JOURNAL OF APPROXIMATION THEORY **24**, 51–77 (1978)

Some Problems in the Approximation of Functions of Two Variables and n-Widths of Integral Operators

CHARLES A. MICCHELLI*

IBM Research Center, Yorktown Heights, New York 10598

AND

ALLAN PINKUS*,†

Department of Mathematics, Technion–Israel Institute of Technology, Haifa, Israel

Communicated by Carl de Boor

Received February 15, 1977

We explicitly obtain, for $K(x, y)$ totally positive, a best choice of functions $u_1, ..., u_n$ and $v_1, ..., v_n$ for the problem $\min_{u_i, v_i} (\int_0^1 (\int_0^1 | K(x, y) - \sum_{i=1}^n u_i(x) v_i(y)| \, dy^p \, dx)^{1/p}$, where $u_i \in L^p[0, 1]$, $v_i \in L^1[0, 1]$, $i = 1, ..., n$, and $p \in [1, \infty]$. We show that an optimal choice is determined by certain sections $K(x, \xi_1), ..., K(x, \xi_n)$, and $K(\tau_1, y), ..., K(\tau_n, y)$ of the kernel K. We also determine the n-widths, both in the sense of Kolmogorov and of Gel'fand, and identify optimal subspaces, for the set $\mathscr{K}_{r,p} = \{f: f(x) = \sum_{i=1}^r a_i k_i(x) + \int_0^1 K(x, y) h(y) \, dy, (a_1, ..., a_r) \in \mathbb{R}^r, \| h \|_p \leq 1\}$, as a subset of $L^q[0, 1]$, with either $p = \infty$ and $q \in [1, \infty]$, or $p \in [1, \infty]$ and $q = 1$, where $\{k_1(x), ..., k_r(x), K(x, y)\}$ satisfy certain restrictions. A particular example is the ball $\mathscr{B}_{r,p} = \{f: f^{(r-1)}$ abs. cont. on $[0, 1]$, $\| f^{(r)} \|_p \leq 1\}$ in the Sobolev space.

1. INTRODUCTION

Our main motivation for this work is the classical result of Schmidt [11] (see also Courant and Hilbert [1, p. 161]) concerning the best approximation of an integral operator by finite rank operators. His problem begins with a real-valued kernel $K(x, y)$ in L^2 of the unit square $[0, 1] \times [0, 1]$. The Schmidt numbers of the associated integral operator

$$(Kf)(x) = \int_0^1 K(x, y) f(y) \, dy \tag{1.1}$$

* Sponsored by the United States Army under Contract DAAG29-75-C-0024.
† Sponsored by the National Science Foundation under Grant MCS75-17385.

0021-9045/78/0241–0051$02.00/0

are defined as the eigenvalues of the operator $K^T K$ ($K^T(x, y) = K(y, x)$, the transpose of K) given by

$$K^T K \phi_i = \lambda_i \phi_i, \; i = 1, 2, \ldots,$$

with

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots$$

and

$$(\phi_i, \phi_j) = \int_0^1 \phi_i(x)\, \phi_j(x)\, dx = \delta_{ij}, \qquad i, j = 1, 2, \ldots.$$

The Hilbert–Schmidt decomposition of $K(x, y)$ is

$$K(x, y) = \sum_{i=1}^{\infty} \psi_i(x)\, \phi_i(y) \qquad \text{a.e.}, \tag{1.2}$$

where $\psi_i = K\phi_i$.

E. Schmidt proved that the best mean square approximation to $K(x, y)$ on the square $[0, 1] \times [0, 1]$ by functions of the form

$$u_1(x)\, v_1(y) + \cdots + u_n(x)\, v_n(y),\; u_i, v_i \in L^2[0, 1] \tag{1.3}$$

is obtained by simply truncating the sum (1.2) after the nth term and the error of approximation is $(\sum_{n+1}^{\infty} \lambda_j)^{1/2}$. In other words,

$$\begin{aligned} \min_{u_i, v_i} \int_0^1 \int_0^1 \Big| K(x, y) - \sum_{i=1}^{n} u_i(x)\, v_i(y) \Big|^2 dx\, dy &= \int_0^1 \int_0^1 \Big| K(x, y) - \sum_{i=1}^{n} \psi_i(x)\, \phi_i(y) \Big|^2 dx\, dy \\ &= \sum_{n+1}^{\infty} \lambda_j. \end{aligned}$$

This result, in the language of operator theory, states that for the *trace norm* on the class of Hilbert–Schmidt operators (1.1), given by

$$|K|_{2,2} = (\operatorname{trace}(K^T K))^{1/2} = \left(\sum_{1}^{\infty} \lambda_j \right)^{1/2},$$

the best rank n approximation to K is

$$K_n f = \sum_{i=1}^{n} \psi_i (f, \phi_i).$$

It is remarkable that this extremal property of the series (1.2) also remains true when we give K the usual *operator norm* defined by

$$\| K \|_{2,2} = \sup_{\|f\|_2 \leq 1} \| Kf \|_2 = \lambda_1^{1/2},$$

$\| f \|_2 = (\int_0^1 | f(x)|^2 \, dx)^{1/2}$. The fact that K_n is the best rank n approximation to K in the operator norm as well, is a familiar result on s-numbers of compact operators on Hilbert spaces, see Pietsch [9]. For the possibility of other choices of best rank n approximations to K in the operator norm see [4].

The problem of approximating real-valued functions $K(x, y)$ in various norms by sums of products of (real-valued) functions of one variable (see (1.3)) and its relationship to n-widths is the subject of this paper. In Section 2, we solve this problem for mean approximation

$$| K |_{1,1} = \int_0^1 \int_0^1 | K(x, y)| \, dx \, dy.$$

We find that a best choice of functions $u_1, \ldots, u_n$, and $v_1, \ldots, v_n$ is determined by certain sections $K(x, \xi_1), \ldots, K(x, \xi_n)$ and $K(\tau_1, y), \ldots, K(\tau_n, y)$ of the kernel K, provided that K is a *nondegenerate* totally positive kernel. This result includes the case announced earlier in [7].

In Section 3, we consider the n-widths of certain subsets of L^p. In particular, for the Sobolev space $W_p{}^r$ defined by

$$W_p{}^r = \Big\{ f : f(x) = \sum_{j=0}^{r-1} a_j x^j + \frac{1}{(r-1)!} \int_0^1 (x - y)_+^{r-1} f^{(r)}(y) \, dy,$$

$$(a_0, a_1, \ldots, a_{r-1}) \in \mathbb{R}^r, \| f^{(r)} \|_p < \infty \Big\},$$

we compute the n-width in the sense of Kolmogorov and Gel'fand and identify optimal subspaces for the set

$$\mathscr{B}_{r,p} = \{ f : f \in W_p{}^r, \| f^{(r)} \|_p \leq 1 \}$$

considered as a subset of $L^q[0, 1]$ with either $p = \infty$ and $q \in [1, \infty]$, or $p \in [1, \infty]$ and $q = 1$. Recall that the Kolmogorov n-width is defined by

$$d_n(\mathscr{B}_{r,p}; L^q[0, 1]) = \inf_{X_n} \sup_{f \in \mathscr{B}_{r,p}} \inf_{g \in X_n} \| f - g \|_q,$$

where X_n is any n-dimensional linear subspace of $L^q[0, 1]$, and the Gel'fand n-width is defined by

$$d^n(\mathscr{B}_{r,p}; L^q[0, 1]) = \inf_{L_n} \sup_{f \in \mathscr{B}_{r,p} \cap L_n} \| f \|_q,$$

where L_n is any subspace of $L^q[0, 1]$ of codimension n.

In Section 4, we return to a discussion of the 2-dimensional approximation problem considered in Section 2, but for mixed (L^p, L^q) norms. Lower bounds for the error are given in terms of certain Kolmogorov n-widths of the integral operator (1.1). The results of Section 3 allow us to show that these lower bounds are sometimes attained. In particular, under the assumptions of Section 2 on the kernel K, we are able to obtain a best choice of functions $u_1,\ldots, u_n$ and $v_1,\ldots, v_n$ for the problem

$$\min_{u_i, v_i} \left(\int_0^1 \left(\int_0^1 \left| K(x, y) - \sum_{i=1}^{n} u_i(x)\, v_i(y)\right| dy\right)^p dx\right)^{1/p}$$

where $u_i \in L^p[0, 1]$, $v_i \in L^1[0, 1]$, $i = 1,\ldots, n$, $p \in (1, \infty]$. As in Section 2, an optimal choice is determined by certain sections $K(x, \xi_1),\ldots, K(x, \xi_n)$ and $K(\tau_1, y),\ldots, K(\tau_n, y)$ of the kernel K. The results of this section extend those discussed in Section 2.

2. Mean Approximation

In this section we find

$$E_{1,1}(K) = \min_{u_i, v_i} \int_0^1 \int_0^1 \left| K(x, y) - \sum_{i=1}^{n} u_i(x)\, v_i(y)\right| dx\, dy, \tag{2.1}$$

where the minimum is taken over $u_i, v_i \in L^1[0, 1]$, $i = 1,\ldots, n$, and identify an optimal choice of functions for a certain class of kernels K.

Definition 2.1. A real-valued kernel $K(x, y)$ defined and continuous on $[0, 1] \times [0, 1]$ is called totally positive if all its Fredholm minors

$$K\begin{pmatrix} s_1, \ldots, s_m \\ t_1, \ldots, t_m \end{pmatrix} = \begin{vmatrix} K(s_1, t_1) & \cdots & K(s_1, t_m) \\ \vdots & & \vdots \\ K(s_m, t_1) & \cdots & K(s_m, t_m) \end{vmatrix}$$

are nonnegative for $0 \leqslant s_1 < \cdots < s_m \leqslant 1$, $0 \leqslant t_1 < \cdots < t_m \leqslant 1$, and all $m \geqslant 1$.

