\phi}(x)|^{q-1}[\operatorname{sgn} \bar{\phi}(x)]\phi_k(x)\,dx = \left[\int_0^1 |\bar{\phi}(x)|^q\,dx\right](-1)^k d_k, \qquad k = 1, \ldots, n + 1. \tag{3.6}$$

Since $\{(-1)^k d_k\}_{k=1}^{n+1}$ weakly alternates in sign, and $\{\phi_k(x)\}_{k=1}^{n+1}$ is a Chebyshev system on $[0, 1]$, it is necessary that $\bar{\phi}(x)$ have n sign changes in $(0, 1)$. The condition that $\{\phi_1, \ldots, \phi_{i-1}, \phi_{i+1}, \ldots, \phi_{n+1}\}$ is a Chebyshev system on $[0,1]$ for $i = 1, \ldots, n + 1$, implies that since $\bar{\phi}(x)$ has n sign changes on $(0,1)$, we must have $\bar{a}_i\bar{a}_{i+1} < 0$, $i = 1, \ldots, n$. The constraint then implies that $\bar{a}_i(-1)^i > 0$, $i = 1, \ldots, n + 1$ so that, in fact, $1 = \sum_{i=1}^{n+1} \bar{a}_i(-1)^i d_i = \sum_{i=1}^{n+1} |\bar{a}_i| d_i$, i.e., the $\{\bar{a}_i\}_{i=1}^{n+1}$ are permissible in the original minimum problem.

Now, assume that $\|\sum_{i=1}^{n+1} a_i^*\phi_i\|_q \leqslant \|\sum_{i=1}^{n+1} \bar{a}_i\phi_i\|_q$, and $\phi^*(x) \neq \pm\bar{\phi}(x)$. Thus, from (3.5) and (3.6),

$$\begin{aligned} &\int_0^1 (|\bar{\phi}(x)|^{q-1}[\operatorname{sgn} \bar{\phi}(x)] \pm |\phi^*(x)|^{q-1}[\operatorname{sgn} \phi^*(x)])\phi_k(x)\,dx \\ &\quad = (\|\bar{\phi}\|_q^q(-1)^k \pm \|\phi^*\|_q^q[\operatorname{sgn} a_k^*])d_k, \qquad k = 1, \ldots, n + 1. \end{aligned} \tag{3.7}$$

Since by assumption, $\{\|\bar{\phi}\|_q^q(-1)^k \pm \|\phi^*\|_q^q[\operatorname{sgn} a_k^*]\}$, $k = 1, \ldots, n + 1$, weakly alternates in sign (the orientation of which is independent of the $\pm$), $d_k \geqslant 0$, $k = 1, \ldots, n + 1$, and $\{\phi_k\}_{k=1}^{n+1}$ is a Chebyshev system, it is necessary, as previously, that $(|\bar{\phi}(x)|^{q-1}[\operatorname{sgn} \bar{\phi}(x)] \pm |\phi^*(x)|^{q-1}[\operatorname{sgn} \phi^*(x)])$ exhibits at least n sign changes in (0,1), and $n + 1$ if $\{\|\bar{\phi}\|_q^q(-1)^k \pm \|\phi^*\|_q^q \cdot \operatorname{sgn} a_k^*\}$, $k = 1, \ldots, n + 1$ is an identically zero sequence for some $\pm$.

However, it is easily seen that

$$\operatorname{sgn}(|\bar{\phi}(x)|^{q-1}[\operatorname{sgn}\bar{\phi}(x)] \pm |\phi^*(x)|^{q-1}[\operatorname{sgn}\phi^*(x)]) = \operatorname{sgn}(\bar{\phi}(x) \pm \phi^*(x)).$$

Since $\bar{\phi}(x) \not\equiv \pm\phi^*(x)$, we cannot have $\|\bar{\phi}\|_q^q(-1)^k = \pm\|\phi^*\|_q^q \operatorname{sgn} a_k^*$ for $k = 1, \ldots, n+1$. Thus, $\bar{\phi}(x) \pm \phi^*(x) = \sum_{i=1}^{n+1} (\bar{a}_i \pm a_i^*)\phi_i(x)$ exhibits exactly n sign changes in (0,1) and the orientation of the sign is fixed, independent of the $\pm$. This implies that $|\bar{a}_i| > |a_i^*|$ for all $i = 1, \ldots, n+1$. Hence $1 = \sum_{i=1}^{n+1} |\bar{a}_i| d_i > \sum_{i=1}^{n+1} |a_i^*| d_i = 1$. This is a contradiction. The proposition, including the uniqueness, has been proven for $1 \leqslant q < \infty$.

We now prove the proposition for $q = \infty$. Let $\bar{\phi}(x) = \sum_{i=1}^{n+1} \bar{a}_i\phi_i(x)$ be the polynomial which equioscillates at $n+1$ points in [0,1], and is normalized such that $\bar{a}_i(-1)^i > 0$, $i = 1, \ldots, n+1$, and $\sum_{i=1}^{n+1} |\bar{a}_i| d_i = 1$. It is well known that $\bar{\phi}$ exists and is unique. Let $\phi^*(x) = \sum_{i=1}^{n+1} a_i^*\phi_i(x)$ be any other polynomial ($\phi^* \not\equiv \pm\bar{\phi}$) for which $\sum_{i=1}^{n+1} |a_i^*| d_i = 1$. If $\|\phi^*\|_\infty \leqslant \|\bar{\phi}\|_\infty$, then, since $\bar{\phi}(x)$ equioscillates at $n+1$ points, $\bar{\phi}(x) \pm \phi^*(x) = \sum_{i=1}^{n+1} (\bar{a}_i \pm a_i^*)\phi_i(x)$ must have at least n weak sign changes with orientation independent of the $\pm$. As before, this implies that $|\bar{a}_i| > |a_i^*|$, for all $i = 1, \ldots, n+1$, and the result follows. Note that for $q = \infty$, $\bar{\phi}(x)$ is independent of the $\{d_i\}_{i=1}^{n+1}$ up to multiplication by a constant. This fact is further examined in Remark 4.1. The proposition is proven. □

Proof of Theorem 3.1. Let $\{a_i^*\}_{i=1}^{n+1}$ and $\{b_i^*\}_{i=1}^{n+1}$ be any solution to: Minimize $\|\sum_{i=1}^{n+1} a_i\phi_i\|_q \|\sum_{i=1}^{n+1} b_i\psi_i\|_p$ subject to $|\sum_{i=1}^{n+1} a_i b_i| = 1$. Assume that $\phi^* = \sum_{i=1}^{n+1} a_i^*\phi_i$ does not have n distinct zeros in (0,1). Then,

$$\frac{\left\|\sum_{i=1}^{n+1} a_i^*\phi_i\right\|_q \left\|\sum_{i=1}^{n+1} b_i^*\psi_i\right\|_p}{\left|\sum_{i=1}^{n+1} a_i^* b_i^*\right|} \geqslant \frac{\left\|\sum_{i=1}^{n+1} a_i^*\phi_i\right\|_q \left\|\sum_{i=1}^{n+1} b_i^*\psi_i\right\|_p}{\sum_{i=1}^{n+1} |a_i^*||b_i^*|}$$

$$> \frac{\left\|\sum_{i=1}^{n+1} \bar{a}_i\phi_i\right\|_q \left\|\sum_{i=1}^{n+1} b_i^*\psi_i\right\|_p}{\sum_{i=1}^{n+1} |\bar{a}_i||b_i^*|},$$

where $\{\bar{a}_i\}_{i=1}^{n+1}$ depends upon $\{|b_i^*|\}_{i=1}^{n+1}$ (chosen as in Proposition 3.1), and are such that $\bar{\phi} = \sum_{i=1}^{n+1} \bar{a}_i\phi_i$ has n distinct zeros in (0,1). Similarily, if $\psi^* = \sum_{i=1}^{n+1} b_i^*\psi_i$ does not exhibit n distinct zeros, we again apply Proposition 3.1 to obtain