Our first theorem below requires a condition on $K(x, y)$ which is stronger than total positivity. This theorem deals with an extremum problem whose solution is guaranteed in a closed simplex. However, we wish to assert that the extremum actually lies within the interior of the simplex. Thus to avoid the possibility that it occurs on the boundary of the simplex we require K to be *nondegenerate* totally positive. Before defining this requirement on K, let us state the theorem.

To this end we define, $\Lambda_n = \{s: s = (s_1, ..., s_n), 0 = s_0 < s_1 < \cdots < s_n < s_{n+1} = 1\}$, the step function

$$h_s(x) = (-1)^i, s_i \leqslant x < s_{i+1}, i = 0, 1, ..., n,$$

and also let $\|f\|_1 = \int_0^1 |f(x)| \, dx$.

THEOREM 2.1. *Let K be a nondegenerate totally positive kernel. Given any $n \geqslant 1$, there exists a $\xi \in \Lambda_n$, such that for any $t \in \Lambda_n$,*

$$\| Kh_\xi \|_1 \leqslant \| Kh_t \|_1 .$$

Moreover, Kh_ξ has exactly n distinct zeros in $(0, 1)$ at $\tau_1, ..., \tau_n$, with $\tau = (\tau_1, ..., \tau_n) \in \Lambda_n$ and

$$\operatorname{sgn} Kh_\xi = h_\tau , \tag{2.2}$$

$$\operatorname{sgn} K^T h_\tau = h_\xi . \tag{2.3}$$

(When Kh_ξ or $K^T h_\tau$ are zero in (2.2) or (2.3) we assign a value to the sgn so that the equations are valid.)

The proof of this theorem requires information on the number of zeros of the function Kh_t. The basic fact needed is the following lemma (see [5, 6, 8] for similar results.)

LEMMA 2.1. *Let K be a totally positive kernel and $t \in \Lambda_m$, be given. If Kh_t vanishes at $s \in \Lambda_m$, then for any $x \in [0, 1]$ either $(-1)^i (Kh_j)(x) > 0$, where $s_i \leqslant x \leqslant s_{i+1}$, for some i, $0 \leqslant i \leqslant m$, or the functions $K(s_1, y), ..., K(s_m, y)$, $K(x, y)$ are linearly dependent on $[0, 1]$.*

Proof. If $x \in [0, 1] \backslash \{s_1, ..., s_m\}$ and $s_i \leqslant x \leqslant s_{i+1}$ for some i, $0 \leqslant i \leqslant m$, and $K(s_1, y), ..., K(s_i, y), K(x, y), K(s_{i+1}, y), ..., K(s_m, y)$ are linearly independent then there exist nontrivial constants $\alpha_1, ..., \alpha_{m+1}$ such that $\alpha_j(-1)^{j+1} \geqslant 0, j = 1, 2, ..., m + 1$, and the function

$$u(y) = \alpha_1 K(s_1, y) + \cdots + \alpha_i K(s_i, y) + \alpha_{i+1} K(x, y) + \cdots + \alpha_{m+1} K(s_m, y)$$

satisfies $(-1)^i u(y) \geqslant 0$, $t_i \leqslant y \leqslant t_{i+1}$, $i = 0, 1, ..., m$. This fact is proven by "smoothing" the kernel K so that the functions $K(s_1, y), ..., K(s_i, y), K(x, y), ..., K(s_m, y)$ form a complete Chebyshev system. We will not go into the details of this standard technique. Let us observe that the function $u(y)$ is nontrivial by virtue of the linear independence of $K(s_1, y), ..., K(s_i, y)$, $K(x, y), ..., K(s_m, y)$.

We use the function $u(y)$ as follows:

$$\begin{aligned}\alpha_{i+1}(Kh_t)(x) &= \alpha_1(Kh_t)(s_1) + \cdots + \alpha_i(Kh_t)(s_i) + \alpha_{i+1}(Kh_t)(x) \\ &\quad + \cdots + \alpha_{m+1}(Kh_t)(s_m) \\ &= \int_0^1 u(y)\, h_t(y)\, dy = \int_0^1 |u(y)|\, dy > 0.\end{aligned}$$

Thus the lemma is proven.

The relationship of the zeros of Kh_t to linear dependence as expressed above, leads us to

DEFINITION 2.2. A totally positive kernel K is called nondegenerate totally positive on [0, 1] provided that

1. For every $m \geqslant 1$ and every choice of t-point, $t \in \Lambda_m$, and s-point $s \in \Lambda_m$, the sets of function $\{K(s_1, y),..., K(s_m, y)\}$, $\{K(x, t_1),..., K(x, t_m)\}$ are linearly independent on [0, 1].

2. For every $m \geqslant 0$ and every choice of t-point and s-point, as above, one of the four sets of functions $\{K(0, y), K(s_1, y),..., K(s_m, y)\}$, $\{K(s_1, y),..., K(s_m, y), K(1, y)\}$, $\{K(x, 0), K(x, t_1),..., K(x, t_m)\}$, $\{K(x, t_1),..., K(x, t_m), K(x, 1)\}$ is linearly independent.

Note that whenever K is nondegenerate totally positive then so is K^T. The kernel $K(x, t) = (x - t)_+^{r-1}$, $r \geqslant 2$, is nondegenerate totally positive, see [12]. However, the totally positive kernel

$$\begin{aligned} K(x, t) &= x(1 - t), \qquad 0 \leqslant x \leqslant t, \\ &= t(1 - x), \qquad t \leqslant x \leqslant 1, \end{aligned} \tag{2.4}$$

is not because it vanishes everywhere on the boundary of the unit square and hence Property 2 is not satisfied. Property 2 is needed to insure that zeros of Kh_t occurring at the ends of the interval [0, 1] may be taken into consideration. Property 1, which holds for (2.4), is insufficient for this purpose.

We draw the following conclusion from Lemma 2.1 which is necessary in the proof of Theorem 2.1.

LEMMA 2.2. *Let K be a nondegenerate totally positive kernel. Then for every $n \geqslant 0$ and t-point, $t \in \Lambda_n$, the function Kh_t has at most n zeros in* (0, 1). *If Kh_t has exactly n zeros at $s \in \Lambda_n$, then*

(i) *these zeros are sign changes,*

(ii) *the orientation of Kh_t is governed by the equation* $\operatorname{sgn} Kh_t = h_s$,

(iii) *at least one of the numbers* $(Kh_t)(0)$, $(Kh_t)(1)$, $(K^Th_s)(0)$, $(K^Th_s)(1)$ *is not zero.*

We are now ready to prove the theorem.

Proof. The minimum of the continuous function $F(t_1, \ldots, t_n) = \| Kh_t \|_1$ is achieved on the *closed* simplex $0 \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant 1$. Hence there are values $0 < \xi_1 < \cdots < \xi_p < 1$, $0 \leqslant p \leqslant n$, such that $\| Kh_\xi \|_1 \leqslant \| Kh_t \|_1$ for all $t = (t_1, \ldots, t_n)$. We claim that $p = n$. To prove this we observe that by Lemma 2.2, Kh_ξ has at most p distinct zeros in (0, 1). Hence F is a differentiable function and by the optimality of Kh_ξ we have

$$0 = \frac{\partial}{\partial t_l} F(t_1, \ldots, t_p)|_{t=\xi} = 2(-1)^{l+1} \int_0^1 \operatorname{sgn}(Kh_\xi)(x)\, K(x, \xi_l)\, dx, \qquad l = 1, \ldots, p. \tag{2.5}$$

Let $0 < \tau_1 < \cdots < \tau_m < 1$, $0 \leqslant m \leqslant p$, denote the location of the sign changes of Kh_ξ, and let μ be a sign, $\mu^2 = 1$, such that $\mu(Kh_\xi)(x)\, h_\tau(x) \geqslant 0$, $x \in [0, 1]$, $\tau = (\tau_1, \ldots, \tau_m)$. Then upon simplification (2.5) reads $(K^T h_\tau)(\xi_i) = 0$, $i = 1, \ldots, p$, and so Lemma 2.2 implies $p \leqslant m$. We conclude that $p = m$ and again by Lemma 2.2, $\operatorname{sgn} Kk_\xi = h_\tau$ and $\operatorname{sgn} K^T h_\tau = h_\xi$. We will now prove that $p = n$. The idea is to show that if $p < n$ then $(Kk_\xi)(0) = (Kh_\xi)(1) = (K^T h_\tau)(0) = (K^T h_\tau)(1) = 0$ which contradicts Lemma 2.2. Let us deal with the left hand endpoint as the argument for the right hand endpoint is similar. By Lemma 2.2 we know that $(Kh_\xi)(0) \geqslant 0$ and $(K^T h_\tau)(0) \geqslant 0$. Now, if $p < n$ then for all ϵ, $0 < \epsilon < \xi_1$,

$$\begin{aligned} \int_0^1 |(Kh_\xi)(x)|\, dx &= F(\xi_1, \ldots, \xi_p) \\ &\leqslant F(\epsilon, \xi_1, \ldots, \xi_p) \\ &= \int_0^1 \left| (Kh_\xi)(x) - 2 \int_0^\epsilon K(x, y)\, dy \right| dx. \end{aligned} \tag{2.6}$$