$$\frac{\left\|\sum_{i=1}^{n+1} a_i^* \phi_i\right\|_q \left\|\sum_{i=1}^{n+1} b_i^* \psi_i\right\|_p}{\left|\sum_{i=1}^{n+1} a_i^* b_i^*\right|} > \frac{\left\|\sum_{i=1}^{n+1} \bar{a}_i \phi_i\right\|_q \left\|\sum_{i=1}^{n+1} \bar{b}_i \psi_i\right\|_p}{\sum_{i=1}^{n+1} |\bar{a}_i||\bar{b}_i|}$$

where the $\{\bar{b}_i\}_{i=1}^{n+1}$ depend upon $\{|\bar{a}_i|\}_{i=1}^{n+1}$. Now, by Proposition 3.1, both the $\{\bar{a}_i\}_{i=1}^{n+1}$ and the $\{\bar{b}_i\}_{i=1}^{n+1}$ alternate in sign. Thus, $\sum_{i=1}^{n+1} |\bar{a}_i||\bar{b}_i| = |\sum_{i=1}^{n+1} \bar{a}_i \bar{b}_i|$. Whence we have shown

$$\frac{\left\|\sum_{i=1}^{n+1} a_i^* \phi_i\right\|_q \left\|\sum_{i=1}^{n+1} b_i^* \psi_i\right\|_p}{\left|\sum_{i=1}^{n+1} a_i^* b_i^*\right|} > \frac{\left\|\sum_{i=1}^{n+1} \bar{a}_i \phi_i\right\|_q \left\|\sum_{i=1}^{n+1} \bar{b}_i \psi_i\right\|_p}{\left|\sum_{i=1}^{n+1} \bar{a}_i \bar{b}_i\right|},$$

a contradiction. The theorem is proven. □

4. Remarks and Examples

RESULT 4.1. In the proof of Proposition 3.1, it was shown that if $\{\phi_i\}_{i=1}^{n+1}$ has the Chebyshev properties of Theorem 3.1, and $d_i \geqslant 0$, $i = 1, \ldots, n+1$, $\sum_{i=1}^{n+1} d_i > 0$, then the unique solution of $\min\|\sum_{i=1}^{n+1} a_i \phi_i\|_\infty$ subject to the constraint $\sum_{i=1}^{n+1} |a_i| d_i = 1$, is $\bar{\phi} = \sum_{i=1}^{n+1} \bar{a}_i \phi_i$. Furthermore $\bar{\phi}$, up to multiplication by a constant, is independent of the $\{d_i\}_{i=1}^{n+1}$. Hence, in particular, the zeros $\{\xi_i\}_{i=1}^{n}$ of $\bar{\phi}$ are independent of $\{d_i\}_{i=1}^{n+1}$. As a result we have the following proposition.

Proposition 4.1. *Let $K = \sum_{i=1}^{n+1} \phi_i \otimes \psi_i$ where $\{\phi_i\}_1^{n+1}$ and $\{\psi_i\}_1^{n+1}$ satisfy the conditions of Theorem 3.1. For every $p \in [1, \infty]$, $[K(\xi_1, y), \ldots, K(\xi_n, y)]$ is an optimal subspace in the Kolmogorov and linear n-widths $d_n(\mathfrak{K}_1^T; L^p)$ and $\delta_n(\mathfrak{K}_1^T; L^p)$. Further, for each $p \in [1, \infty]$, there exist $\{\eta_i^p\}_{i=1}^{n}$, such that the set $[K(x, \eta_1^p), \ldots, K(x, \eta_n^p)]$ is an optimal subspace for the n-widths $d_n(\mathfrak{K}_p; L^\infty)$ and $\delta_n(\mathfrak{K}_p; L^\infty)$, and interpolation at $\xi_1, \ldots, \xi_n$ (independent of p) is an optimal method of approximating $\mathfrak{K}_p$.*

REMARK 4.2. We have concerned ourselves, in Section 3, with a particular type of optimal subspace. Many other optimal subspaces may of course be constructed (see (3.1)–(3.4)). For example, it is possible, under the conditions of Theorem 3.1 and since $\bar{\phi}$ and $\bar{\psi}$ have n sign changes in (0, 1), to choose

$$d\bar{\alpha}_i = d\mu_{\tau_i} + d\mu_{\tau_{i+1}} \qquad , i = 1 \cdots, n$$

$$d\bar{\beta}_i = d\mu_{\zeta_i} + d\mu_{\zeta_{i+1}} \qquad , i = 1, \ldots, n$$

for some $0 \leqslant \tau_1 < \cdots < \tau_{n+1} \leqslant 1$, $0 \leqslant \zeta_1 < \cdots < \zeta_{n+1} \leqslant 1$. If $q = \infty$, we may further choose the $\{\tau_i\}_1^{n+1}$, as the points of equioscillation of $\bar{\phi}$.

REMARK 4.3. The results of Section 3 may be extended to weak Chebyshev systems. The major difficulty in such an extension is to prove the rank condition given in (3.3) and (3.4). This condition may be verified provided that the convexity cone associated with the weak Chebyshev system $\{\phi_1, \ldots, \phi_{n+1}\}$, contains a weak Chebyshev system of dimension at least n. For details of this type of argument and the meaning of the condition on the convexity cone, see [6].

REMARK 4.4. Note that no claim of uniqueness has been made in Theorem 3.1. However, if p or $q = \infty$, then uniqueness (by which we mean uniqueness up to a multiplicative constant), does in fact obtain. For example, if $p = \infty$, then the $\{\bar{b}_i\}_{i=1}^{n+1}$ are determined independently of q, and for fixed $\{|\bar{b}_i|\}_1^{n+1}$, the $\{\bar{a}_i\}_1^{n+1}$ are unique by Proposition 3.1. In addition, if $p = q = 2$, then we have from (3.6)

$$\begin{aligned} \int_0^1 \bar{\Phi}(x)\phi_k(x)\,dx &= \left[\int_0^1 |\bar{\Phi}|^2\,dx\right]\bar{b}_k, \qquad k = 1, \ldots, n+1 \\ \int_0^1 \bar{\Psi}(y)\psi_k(y)\,dy &= \left[\int_0^1 |\ \psi|^2\,dy\right]\bar{a}_k, \qquad k = 1, \ldots, n+1. \end{aligned} \tag{4.1}$$

By definition, $K(x,y) = \sum_{k=1}^{n+1} \phi_k(x)\psi_k(y)$. Now set

$$KK^T(x,z) = \int_0^1 K(x,y)K(z,y)\,dy = \sum_{k,j=1}^{n+1} \phi_k(x)\phi_j(z)\left[\int_0^1 \psi_k(y)\psi_j(y)\,dy\right].$$

An application of (4.1) yields

$$\int_0^1 KK^T(x,z)\bar{\Phi}(z)\,dz = \bar{\Phi}(x)\left[\int_0^1 |\bar{\Phi}(x)|^2\,dx\right]\left[\int_0^1 |\bar{\Psi}(y)|^2\,dy\right],$$

i.e., $\bar{\Phi}(x)$ is an eigenfunction of $KK^T(x,z)$ with eigenvalue $[\int_0^1 |\bar{\Phi}(x)|^2\,dx] \cdot [\int_0^1 |\bar{\Psi}(y)|^2\,dy]$. $KK^T(x,z)$ is a positive semidefinite kernel of rank $n+1$ and $[\int_0^1 |\bar{\Phi}(x)|^2\,dx][\int_0^1 |\bar{\Psi}(y)|^2\,dy]$ is its smallest nonzero eigenvalue (see Kolmogorov [3], and Melkman and Micchelli [4]). Thus, if we can show that this eigenvalue has multiplicity 1, then $\bar{\Phi}(x)$ is unique. The Chebyshevian conditions on the $\{\phi_i\}_1^{n+1}$ and $\{\psi_i\}_1^{n+1}$ are sufficient to prove the result.