The function $P_\epsilon(x) = (Kh_\xi)(x) - 2 \int_0^\epsilon K(x, y)\, dy$ has, by Lemma 2.2, at most $p + 1$ zeros. Furthermore, for ϵ small, P_ϵ has p sign changes near the sign changes of $(Kh_\xi)(x)$ and slightly to the left of the first sign change of Kh_ξ, P_ϵ is positive (because Kh_ξ begins positively). Now, if P_ϵ has no more zeros in (0, 1), then $\operatorname{sgn} P_\epsilon = h_{\tau(\epsilon)}$, for some $0 < \tau_1(\epsilon) < \cdots < \tau_p(\epsilon) < 1$ and $\tau_i(\epsilon) \to \tau_i$ as $\epsilon \to 0^+$. Thus

$$\begin{aligned} F(\epsilon, \xi_1, \ldots, \xi_p) &= \int_0^1 h_{\tau(\epsilon)}(x)(Kh_\xi)(x) - 2 \int_0^\epsilon (K^T h_{\tau(\epsilon)})(y)\, dy \\ &\leqslant \int_0^1 |(Kh_\xi)(x)|\, dx - 2 \int_0^\epsilon (K^T h_{\tau(\epsilon)})(y)\, dy \\ &= F(\xi_1, \ldots, \xi_p) - 2 \int_0^\epsilon (K^T h_{\tau(\epsilon)})(y)\, dy. \end{aligned}$$

For ϵ sufficiently small, $K^T h_{\tau(\epsilon)}(y) > 0$, $0 < y < \epsilon$. This inequality contradicts (2.6) and we conclude that P_ϵ has exactly one more zero in (0, 1). Moreover, since $P_\epsilon = -Kh_{\xi(\epsilon)}$ where $\xi(\epsilon) = (\epsilon, \xi_1, ..., \xi_p)$, Lemma 2.2 implies that this zero is a sign change, that it must be the first sign change of P_ϵ and to the left of it P_ϵ is negative in (0, 1). Hence we conclude that for ϵ small, $P(0; \epsilon) \leqslant 0$. Thus $(Kh_\xi)(0) \leqslant 0$ and so it follows that $(Kh_\xi)(0) = 0$. Returning to (2.6) we have by an easy computation

$$0 \leqslant \lim_{\epsilon \to 0+} \epsilon^{-1}(F(\epsilon, \xi_1, ..., \xi_p) - F(\xi_1, ..., \xi_p)) = -2(K^T h_\tau)(0),$$

whence we conclude $(K^T h_\tau)(0) = 0$. We now apply the above analysis to the right hand endpoint to obtain $(Kh_\xi)(1) = (K^T h_\tau)(1) = 0$. This contradicts Lemma 2.2 and the theorem is proven.

Let us remark, that if K is totally positive and only Property 1 of Definition 2.2 is satisfied, then we may show that p defined above is $\geqslant n - 1$. This is accomplished by comparing $F(\xi_1, ..., \xi_p)$ to $F(\xi - \epsilon, \xi + \epsilon, \xi_1, ..., \xi_p)$ for any ξ, $0 < \xi < \xi_1$ and ϵ sufficiently small.

The following corollary, although not explicitly used in the solution of the mean approximation problem (2.1), is an expression of the symmetry of Theorem 2.1 under replacement of K by K^T.

COROLLARY 2.1. *Let $\tau = (\tau_1, ..., \tau_n)$ be the τ-point defined by Theorem* 2.1. *Then*

$$\| K^T h_\tau \|_1 \leqslant \| K^T h_s \|_1$$

for every $s \in \Lambda_n$.

Proof. Given any $s \in \Lambda_n$ there then exists a $t \in \Lambda_n$, such that $(Kh_t)(s_i) = 0$, $i = 1, 2, ..., n$, see [8].

Hence by Lemma 2.2, sgn $Kh_t = h_s$ and by the optimality of $\xi = (\xi_1, ..., \xi_n)$ given by Theorem 2.1, we have

$$\begin{aligned} \| K^T h_\tau \|_1 &= (K^T h_\tau, h_\xi) = \| Kh_\xi \|_1 \\ &\leqslant \| Kh_t \|_1 = (h_s, Kh_t) = (K^T h_s, h_t) \\ &\leqslant \| K^T h_s \|_1 . \end{aligned}$$

To proceed further we require one final lemma (see [5, 6, 8] for similar results).

LEMMA 2.3. *Let $\tau = (\tau_1, ..., \tau_n)$, $\xi = (\xi_1, ..., \xi_n)$ be defined by Theorem* 2.1. *Then*

$$K\begin{pmatrix} \tau_1, ..., \tau_n \\ \xi_1, ..., \xi_n \end{pmatrix} > 0.$$

Proof. Suppose to the contrary that there exist nontrivial constants $\alpha_1, ..., \alpha_n$ such that the function $u(x) = \sum_{i=1}^n \alpha_i K(x, \xi_i)$ vanishes at $\tau_1, ..., \tau_n$.

There exists a $z \in (0, 1)$ such that $u(z) \neq 0$. Thus for some constant c, $Kh_\xi - cu$ vanishes at $\tau_1, \ldots, \tau_n, z$. Let $v(y)$ be a nontrivial function, $v(y) = \beta_1 K(\tau_1, y) + \cdots + \beta_n K(\tau_n, y) + \beta_{n+1} K(z, y)$ such that $(-1)^i v(y) \geqslant 0$, $\xi_i \leqslant y \leqslant \xi_{i+1}$, $i = 0, 1, \ldots, n$. Hence

$$0 = \sum_{i=1}^{n} \beta_i((Kh_\xi)(\tau_i) - cu(\tau_i)) + \beta_{n+1}((Kh_\xi)(z) - cu(z))$$

$$= \int_0^1 v(y)\, h_\xi(y)\, dy - c \sum_{i=1}^{n} \alpha_i v(\xi_i)$$

$$= \int_0^1 |v(y)|\, dy > 0.$$

This contradiction proves the lemma.

We are now prepared to state and prove the main theorem of the section. To this end, observe that the function

$$E(x, y) = K\begin{pmatrix} x, \tau_1, \ldots, \tau_n \\ y, \xi_1, \ldots, \xi_n \end{pmatrix} \Big/ K\begin{pmatrix} \tau_1, \ldots, \tau_n \\ \xi_1, \ldots, \xi_n \end{pmatrix} \tag{2.7}$$

may be expressed as

$$= K(x, y) - \sum_{i,j=1}^{n} c_{ij} K(x, \xi_i)\, K(\tau_j, y),$$

where

$$c_{ij} = (-1)^{i+j} K\begin{pmatrix} \tau_1, \ldots, \tau_{j-1}, \tau_{j+1}, \ldots, \tau_n \\ \xi_1, \ldots, \xi_{i-1}, \xi_{i+1}, \ldots, \xi_n \end{pmatrix} \Big/ K\begin{pmatrix} \tau_1, \ldots, \tau_n \\ \xi_1, \ldots, \xi_n \end{pmatrix}.$$

Therefore,

$$E_{1,1}(K) = \min_{u_i, v_i} \int_0^1 \int_0^1 \left| K(x, y) - \sum_{i=1}^{n} u_i(x)\, v_i(y) \right| dx\, dy$$

$$\leqslant \int_0^1 \int_0^1 |E(x, y)|\, dx\, dy. \tag{2.8}$$

Actually, we have

THEOREM 2.2. *If K is a nondegenerate totally positive kernel and $E_{1,1}(K)$ is as defined in* (2.1), *then*

$$E_{1,1}(K) = \int_0^1 \int_0^1 |E(x, y)|\, dx\, dy = \| Kh_\xi \|_1$$

$$= \int_0^1 \int_0^1 \left| K(x, y) - \sum_{i=1}^{n} u_i^0(x)\, v_i^0(y) \right| dx\, dy,$$

where $u_i^0(x) = K(x, \xi_i)$ *and* $v_i^0(y) = \sum_{j=1}^{n} c_{ij} K(\tau_j, y)$, $i = 1,..., n$, *and the* $\{\xi_i\}_1^n$, $\{\tau_i\}_1^n$ *are as defined in Theorem* 2.1.

Proof. By the Hobby–Rice theorem [2], we know that given any n functions $v_1, ..., v_n \in L^1[0, 1]$ there exists a $t \in \Lambda_k$, $0 \leqslant k \leqslant n$, such that $\int_0^1 v_i(y)\, h_t(y)\, dy = 0$, $i = 1, 2,..., n$. Let $h(x, y) = h_t(y) \operatorname{sgn}(Kh_t)(x)$. Then for $u_1, ..., u_n \in L^1[0, 1]$,

$$\begin{aligned} \| Kh_\xi \|_1 &\leqslant \int_0^1 |(Kh_t)(x)|\, dx \\ &= \int_0^1 \int_0^1 \left(K(x, y) - \sum_{i=1}^{n} u_i(x)\, v_i(y) \right) h(x, y)\, dx\, dy \\ &\leqslant \int_0^1 \int_0^1 \left| K(x, y) - \sum_{i=1}^{n} u_i(x)\, v_i(y) \right| dx\, dy. \end{aligned}$$

Thus, since $u_1, ..., u_n, v_1, ..., v_n$ were arbitrarily chosen in $L^1[0, 1]$, we have

$$\| Kh_\xi \|_1 \leqslant E_{1,1}(K).$$

Also, we have, in view of (2.2), (2.3) and (2.7),

$$\begin{aligned} &\int_0^1 \int_0^1 | E(x, y)|\, dx\, dy \\ &\quad = \int_0^1 \int_0^1 E(x, y)\, h_\tau(x)\, h_\xi(y)\, dx\, dy \\ &\quad = \int_0^1 (Kh_\xi)(x)\, h_\tau(x)\, dx - \sum_{i,j} c_{ij} (K^T h_\tau)(\xi_i)(Kh_\xi)(\tau_i) \\ &\quad = \int_0^1 |(Kh_\xi)(x)|\, dx = \| Kh_\xi \|_1, \end{aligned}$$

which together with (2.8) finishes the proof.

This theorem states that the best approximation in the mean on the square $[0, 1] \times [0, 1]$ may be accomplished by interpolating $K(x, y)$ with the sections $K(\tau_i, y)$, $K(x, \xi_j)$ at (τ_i, ξ_j), $i, j = 1,..., n$.