REMARK 4.5. Let A be any $(n+1)\times(n+1)$ nonsingular matrix. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two vector norms on R^{n+1} and let

$$d_k(A; X_1; X_2) = \inf_{Y_k} \sup_{\|x\|_1 \leq 1} \inf_{y \in Y_k} \|Ax - y\|_2,$$

where Y_k is any k-dimensional subspace of R^{n+1}. The following is a simple consequence of the previously mentioned result of Brown [1].

Lemma 4.1. $d_n(A; X_1; X_2)d_0(A^{-1}; X_2; X_1) = 1$.

PROOF.

$$d_n(A; X_1; X_2) = \min_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_1} = \left[\max_{x\neq 0} \frac{\|x\|_1}{\|Ax\|_2}\right]^{-1} = \left[\max_{y\neq 0} \frac{\|A^{-1}y\|_1}{\|y\|_2}\right]^{-1}$$
$$= [d_0(A^{-1}; X_2; X_1)]^{-1}.$$

If we let $\|\cdot\|_p$ denote the usual l_p norm in R^{n+1}, then from the above lemma $d_n(A; l_p; l_q) = [d_0(A^{-1}; l_q; l_p)]^{-1}$.

From Section 2 we in fact obtain $d_n(A; l_p; l_q) = \delta_n(A; l_p; l_q) = E^n_{p',q}(A)$, where $\delta_n(A; l_p; l_q)$ is as earlier defined, and

$$E^n_{p',q}(A) = \inf_{\substack{B \\ \operatorname{rank} B \leqslant n}} \|A - B\|_{p',q}$$

where B is any $(n + 1) \times (n + 1)$ matrix of rank $\leqslant n$, and

$$\|C\|_{p',q} = \left[\sum_{j=1}^{n+1} \left(\sum_{i=1}^{n+1} |c_{ij}|^q\right)^{p'/q}\right]^{1/p'}.$$

One may obtain Lemma 4.1 via the analysis of Sections 2 and 3 by, for example, the choice of $\psi_i(y) = X[(i-1)/(n+1), i/(n+1)](y)$, $i = 1, \ldots, n+1$, and $\phi_i(x) = \sum_{k=1}^{n+1} a_k X[(k-1)/(n+1), k/(n+1)](x)$, where $X_{[a,b]}$ is the usual characteristic function of the interval $[a, b]$. Hence,

$$\min \frac{\|\sum_{i=1}^{n+1} x_i\phi_i\|_q \|\sum_{i=1}^{n+1} y_i\psi_i\|_{p'}}{\sum_{i=1}^{n+1} |x_i||y_i|} = \min_{x_i, y_i} \frac{\|Ax\|_q \|y\|_{p'}}{|(x, y)|} \cdot \frac{1}{(n+1)^2}$$
$$= \frac{1}{(n+1)^2} \min_{x\neq 0} \frac{\|Ax\|_q}{\|x\|_p}. \quad \square$$

REMARK 4.6. The problem of determining $\min_{x\neq 0}(\|Ax\|_q)/(\|x\|_p)$ where A is an $(n + 1) \times (n + 1)$ nonsingular matrix cannot, in general, be put in a more closed form. However, in certain cases, something more may be established.

Case I. Let $A = \operatorname{diag}[\lambda_1, \ldots, \lambda_{n+1}]$, i.e., A is an $(n + 1) \times (n + 1)$ diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_{n+1}$. We assume that $|\lambda_i| > 0$, $i = 1, \ldots, n + 1$, i.e., that A is nonsingular. Then

Proposition 4.2. For A as above,

$$d_n(A; l_p; l_q) = \begin{cases} \min\limits_{i=1,\ldots,n+1} |\lambda_i|, & 1 \leqslant q \leqslant p \leqslant \infty \\ \left(\sum\limits_{i=1}^{n+1} |\lambda_i|^r\right)^{1/r}, & 1 \leqslant p < q \leqslant \infty \end{cases}$$

where $1/r = 1/q - 1/p < 0$.