The condition of nondegenerate total positivity was imposed so as to insure that $0 < \xi_1 < \cdots < \xi_n < 1$, and $0 < \tau_1 < \cdots < \tau_n < 1$. However, if $K(x, y)$ is only totally positive on $[0, 1] \times [0, 1]$, and if there exist $0 \leqslant x_1 < \cdots < x_n \leqslant 1$ and $0 \leqslant y_1 < \cdots < y_n \leqslant 1$ such that

$$K\begin{pmatrix} x_1, ..., x_n \\ y_1, ..., y_n \end{pmatrix} > 0$$

(if not, then $E_{1,1}(K) = 0$), then by "smoothing" $K(x, y)$, both with respect to x and y, it is possible to prove that $E_{1,1}(K) = \| Kh_\xi \|_1$, where as in Theorem 2.1, $\| Kh_\xi \|_1 = \inf_{0<\eta_1<\dots<\eta_n<1} \| Kh_1 \|_1$.

Specifically, the method of smoothing we have in mind replaces $K(x, y)$ by

$$K_\epsilon(x, y) = \int_0^1 \int_0^1 G_\epsilon(x, \sigma)\, K(\sigma, \tau)\, G_\epsilon(\tau, y)\, d\sigma\, d\tau,$$

where

$$G_\epsilon(x, y) = \frac{1}{\epsilon(2\pi)^{1/2}} \exp\left[-\frac{1}{2}\left(\frac{x-y}{\epsilon}\right)^2\right], \qquad \epsilon > 0.$$

Then K_ϵ is strictly totally positive (because G_ϵ is) and thus certainly satisfies the hypotheses of Theorems 2.1 and 2.2. Since K_ϵ converges to K as $\epsilon \downarrow 0$, the above assertion follows directly.

In the remainder of the paper we will show the relationship of the previous problem, as well as a general version of it for mixed (L^p, L^q) norms (see Section 4), to certain Gel'fand and Kolmogorov widths. As these results on widths are of independent interest we devote the next section to their discussion.

3. Widths

In this section we compute exactly the Kolmogorov and Gel'fand widths and identify optimal subspaces for certain subsets of $L^p[0, 1]$, $1 \leqslant p \leqslant \infty$. The norm of $f \in L^p$ is denoted by $\| f \|_p$ and p' is used to denote the conjugate exponent defined by $1/p + 1/p' = 1$.

We begin by recalling the definition of Kolmogorov and Gel'fand n-widths. Let X be a normed linear space, $\mathfrak{A}$ a subset of X, and X_n any n-dimensional linear subspace of X. Then, the n-width of $\mathfrak{A}$ relative to X, in the sense of Kolmogorov, is defined to be

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

and X_n is called an optimal subspace for $\mathfrak{A}$ provided that

$$d_n(\mathfrak{A}; X) = \sup_{x \in \mathfrak{A}} \inf_{y \in X_n} \| x - y \|$$

$$\equiv \delta(\mathfrak{A}; X_n).$$

The n-width of $\mathfrak{A}$ relative to X, in the sense of Gel'fand, is defined as

$$d^n(\mathfrak{A}; X) = \inf_{L_n} \sup_{x \in \mathfrak{A} \cap L_n} \| x \|,$$

where L_n is any subspace of X of codimension n. If

$$d^n(\mathfrak{A}; X) = \sup_{x \in \mathfrak{A} \cap L_n} \| x \| ,$$

then L_n is an optimal subspace for the Gel'fand n-width of $\mathfrak{A}$.

Our sets have the following form. Given functions $k_1(x),..., k_r(x)$ defined and continuous on $[0, 1]$, and a kernel $K(x, y)$ jointly continuous in $x, y \in [0, 1]$, we define

$$\mathscr{K}_{r,p} = \left\{ \sum_{j=1}^{r} a_j k_j(x) + \int_0^1 K(x, y)\, h(y)\, dy : (a_1 ,..., a_r) \in R^r, \| h \|_p \leqslant 1 \right\}. \quad (3.1)$$

The prototype of this class of sets is the choice $k_j(x) = x^{j-1}, j = 1,..., r$ and $K(x, y) = (1/(r-1)!)(x-y)_+^{r-1}$. In this case $\mathscr{K}_{r,p}$ is simply the ball

$$\mathscr{B}_{r,p} = \{f : f^{(r-1)} \text{ abs. cont.}, \| f^{(r)} \|_p \leqslant 1\}. \quad (3.2)$$

In the general case, we will consider $\mathscr{K}_{r,p}$ as a subset of $L^q[0, 1]$ for some q, $1 \leqslant q \leqslant \infty$, and as such compute its Kolmogorov and Gel'fand n-widths when certain addition hypotheses are satisfied.

For our purposes, in Section 4, where we study mixed (L^p, L^q) approximation to $K(x, y)$ by functions of the form (1.3), we will only need the results of this section when $r = 0$. However, for the sake of (3.2) we deal here with $r > 0$ as well and require that the following properties hold.

I.

$$K\begin{pmatrix} x_1 ,..., x_r , x_{r+1} ,..., x_{r+m} \\ 1,..., r, \quad y_1 ,..., y_m \end{pmatrix} = \begin{vmatrix} k_1(x_1) & \cdots & k_r(x_1) & K(x_1 , y_1) & \cdots & K(x_1 , y_m) \\ \vdots & & \vdots & \vdots & & \vdots \\ k_1(x_{r+m}) & \cdots & k_r(x_{r+m}) & K(x_{r+m} , y_1) & \cdots & K(x_{r+m} , y_m) \end{vmatrix}$$

is non-negative for any points $0 \leqslant y_1 < \cdots < y_m \leqslant 1$, $0 \leqslant x_1 < \cdots < x_{r+m} \leqslant 1$ and integer $m \geqslant 0$. Furthermore, we require that for any given y-point $0 < y_1 < \cdots < y_m < 1$ (x-point, $0 < x_1 < \cdots < x_{r+m} \leqslant 1$) the above determinant is not identically zero for all x-points (y-points).

II. $\{k_i(x)\}_{i=1}^r$ is a Chebyshev system on $(0, 1)$, i.e., for any $0 < x_1 < \cdots < x_r < 1$,

$$K\begin{pmatrix} x_1 ,..., x_r \\ 1,..., r \end{pmatrix} > 0.$$

In particular, when $r = 0$, Property I implies that K is a nondegenerate totally positive kernel since Property I above implies that the functions

$K(x_1, y), \ldots, K(x_m, y), K(1, y)$ are linearly independent on $[0, 1]$. This property is more restrictive than the requirement of nondegenerate total positivity and we could relax the hypotheses I and II somewhat in what follows. However, for us it is important that these properties hold for the special case (3.2), see [12], and they shall always be assumed to hold in this section.

3.1. *Kolmogorov n-width*, $p = \infty$, $1 \leqslant q \leqslant \infty$

Our objective is to find

$$d_n(\mathscr{K}_{r,\infty}; L^q[0, 1]) = \inf_{X_n} \sup_{f \in \mathscr{K}_{r,\infty}} \inf_{g \in X_n} \| f - g \|_q .$$

The computation of the n-width when $q = \infty$ was done in [5] and so we here restrict ourselves to considering $q < \infty$.

We introduce the class

$$\mathscr{P}_n = \{k + Kh_t : t \in \Lambda_p , 0 \leqslant p \leqslant n, k \in Q_r\},$$

where $Q_r = [k_1, \ldots, k_r]([f_1, \ldots, f_m] =$ the linear space spanned by $f_1, \ldots, f_m)$. A typical element of $\mathscr{P}_n$ will be denoted by P or by P_t. Thus if $P_t \in \mathscr{P}_n$, then $P_t = k + Kh_t$ for some $k \in Q_r$.

THEOREM 3.1. *Given integers $m, r \geqslant 0$ and a number q, $1 \leqslant q < \infty$, then there exists $\xi \in \Lambda_m$ and $k \in Q_r$ such that $P_\xi^* = k + Kh_\xi$ satisfies*

$$\| P_\xi^* \|_q \leqslant \| P \|_q , \tag{3.3}$$

for every $P \in \mathscr{P}_m$. Moreover, P_ξ^* has exactly $m + r$ simple zeros in $(0, 1)$ *at* $0 < \tau_1 < \cdots < \tau_{m+r} < 1$ *and hence*

$$\operatorname{sgn} P_\xi^*(x) = (-1)^r h_\tau(x), \tag{3.4}$$

$$\operatorname{sgn} \left(\int_0^1 | P_\xi^*(x)|^{q-1} h_\tau(x) K(x, y)\, dx \right) = (-1)^r h_\xi(y), \tag{3.5}$$

and

$$\int_0^1 | P_\xi^*(x)|^{q-1} h_\tau(x) k_i(x)\, dx = 0, \qquad i = 1, \ldots, r. \tag{3.6}$$

Note that when $r = 0$, $q = 1$, and $m = n$, this theorem reduces to Theorem 2.1. The proof of the general case follows the proof of Theorem 2.1 with only slight modifications. The details require the following generalized versions of Lemma 2.2.

LEMMA 3.1. *For given $m, r \geqslant 0$, $P \in \mathscr{P}_m$ has at most $m + r$ zeros in* $(0, 1)$. *If P has exactly $m + r$ zeros at $s \in \Lambda_{m+r}$, then these zeros are sign changes, the orientation of P is governed by the equation* $\operatorname{sgn} P = (-1)^r h_s$, *and* $P(1) \neq 0$.

Proof. Let $P = k + Kh_t$, $k \in Q_r$, $t \in \Lambda_m$. Assume P has at least r zeros (otherwise there is nothing to prove) and let $\mathbf{s} = (s_1, \ldots, s_r)$, $0 < s_1 < \cdots < s_r < 1$. Then it follows that $P = Jh_t$, where $J(x, y)$ is the compound kernel

$$J(x, y) = K\begin{pmatrix} s_1, \ldots, s_r, x \\ 1, \ldots, r, y \end{pmatrix} \Big/ K\begin{pmatrix} s_1, \ldots, s_r \\ 1, \ldots, r \end{pmatrix} .$$

Now, the kernel $\bar{J}(x, y) = (-1)^r h_{\mathbf{s}}(x) J(x, y)$ is totally positive, because Sylvester's determinant identity tells us that if $0 \leqslant x_1 < \cdots < x_l \leqslant 1$, $0 \leqslant y_1 < \cdots < y_l \leqslant 1$, then

$$\bar{J}\begin{pmatrix} x_1, \ldots, x_l \\ y_1, \ldots, y_l \end{pmatrix} = K\begin{pmatrix} z_1, \quad \ldots \quad, z_{l+r} \\ 1, \ldots, r, y_1, \ldots, y_l \end{pmatrix} \Big/ K\begin{pmatrix} s_1, \ldots, s_r \\ 1, \ldots, r \end{pmatrix},$$

where $0 \leqslant z_1 < \cdots < z_{l+r} \leqslant 1$ are the points of the set $\{s_1, \ldots, s_r, x_1, \ldots, x_l\}$ arranged in increasing order. (Note that $\bar{J}(s_i, y) \equiv 0$, $i = 1, \ldots, r$.) Now, if P also vanishes at $0 < s_{r+1} < \cdots < s_{r+m} < 1$, (say $s_r < s_{r+1}$), then it follows directly from Lemma 2.1 and Property I that $(-1)^i (\bar{J}h_t)(x) > 0$ for $x \in (s_{r+i}, s_{r+i+1})$, $i = 1, \ldots, m$, $(\bar{J}h_t)(1) \neq 0$, and $(\bar{J}h_t)(x) > 0$ for $x \in (s_i, s_{i+1})$, $i = 0, 1, \ldots, r$. These facts immediately imply the results of the lemma.

We need another lemma similar to Lemma 3.1 which also reduces to Lemma 2.2 in the case $r = 0$.

LEMMA 3.2. *Given $t \in \Lambda_m$ and $g(x) \in L^\infty[0, 1]$ such that* $\operatorname{sgn} g = h_t$. *Assume that $(g, k_i) = 0$, $i = 1, \ldots, r$. Then $m \geqslant r$, and $K^T g$ has at most $m - r$ zeros in* $(0, 1)$. *If $K^T g$ has $m - r$ zeros at $s \in \Lambda_{m-r}$ then the zeros are sign changes and* $\operatorname{sgn} K^T g = (-1)^r h_s$.

Proof. The fact that $(g, k_i) = 0$, $i = 1, \ldots, r$, implies that g has at least r sign changes is a well-known result obtained from the Chebyshev property of $\{k_i(x)\}_1^r$. The remaining proof is quite similar to that of Lemma 2.1. Assume $(K^T g)(s_i) = 0$, $i = 1, \ldots, m - r$. Since $k_1(x), \ldots, k_r(x)$, $K(x, s_1), \ldots, K(x, s_{m-r})$, $K(x, y)$ are linearly independent for $y \in (0, 1)\backslash\{s_1, \ldots, s_{m-r}\}$, there exists a nontrivial linear combination $u(x) = \sum_{i=1}^r a_i k_i(x) + \sum_{i=1}^{m-r} b_i K(x, s_i) + cK(x, y)$ such that $u(x) h_t(x) \geqslant 0$, $x \in [0, 1]$. Since $c(K^T g)(y) = (u, g) > 0$, it follows that $(K^T g)(y) \neq 0$ for $y \in (0, 1)\backslash\{s_1, \ldots, s_{m-r}\}$. It is easily shown, by determining the sign of c, that $\operatorname{sgn}(K^T g) = (-1)^r h_s$ in $(0, 1)$.

We are now ready to prove Theorem 3.1.

Proof. The existence of a minimum $P_\xi^* = k + Kh_\xi$, where $\xi = (\xi_1,..., \xi_p)$, $0 < \xi_1 < \cdots < \xi_p < 1$, $0 \leqslant p \leqslant m$, follows directly. Using the minimality of ξ we have that

$$\int_0^1 | P_\xi^*(x)|^{q-1} \operatorname{sgn} P_\xi^*(x)\, k_i(x)\, dx = 0, \qquad i = 1,..., r$$

and

$$\int_0^1 | P_\xi^*(x)|^{q-1} \operatorname{sgn} P_\xi^*(x)\, K(x, \xi_i)\, dx = 0, \qquad i = 1,..., p.$$

Let $0 < \tau_1 < \cdots < \tau_l < 1$, $l \geqslant r$, be the location of the sign changes of P_α^* on (0, 1). Then according to Lemma 3.1, $l \leqslant p + r$, while Lemma 3.2 implies that $l \geqslant p + r$. Thus $l = r + p$ and by Lemma 3.2, $\operatorname{sgn} P_\xi^* = (-1)^r h_\tau$. Moreover, if $p < m$ then the function $P_{\xi(\epsilon)}$, $\xi(\epsilon) = (\xi_1,..., \xi_p, 1 - \epsilon)$ may be compared to P_ξ^* for ϵ small and positive to show, as in the proof of Theorem 2.1, that $P_\xi^*(1) = 0$, This conclusion contradicts Lemma 3.1 and hence $p = m$.

We now turn to the computation of the Kolmogorov n-width of $\mathscr{K}_{r,\infty}$.

Let r and q be as given, and apply Theorem 3.1 for each $n \geqslant r$ with $m = n - r$, to obtain points $0 < \xi_1 < \cdots < \xi_{n-r} < 1$, $0 < \tau_1 < \cdots < \tau_n < 1$ and a function P_ξ^* which satisfies (3.3)-(3.6). Since P_ξ^* plays a distinguished role in computing the n-width of $\mathscr{K}_{r,\infty}$ we give it the special designation $g_{n,r,q}(x)$. We will also use the notation $g_n(x)$ for $g_{n,r,q}(x)$ suppressing its dependence on r and q. In addition, we define the n-dimensional subspace

$$X_n{}^0 = [k_1,..., k_r, K(\cdot, \xi_1),..., K(\cdot, \xi_{n-r})].$$

THEOREM 3.2.

$$\begin{aligned} d_n(\mathscr{K}_{r,\infty}; L^q[0, 1]) &= \infty, & n \leqslant r,\\ &= \| g_n \|_q, & n \geqslant r, \end{aligned}$$

and for $n \geqslant r$, $X_n{}^0$ is an optimal subspace for the n-width of $\mathscr{K}_{r,\infty}$.

Proof. Since the subspace Q_r spanned by $k_1,..., k_r$ is contained in $\mathscr{K}_{r,\infty}$, the n-width of $\mathscr{K}_{r,\infty}$, when $n < r$, must be ∞. Now, suppose $n \geqslant r$. We will first prove that $\| g_n \|_q$ is a lower bound for the n-width. We proceed as follows: The only n-dimensional subspaces in contention for approximating $\mathscr{K}_{r,\infty}$ are those which contain Q_r. Let X_n be such a subspace and assume for the moment that $q > 1$. Let X_n be spanned by the functions $k_1,..., k_r, u_1,..., u_{n-r}$.

For every $z = (z_1,..., z_{n-r+1})$ with $\sum_{i=1}^{n-r+1} z_i{}^2 = 1$, we define $t_0(z) = 0$, $t_i(z) = \sum_{j=1}^{i} z_j{}^2$, $i = 1, 2,..., n - r + 1$ and $f_z(y) = f(y; z) = \operatorname{sgn} z_j$, for $t_{j-1}(z) < y < t_j(z)$, $j = 1, 2,..., n - r + 1$. Note that $f_z(y) = \pm h_s(y)$ for some

$s \in \Lambda_k$, $0 \leqslant k \leqslant n - r$. Moreover, $f_z(y, -z) = -f(y, z)$ for all z and y. (This particular odd embedding of the surface of the $n - r + 1$ sphere into the set of extreme points of the unit ball in L^∞ is used in [10] to simplify the proof of the Hobby–Rice theorem [2].)

The function Kf_z has a *unique* best approximation in $L^q[0, 1]$ from the subspace X_n (because $1 < q < \infty$) which we denote by

$$\sum_{i=1}^{r} \alpha_i(z)\, k_i + \sum_{i=1}^{n-r} \beta_i(z)\, u_i\, .$$

Thus the mapping $(z_1, \ldots, z_{n-r+1}) \to (\beta_1(z), \ldots, \beta_{n-r}(z))$ is a continuous odd mapping defined on the $n - r + 1$ sphere $S^{n-r} = \{z : \sum_{i=1}^{n-r+1} z_i^2 = 1\}$. Hence, by the Borsuk Antipodality Theorem (cf. [3]), there is a $z_0 \in S^{n-r}$ for which $\beta_i(z_0) = 0$, $i = 1, 2, \ldots, n - r$. Moreover, by the definition of g_n we have

$$\begin{aligned} \| g_n \|_q &\leqslant \Big\| Kf_{z_0} - \sum_{i=1}^{r} \alpha_i(z_0)\, k_i \Big\|_q \\ &\leqslant \| Kf_{z_0} - v \|_q \end{aligned}$$

for every function $v \in X_n$. Letting $q \to 1^+$ we have, for all q, $1 \leqslant q < \infty$

$$\| g_n \|_q \leqslant \sup_{f \in \mathscr{K}_{r,\infty}} \inf_{v \in X_n} \| f - v \|_q\, .$$

Since X_n was chosen arbitrarily we obtain the desired conclusion,

$$\| g_n \|_q \leqslant d_n(\mathscr{K}_{r,\infty}\, ;\, L^q[0, 1]).$$

The proof of the upper bound for the n-width requires

LEMMA 3.3. *Let* $0 < \xi_1 < \cdots < \xi_{n-r} < 1$, $0 < \tau_1 < \cdots < \tau_n < 1$, *be the points given by Theorem* 3.1 *corresponding to* g_n. *Then*

$$K\begin{pmatrix} \tau_1, & \cdot & \cdot & \cdot & , \tau_n \\ 1, \ldots, r, & \xi_1 & , \ldots, & \xi_{n-r} \end{pmatrix} > 0.$$

The proof of this lemma is similar to the proof of Lemma 2.3. We omit the details (see [5, 6, 8] for related results). Using this lemma we define the unique linear interpolation operator S from $C[0, 1]$ onto X_n^0 by the condition that

$$(Sf)(\tau_i) = f(\tau_i), \qquad i = 1, \ldots, n.$$

We shall show that $\sup_{f \in \mathscr{K}_{r,\infty}} \| f - Sf \|_q \leqslant \| g_n \|_q$ and since $d_n(\mathscr{K}_{r,\infty}\, ;\, L^q\, [0, 1]) \leqslant \sup_{f \in \mathscr{K}_{r,\infty}} \| f - Sf \|_q$, this will prove the theorem.