PROOF. By previous remarks, for $p, q \neq \infty$,

$$d_n(A; l_p; l_q) = \min_{x \neq 0} \frac{\|Ax\|_q}{\|x\|_p}$$

$$= \min_{x \neq 0} \frac{\left(\sum\limits_{j=1}^{n+1} |\lambda_j x_j|^q\right)^{1/q}}{\left(\sum\limits_{j=1}^{n+1} |x_j|^p\right)^{1/p}}$$

$$= \min_{\substack{\tau_j \geqslant 0 \\ \sum_1^{n+1} \tau_j = 1}} \left(\sum_{j=1}^{n+1} |\lambda_j|^q \tau_j^{q/p}\right)^{1/q}$$

$$= \left(\min_{\substack{\tau_j \geqslant 0 \\ \sum_1^{n+1} \tau_j = 1}} \sum_{j=1}^{n+1} |\lambda_j|^q \tau_j^{q/p}\right)^{1/q}.$$

Now, if $q/p \leqslant 1$, then, since $\sum_{j=1}^{n+1} |\lambda_j|^q \tau_j^{q/p}$ is concave in $\boldsymbol{\tau}$, the minimum must occur at an extreme point and thus

$$d_n(A; l_p; l_q) = \min_{i=1,\ldots,n+1} |\lambda_i|.$$

If $q/p > 1$, then $\sum_{j=1}^{n+1} |\lambda_j|^q \tau_j^{q/p}$ is strictly convex and thus the minimum occurs at an interior point of the simplex, $\tau_i \geqslant 0$, $\sum_1^{n+1} \tau_i = 1$. By Lagrange multipliers, the $\{\tau_i\}_{i=1}^{n+1}$ must satisfy

$$|\lambda_j|^q \tau_j^{q/p-1} = \mu, \qquad j = 1, \ldots, n$$

$$\sum_{j=1}^{n+1} \tau_j = 1.$$

Hence,

$$\tau_j = \frac{|\lambda_j|^r}{\left[\sum\limits_{i=1}^{n+1} |\lambda_i|^r\right]}$$

and therefore,

$$\left(\sum_{j=1}^{n+1} |\lambda_j|^q \tau_j^{q/p}\right)^{1/q} = \left[\frac{\sum_{j=1}^{n+1} |\lambda_j|^q |\lambda_j|^{r\cdot q/p}}{\left(\sum_{i=1}^{n+1} |\lambda_i|^r\right)^{q/p}}\right]^{1/q}$$

$$= \left[\frac{\sum_{j=1}^{n+1} |\lambda_j|^r}{\left(\sum_{j=1}^{n+1} |\lambda_j|^r\right)^{q/p}}\right]^{1/q} = \left[\sum_{j=1}^{n+1} |\lambda_j|^r\right]^{1/r}.$$

Case II. If $p = q = 2$, then $d_n(A; l_2; l_2) = \lambda_{n+1}^{1/2}$, where λ_{n+1} is the smallest positive eigenvalue of the positive definite matrix $A^T A$. This follows from

$$\min_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \left[\min_{x\neq 0} \frac{(Ax,Ax)}{(x,x)}\right]^{1/2} = \left[\min_{x\neq 0} \frac{(A^T Ax,x)}{(x,x)}\right]^{1/2} = [\lambda_{n+1}]^{1/2},$$

by the min–max characterization of eigenvalues for symmetric matrices.

Case III. Let A be such that all $n \times n$ minors of A are nonnegative, and of course, $\det A = |A| \neq 0$. (This is the analogue to the Chebyshevian requirements on $\{\phi_i\}_1^{n+1}$ and $\{\psi_i\}_1^{n+1}$).

Let $A\binom{1,\ldots,\hat{\imath},\ldots,n+1}{1,\ldots,\hat{\jmath},\ldots,n+1}$ denote the determinant of the $n \times n$ submatrix of A obtained by deleting the ith row and jth column. Then the following hold:

1.

$$d_n(A; l_p; l_\infty) = \frac{|A|}{\left(\sum_{j=1}^{n+1}\left[\sum_{i=1}^{n+1} A\binom{1\cdots,\hat{\imath},\ldots,n+1}{1\cdots,\hat{\jmath},\ldots,n+1}\right]^p\right)^{1/p}}$$

2.

$$d_n(A; l_1; l_q) = \frac{|A|}{\left(\sum_{i=1}^{n+1}\left[\sum_{j=1}^{n+1} A\binom{1,\ldots,\hat{\imath},\ldots,n+1}{1,\ldots,\hat{\jmath},\ldots,n+1}\right]^{q'}\right)^{1/q'}},$$

where $1/q + 1/q' = 1$.