To this end, observe that if $f \in \mathscr{K}_{r,\infty}$ has the representation $f = k + Kh$ for some $k \in Q_r$ and $\| h \|_\infty \leqslant 1$ then

$$f(x) - (Sf)(x) = \int_0^1 \frac{K\begin{pmatrix} \tau_1, & \cdot & \cdot & \cdot & \cdot & , \tau_n, x \\ 1,\ldots, r, \xi_1, \ldots, \xi_{n-r}, y \end{pmatrix}}{K\begin{pmatrix} \tau_1, & \cdot & \cdot & \cdot & \cdot & , \tau_n \\ 1,\ldots, r, \xi_1, \ldots, \xi_{n-r} \end{pmatrix}} h(y)\, dy.$$

Therefore

$$\sup_{f \in \mathscr{K}_{r,\infty}} \| f - Sf \|_q^q \leqslant \int_0^1 \left(\int_0^1 \frac{\left| K\begin{pmatrix} \tau_1, & \cdot & \cdot & \cdot & \cdot & , \tau_n, x \\ 1,\ldots, r, \xi_1, \ldots, \xi_{n-r}, y \end{pmatrix} \right|}{K\begin{pmatrix} \tau_1, & \cdot & \cdot & \cdot & \cdot & , \tau_n \\ 1,\ldots, r, \xi_1, \ldots, \xi_{n-r} \end{pmatrix}}\, dy \right)^q dx$$

$$= \int_0^1 \left| \int_0^1 \frac{K\begin{pmatrix} \tau_1, & \cdot & \cdot & \cdot & \cdot & , \tau_n, x \\ 1,\ldots, r, \xi_1, \ldots, \xi_{n-r}, y \end{pmatrix}}{K\begin{pmatrix} \tau_1, & \cdot & \cdot & \cdot & \cdot & , \tau_n \\ 1,\ldots, r, \xi_1, \ldots, \xi_{n-r} \end{pmatrix}} h_\xi(y)\, dy \right|^q dx$$

and because $g_n = P_\xi = k + Kh_\xi$ for some $k \in Q_r$

$$= \| g_n - Sg_n \|_q^q = \| g_n \|_q^q .$$

The last equality follows since $g_n(\tau_i) = 0$, $i = 1,\ldots, n$, and hence $Sg_n = 0$. Thus the theorem is proven.

We now turn to the computation of the Gel'fand n-width of $\mathscr{K}_{r,\infty}$.

3.2. *Gel'fand n-width*, $p = \infty$, $1 \leqslant q \leqslant \infty$.

The case $q = \infty$ was done in [5]. We again assume that $q < \infty$ and define, for $n \geqslant r$, the subspace

$$L_n^0 = \{f : f \in C[0, 1], f(\tau_i) = 0, i = 1, 2,\ldots, n\}.$$

L_n^0 is a subspace of $C[0, 1]$ and since $Sf = 0$, if $f \in L_n^0$, the proof of Theorem 3.2 implies that

$$\sup_{f \in L_n^0 \cap \mathscr{K}_{r,\infty}} \| f \|_q \leqslant \| g_n \|_q .$$

This inequality does not give an upper bound for the Gel'fand n-width of $\mathscr{K}_{r,\infty}$ since by definition

$$d^n(\mathscr{K}_{r,\infty}; L^q[0, 1]) = \inf_{L_n} \sup_{f \in L_n \cap \mathscr{K}_{r,\infty}} \| f \|_q , \tag{3.7}$$

where the infimum is taken over all subspaces of $L^q[0, 1]$ of codimension n.

Clearly, L_n^0 does not fit this requirement. However, let us "smooth" L_n^0 slightly to

$$L_n^0(\epsilon) = \left\{f : f \in L^q[0, 1], \int_{\tau_i}^{\tau_i+\epsilon} f(x)\, dx = 0,\ i = 1, 2,..., n\right\}.$$

For $\epsilon > 0$, ϵ small, define

$$f(x) - (S_\epsilon f)(x) = \int_0^1 R(x, y : \epsilon)\, h(y)\, dy,$$

where

$$R(x, y : \epsilon) = \frac{\int_{\tau_1}^{\tau_1+\epsilon} \cdots \int_{\tau_n}^{\tau_n+\epsilon} K\begin{pmatrix}\sigma_1, & . & . & . & , \sigma_n, x\\ 1,..., r, & \xi_1, ..., & & & \xi_{n-r}, y\end{pmatrix} d\sigma_1 \cdots d\sigma_n}{\int_{\tau_1}^{\tau_1+\epsilon} \cdots \int_{\tau_n}^{\tau_n+\epsilon} K\begin{pmatrix}\sigma_1, & . & . & . & , \sigma_n\\ 1,..., r, & \xi_1, ..., & & & \xi_{n-r}\end{pmatrix} d\sigma_1 \cdots d\sigma_n}.$$

$S_\epsilon f$ is the unique element in X_n^0 such that

$$\int_{\tau_i}^{\tau_i+\epsilon} (f - S_\epsilon f)(x)\, dx = 0, \qquad i = 1, 2,..., n.$$

When $\epsilon = 0$, $S_\epsilon = S$ and $\| S_\epsilon f - Sf \|_q \leqslant \max_{x,y} | R(x, y; \epsilon) - R(x, y; 0) |\, \| h \|_\infty$. Thus

$$\sup_{f \in L_n^0(\epsilon) \cap \mathscr{K}_{r,\infty}} \| f \|_q \leqslant \| g_n \|_q + \max_{x,y} | R(x, y : \epsilon) - R(x, y : 0) |.$$

The expression $\max_{x,y} | R(x, y; \epsilon) - R(x, y; 0)|$ goes to zeros as $\epsilon \to 0^+$ and thus $\| g_n \|_q$ does provide an upper bound for $d^n(\mathscr{K}_{r,\infty}\,; L^q[0, 1])$.

The fact that $\| g \|_q$ is a lower bound for the Gel'fand n-width is proven in a fashion similar to the proof of Theorem 3.2. The argument goes as follows: if L_n is a subspace of codimension n of $L^q[0, 1]$ with $\sup_{f \in L_n^0 \cap \mathscr{K}_{r,\infty}} \| f \|_q < \infty$, then $L_n \cap Q_r = \{0\}$. Thus if

$$L_n = \{f : f \in L^q[0, 1], (u_i\,, f) = 0, i = 1,..., n\},$$

where $u_1\,,..., u_n$ are linearly independent functions contained in $L^{q'}[0, 1]$, then the matrix $((u_i\,, k_j))$ has full rank r. We may assume without loss of generality that $\det((u_i\,, k_j))_{i,j=1,...,r} \neq 0$. Setting

$$N(x, y) = \left.\begin{vmatrix} K(x, y) & k_1(x) & \cdots & k_r(x) \\ (u_1\,, K(\cdot, y)) & (u_1\,, k_1) & \cdots & (u_1\,, k_r) \\ \vdots & \vdots & & \cdot \\ (u_r\,, K(\cdot, y)) & (u_r\,, k_1) & \cdots & (u_r\,, k_r) \end{vmatrix} \middle/ \begin{vmatrix} (u_1\,, k_1) & \cdots & (u_1\,, k_r) \\ \vdots & & \vdots \\ (u_r\,, k_1) & \cdots & (u_r\,, k_r) \end{vmatrix}\right., \tag{3.8}$$

then $f \in L_n \cap \mathscr{K}_{r,\infty}$, i.e., $f = k + Kh \in L_n$, for some $k \in Q_r$, $\| h \|_\infty \leqslant 1$, if and only if $f = Nh$ and $(v_i, h) = 0$, $i = r+1,..., n$ where $v_i = N^T u_i$. Now, by the Hobby–Rice theorem, [2], there is an h_s, $s = (s_1,..., s_k)$, $0 \leqslant k \leqslant n - r$, such that $(v_i, h_s) = 0$, $i = r + 1,..., n$. Hence $f_0 \equiv k + Kh_s \in L_n$ for some $k \in Q_r$. Therefore we conclude by the minimality property of g_n that

$$\| g_n \|_q \leqslant \| f_0 \|_q \leqslant \sup_{f \in L_n \cap \mathscr{K}_{r,\infty}} \| f \|_q .$$

Since L_n was an arbitrary subspace of codimension n of $L^q[0, 1]$, we finally obtain

$$\| g_n \|_q = d^n(\mathscr{K}_{r,\infty}; L^q[0, 1]).$$

Incidentally, we may in the proof of the lower bound allow L_n to be chosen from the larger class of subspaces of codimension n of $C[0, 1]$ and still obtain the same result. Perhaps, it is best that we extend the definition of the Gel'fand n-width to make this remark precise.

For a subset $\mathfrak{A}$ of a normed linear space $(X, \| \cdot \|)$ and a set $\mathfrak{F}$ of linear functionals defined on $\mathfrak{A}$ we let the Gel'fand n-width of $\mathfrak{A}$ relative to X and $\mathfrak{F}$ be

$$d^n(\mathfrak{A}; X, \mathfrak{F}) = \inf_{L_n} \sup_{x \in L_n} \| x \|$$

where $L_n = \{x : x \in \mathfrak{A}, F_i x = 0, i = 1,..., n\}$ and the infimum is taken over all $F_1,..., F_n \in \mathfrak{F}$. If $\mathfrak{F} = X^*$ (norm dual of X) then from our previous definition

$$d^n(\mathfrak{A}; X, X^*) = d^n(\mathfrak{A}; X).$$

Thus we may conveniently summarize our previous remarks in

THEOREM 3.3.

$$d^n(\mathscr{K}_{r,\infty}; L^q[0, 1]) = d^n(\mathscr{K}_{r,\infty}; L^q[0, 1], C^*[0, 1]) = \begin{cases} \infty, & n < r, \\ \| g_n \|_q, & n \geqslant r, \end{cases}$$

and for $n \geqslant r$, L_n^0 *is an optimal subspace (of* $C^*[0, 1]$*) for the Gel'fand* n-*width of* $\mathscr{K}_{r,\infty}$.

3.3. *Kolmogorov n-width,* $1 \leqslant p \leqslant \infty$, $q = 1$

The case $p = 1$ was previously done in [6]. Although $p = \infty$ was done in Section 3.1 the following discussion also holds in this case. Thus we assume $1 < p \leqslant \infty$. For this problem we need

THEOREM 3.4. *Given any n, r with $n \geqslant r$ and p, $1 < p \leqslant \infty$, there exists an $\eta \in \Lambda_n$ such that for any $t \in \Lambda_n$ satisfying the condition*

$$(k_i, h_t) = 0, \; i = 1,..., r, \tag{3.9}$$

we have

$$\| K^T h_\eta \|_{p'} \leqslant \| K^T h_t \|_{p'} . \tag{3.10}$$

Moreover, $t = \eta$ satisfies (3.9) *and $K^T h_\eta$ has exactly $n - r$ distinct zeros in* (0, 1) *at $\zeta \in \Lambda_{n-r}$. Hence*

$$\operatorname{sgn} K^T h_\eta = (-1)^r h_\zeta .$$

The proof of this theorem is similar to the proofs of Theorems 2.1 and 3.1. We omit the details.

We are now ready to compute n-widths. Let us first define

$$X_n^1 = [k_1 ,..., k_r , K(\cdot, \zeta_1),..., K(\cdot, \zeta_{n-r})]$$

and

$$L_n^1 = \{f : f \in C[0, 1], f(\eta_i) = 0, i = 1, 2,..., n\}.$$

THEOREM 3.5.

$$\begin{aligned} d_n(\mathscr{K}_{r,p}; L^1[0, 1]) &= \infty, && n < r, \\ &= \| K^T h_\eta \|_{p'}, && n \geqslant r, \end{aligned}$$

and for $n \geqslant r$, X_n^1 is an optimal subspace for the n-width of $\mathscr{K}_{r,p}$.

Proof. We first prove the lower bound. Let X_n be any n-dimensional subspace of $L^1[0, 1]$ such that $\delta(\mathscr{K}_{r,p}; X_n) < \infty$. Then $Q_r \subseteq X_n$ and by the Hobby-Rice Theorem there exists a $t \in \Lambda_k$, $0 \leqslant k \leqslant n$, such that the norm one linear functional $F(y) = (y, h_t)$ annihilates X_n . Thus we conclude that

$$\delta(\mathscr{K}_{r,p}; X_n) \geqslant \sup_{f \in \mathscr{K}_{r,p}} |(f, h_t)|$$

and keeping in mind that $Q_r \subseteq X_n$ this simplifies to

$$\begin{aligned} \delta(\mathscr{K}_{r,p}; X_n) &\geqslant \| K^T h_t \|_{p'} \\ &\geqslant \| K^T h_\eta \|_{p'}. \end{aligned}$$

The arbitrariness of X_n implies that the desired lower bound is valid.

The reverse inequality requires

LEMMA 3.4.

$$K \begin{pmatrix} \eta_1 , & . & . & . & , \eta_n \\ 1,..., r, & \zeta_1 ,..., & \zeta_{n-r} \end{pmatrix} > 0.$$

The proof of this lemma is similar to that of Lemma 2.3 (see [5, 6, 8] for similar results). Therefore we may define an interpolation operator $T: C[0, 1] \to X_n^1$ by the conditions

$$(Tf)(\eta_i) = f(\eta_i),\ i = 1,\dots, n.$$

Then as in Theorem 3.2, if $f = k + Kh$, $k \in Q_r$, $\| h \|_p \leqslant 1$, we have

$$f(x) - (Tf)(x) = \int_0^1 \frac{K\begin{pmatrix}\eta_1, & \cdot & \cdot & \cdot & \cdot & , \eta_n, x\\ 1,\dots, r, & \zeta_1, & \dots, & \zeta_{n-r}, & y\end{pmatrix}}{K\begin{pmatrix}\eta_1, & \cdot & \cdot & \cdot & , \eta_n\\ 1,\dots, r, & \zeta_1, & \dots, & \zeta_{n-r}\end{pmatrix}}\, h(y)\, dy$$

and

$$\sup_{f \in \mathscr{K}_{r,p}} \| f - Tf \|_1$$

$$\leqslant \left(\int_0^1 \left(\int_0^1 \frac{\left| K\begin{pmatrix}\eta_1, & \cdot & \cdot & \cdot & \cdot & , \eta_n, x\\ 1,\dots, r, & \zeta_1, & \dots, & \zeta_{n-r}, & y\end{pmatrix}\right|}{K\begin{pmatrix}\eta_1, & \cdot & \cdot & \cdot & , \eta_n\\ 1,\dots, r, & \zeta_1, & \dots, & \zeta_{n-r}\end{pmatrix}}\, dx \right)^{p'} dy \right)^{1/p}$$

$$= \left(\int_0^1 \left(\left| \int_0^1 \frac{K\begin{pmatrix}\eta_1, & \cdot & \cdot & \cdot & \cdot & , \eta_n, x\\ 1,\dots, r, & \zeta_1, & \dots, & \zeta_{n-r}, & y\end{pmatrix}}{K\begin{pmatrix}\eta_1, & \cdot & \cdot & \cdot & , \eta_n\\ 1,\dots, r, & \zeta_1, & \dots, & \zeta_{n-r}\end{pmatrix}}\, h_\eta(x)\, dx \right| \right)^{p'} dy \right)^{1/p'}.$$

Since $(k_i, h_\eta) = 0$, $i = 1,\dots, r$, and $(K(\cdot, \zeta_i), h_\eta) = 0$, $i = 1, 2,\dots, n - r$, the above simplifies to

$$\sup_{f \in \mathscr{K}_{r,p}} \| f - Tf \|_1 \leqslant \| K^T h_\eta \|_{p'}.$$

Thus

$$d_n(\mathscr{K}_{r,p}; L^1[0, 1]) = \sup_{f \in \mathscr{K}_{r,p}} \| f - Tf \|_1 = \| K^T h_\eta \|_{p'}.$$

Finally, we have

3.4. *Gel'fand n-width,* $1 \leqslant p \leqslant \infty$, $q = 1$

Again $p = 1$ was done in [6] while $p = \infty$ is included in Subsection 3.2. We assume here that $1 < p \leqslant \infty$.

THEOREM 3.6.

$$\begin{aligned} d^n(\mathscr{K}_{r,p}; L^1[0, 1]) &= \infty, && n < r,\\ &= \| K^T h_\eta \|_{p'}, && n \geqslant r, \end{aligned}$$

and for $n \geqslant r$, L_n^1 *is an optimal subspace for the n-width of* $\mathscr{K}_{r,p}$.

Actually (see the proof of Theorem 3.3) L_n^1 is a "nearly" optimal subspace.

Proof. The upper bound

$$\sup_{f\in L_n^1\cap\mathscr{K}_{r,p}} \|f\|_1 \leqslant \sup_{f\in\mathscr{K}_{r,p}} \|f - Tf\|_1 \leqslant \|K^T h_n\|_{p'},$$

follows from the proof of Theorem 3.5.

For the lower bound, we let L_n be any subspace of finite codimension n of $L^1[0, 1]$ such that $\sup_{f\in L_n\cap\mathscr{K}_{r,p}} \|f\|_1 < \infty$. Hence

$$L_n \cap Q_r = \{0\} \qquad \text{and} \qquad L_n = \{f : f \in L^1[0, 1], (u_i, f) = 0, i = 1,\dots, n\}$$

for some linearly independent functions $u_1,\dots, u_n \in L^\infty[0, 1]$. Let $N(x, y)$ be as defined in (3.8) and set $v_i = N^T u_i$. The lower bound argument given in Section 3.1 may be modified to prove that there is an $s = (s_1,\dots, s_k)$, $0 \leqslant k \leqslant n$, such that $(k_i, h_s) = 0$, $i = 1,\dots, r$, and

$$\| K^T h_1 \|_{p'} = \min_{\alpha_{r+1},\dots,\alpha_n} \Big\| K^T h_1 - \sum_{i=r+1} \alpha_i v_i \Big\|_{p'}.$$

To accomplish this, let $f_z(x)$ be as in Section 3.1 for $z \in S^n = \{z = (z_1,\dots, z_{n+1}): \sum_{i=1}^{n+1} z_i^2 = 1\}$. For $1 < p' < \infty$, let $\alpha_{r+1}(z),\dots, \alpha_n(z)$ be the unique coefficients in the best $L^{p'}$ approximation to $K^T f_z$ from the subspace spanned by $v_{r+1},\dots, v_n$. We define an odd, continuous mapping of S^n into R^n by $z \to ((k_1, f_z),\dots, (k_r, f_z), \alpha_{r+1}(z),\dots, \alpha_n(z))$ and again apply the Borsuk Antipodality Theorem and obtain a $z_0 \in S^n$ for which $(k_i, f_z) = 0$, $i = 1,\dots, r$, $\alpha_i(z_0) = 0$, $i = r + 1,\dots, n$. Then $h_s = \pm f_z$ serves our purpose.