3.

$$d_n(A; l_\infty; l_q) = \min_j \frac{|A|}{\left(\sum_{i=1}^{n+1}\left[A\begin{pmatrix}1,\ldots,\hat{\imath},\ldots,n+1\\1,\ldots,\hat{\jmath},\ldots,n+1\end{pmatrix}\right]^{q'}\right)^{1/q'}}$$

4.

$$d_n(A; l_p; l_1) = \min_i \frac{|A|}{\left(\sum_{j=1}^{n+1}\left[A\begin{pmatrix}1,\ldots,\hat{\imath},\ldots,n+1\\1,\ldots,\hat{\jmath},\ldots,n+1\end{pmatrix}\right]^{p}\right)^{1/p}}.$$

REMARK 4.7. It is worth noting that (2.1) may be simplified in the case that $\|\cdot\|_1 = \|\cdot\|_2$ and $\phi_i = \psi_i$. For here it is easily shown that it is possible to choose $\bar{a}_i = \bar{b}_i$, $i = 1, \ldots, n+1$, since

$$\min_{\sum_{i=1}^{n+1} a_i b_i = 1} \left\|\sum_{i=1}^{n+1} a_i \phi_i\right\|_1 \left\|\sum_{i=1}^{n+1} b_i \phi_i\right\|_2 \geqslant \min_{a_i}\left[\frac{\left\|\sum_{i=1}^{n+1} a_i\phi_i\right\|_1}{\left(\sum_{i=1}^{n+1} a_i^2\right)^{1/2}}\right]^2$$

and equality is, in fact, obtained. Thus, if $\{\phi_i\}_{i=1}^{n+1}$ satisfies the usual Chebyshevian properties of Section 3, then it follows that in $d_n(\mathfrak{K}_p; L^{p'})$, the nodes which determine the optimal subspace in Theorem 3.2, and the points of interpolation, are one and the same.

References

1. A. L. Brown, Best n-dimensional approximation to sets of functions, *Proc. London Math. Soc.* 14 (1964), 577–594.
2. S. Karlin, *Total Positivity Vol. I*, Stanford Univ. Press, Stanford, CA, 1968.
3. A. Kolmogorov, Uber die beste Annaherung von Funktionen einer gegebenen Functionenklasse, *Annals of Math.* 37 (1936), 107–110.
4. A. A. Melkman and C. A. Micchelli, Spline spaces are optimal for L^2 n-widths, *Illinois J. Math.* 22 (1978), 541–564.
5. C. A. Micchelli and A. Pinkus, Some problems in the approximation of functions of two variables and n-widths of integral operators, *J. Approx. Th.*, to be published.
6. C. A. Micchelli, Best L^1 approximation by weak Chebyshev systems and the uniqueness of interpolating perfect splines, *J. Approx. Th.* 19 (1977), 1–14.
7. A. Pietsch, s-Numbers of operators in Banach spaces, *Studia Math.* 51 (1974), 201–223.

8. E. Schmidt, Zur theorie der linearen und nichtlinearen integralgleichungen, I., *Math. Annalen*, 63 (1907), 433–476.
9. V. M. Tichomirov, Best methods of approximation and interpolation of differentiable functions in the space $C[-1, 1]$, *Math. USSR Sbornik* 9 (1969), 275–289.
10. V. M. Tikhomirov, Some problems in approximation theory, *Math. Notes* 9 (1971) 343–350.
11. H. Triebel, Interpolationseigenschaffen von entropie-und durchmesseridealen kompakter operatoren, *Studia Math.* 34 (1970), 89–107.