Since the best $L^{p'}$ approximation to $K^T h_s$ by the subspace spanned by $v_{r+1},\dots, v_n$ is zero, we necessarily have the orthogonality relations $(g, v_i) = 0$, $i = r + 1,\dots, n$ where $g = \operatorname{sgn} K^T h_s \, | K^T h_s |^{p'-1}$. Let $w = g/\| g \|_p$. Then $w \in L^p[0, 1]$ with $\| w \|_p = 1$ and $f_0 = Nw \in L_n \cap \mathscr{K}_{r,p}$ (see the discussion in Section 3.2). Hence

$$\sup_{f\in L_n\cap\mathscr{K}_{r,p}} \|f\|_1 \geqslant \|f_0\|_1 \geqslant (h_s, f_0)$$

and because $(k_i, h_s) = 0$, $i = 1,\dots, r$, we have

$$= (K^T h_s, w) = \| K^T h_s \|_{p'} \geqslant \| K^T h_1 \|_{p'}.$$

Letting $p' \to 1^+$ completes the proof.

As was indicated, the prototype for the class of sets considered in this section is $k_j(x) = x^{j-1}$, $j = 1,\dots, r$ and $K(x, y) = (1/(r - 1)!)(x - y)_+^{r-1}$ since, in this case, $\mathscr{K}_{r,p}$ is simply a ball of the Sobolev space. We specialize below the results of this section for this specific class of functions.

DEFINITION. A perfect spline on [0, 1] of degree r with m knots $\{\xi_i\}_{i=1}^m$, $0 = \xi_0 < \xi_1 < \cdots < \xi_m < \xi_{m+1} = 1$, is any function $P(x)$ of the form

$$P(x) = \sum_{i=0}^{r-1} a_i x^i + c \sum_{j=0}^{m} (-1)^j \int_{\xi_j}^{\xi_{j+1}} (x-y)_+^{r-1}\, dy,$$

where, as usual, $x_+{}^r = x^r$ if $x \geqslant 0$, and zero otherwise.

Let $\mathscr{P}_m$ denote the class of perfect splines of degree r with at most m knots with $|P^{(r)}(x)| = 1$ a.e. on [0, 1], and let $Q_m = \{P : P \in \mathscr{P}_m,\ P^{(i)}(0) = P^{(i)}(1) = 0,\ i = 0, 1,\ldots, r-1\}$. Theorems 3.1 and 3.4 reduce to the following

COROLLARY 3.1. *Let $1 \leqslant p < \infty$, and $P_{m,p} \in \mathscr{P}_m$ be any perfect spline which attains $\min_{P\in\mathscr{P}_m} \| P \|_p$. Then $P_{m,p}$ has m distinct knots in* (0, 1), *and exactly $m + r$ zeros in* (0, 1), *each one a sign change.*

COROLLARY 3.2. *Let $1 \leqslant p < \infty$ and $m \geqslant r$, and let $Q_{m,p} \in Q_m$ be any perfect spline which attains $\min_{Q\in Q_m} \| Q \|_p$. Then $Q_{m,p}$ has m distinct knots in* (0, 1) *and exactly $m - r$ zeros in* (0, 1), *each one a sign change.*

Let $\mathscr{B}_{r,p} = \{f : f^{(r-1)}$ abs. cont., $\| f^{(r)} \|_p \leqslant 1\}$. Then from Theorems 3.2 and 3.3 we have

COROLLARY 3.3. *For $1 \leqslant q < \infty$,*

$$d_n(\mathscr{B}_{r,\infty}; L^q[0,1]) = d^n(\mathscr{B}_{r,\infty}; L^q[0,1]) = \begin{cases} \infty, & n < r, \\ \| P_{n-r,q} \|_q, & n \geqslant r, \end{cases}$$

and for $n \geqslant r$,

(i) $X_n{}^0 = [1, x,\ldots, x^{r-1}, (x-\xi_1)_+^{r-1},\ldots, (x-\xi_{n-r})_+^{r-1}]$, *where the $\{\xi_i\}_{i=1}^{n-r}$ are the knots of $P_{n-r,q}$, is an optimal subspace for the n-width d_n.*

(ii) $L_n{}^0 = \{f : f \in C[0,1], f(\tau_i) = 0, i = 1,\ldots, n\}$, *where the $\{\tau_i\}_{i=1}^n$ are the sign changes of $P_{n-r,q}$, is an optimal subspace for the n-width d^n.*

From Theorems 3.5 and 3.6, we have

COROLLARY 3.4. *For $1 < p \leqslant \infty$,*

$$d_n(\mathscr{B}_{r,p}; L^1[0,1]) = d^n(\mathscr{B}_{r,p}; L^1[0,1]) = \begin{cases} \infty, & n < r, \\ \| Q_{n,p'} \|_{p'}, & n \geqslant r, \end{cases}$$

where $1/p + 1/p' = 1$, and for $n \geqslant r$

(i) $X_n{}^1 = [1, x,\ldots, x^{r-1}, (x-\zeta_1)_+^{r-1},\ldots, (x-\zeta_{n-r})_+^{r-1}]$, *where the $\{\zeta_i\}_{i=1}^{n-r}$ are the sign changes of $Q_{n,p'}$, is an optimal subspace for the n-width d_n.*

(ii) $L_n^1 = \{f: f \in C[0, 1], f(\eta_i) = 0, i = 1,..., n\}$, *where* $\{\eta_i\}_{i=1}^n$ *are the knots of* $Q_{n,p'}$, *is an optimal subspace for the n-width* d^n.

Note that by setting $q = 1$ in Corollary 3.3 and $p = \infty$ in Corollary 3.4, it follows that $\| P_{n-r,1} \|_1 = \| Q_{n,1} \|_1$ and the knots of P_{n-r-1} may be taken as the sign changes of $Q_{n,1}$ and vice versa.

4. Mixed (L^p, L^q) Norms

Let

$$| K |_{p,q} = \left(\int_0^1 \left(\int_0^1 | K(x, y)|^q \, dy \right)^{p/q} dx \right)^{1/p}$$

where $1 \leqslant p < \infty$, $1 \leqslant q < \infty$. If $q = \infty$ and/or $p = \infty$, then the usual definitions apply. We use, as before, the pairing $(u, v) = \int_0^1 u(x)\, v(y)\, dy$ for $u \in L^p$, $v \in L^{p'}$, $1/p + 1/p' = 1$.

We study

$$E_{p,q}^n(K) = \inf \left\{ \left| K - \sum_{i=1}^n u_i \otimes v_i \right|_{p,q} : u_1 ,..., u_n \in L^p[0, 1], \right. \\ \left. v_1 ,..., v_n \in L^q[0, 1] \right\}, \tag{4.1}$$

where

$$(u_i \otimes v_j)(x, y) = u_i(x)\, v_j(y),$$

and shall make use of the results of Section 3 with $r = 0$. For convenience, $\mathscr{K}_{0,p}$ shall be denoted by $\mathscr{K}_p$. Thus

$$\mathscr{K}_p = \{Kh : \| h \|_p \leqslant 1\}.$$

Also, let

$$\mathscr{K}_p^T = \{K^T h : \| h \|_p \leqslant 1\}.$$

THEOREM 4.1. $\max\{d_n(\mathscr{K}_{q'}; L^p[0, 1]), d_n(\mathscr{K}_{p'}^T; L^q[0, 1])\} \leqslant E_{p,q}^n(K)$.

Before proving this theorem let us observe that the above n-widths, when $n = 0$, are given by

$$\begin{aligned} d_0(\mathscr{K}_{q'}; L^p[0, 1]) &= d_0(\mathscr{K}_{p'}^T; L^q[0, 1]) \\ &= \sup_{\|h\|_{q'} \leqslant 1} \| Kh \|_p \\ &= \| K \|_{p,q} . \end{aligned}$$

The right-hand side is the *operator norm* of K as an integral operator acting on $L^{q'}[0, 1]$ into $L^p[0, 1]$. Now, by Hölder's inequality, for $h \in L^{q'}[0, 1]$, $g \in L^{p'}[0, 1]$

$$\begin{aligned} |(Kh, g)| &= \left| \int_0^1 \int_0^1 g(x)\, K(x, y)\, h(y)\, dx\, dy \right| \\ &\leqslant \int_0^1 | g(x) | \left(\int_0^1 | K(x, y) |^q\, dy \right)^{1/q} dx\, \| h \|_{q'} \qquad (4.2) \\ &\leqslant | K |_{p,q} \| g \|_{p'} \| h \|_{q'} . \end{aligned}$$

Thus since

$$\sup_{\substack{\|h\|_{q'} \leqslant 1 \\ \|g\|_{p'} \leqslant 1}} |(Kh, g)| = \| K \|_{p,q} ,$$

we have

$$\| K \|_{p,q} \leqslant | K |_{p,q} = E^0_{p,q}(K)$$

which proves the theorem for $n = 0$.

Now, for general n we prove the theorem by returning to (4.2) to see that for $u_1, \ldots, u_n \in L^p[0, 1]$, $v_1, \ldots, v_n \in L^q[0, 1]$

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

Thus we have

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

and

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

The first inequality implies that

$$d_n(\mathscr{K}_{q'}; L^p[0, 1]) \leqslant E^n_{p,q}(K),$$

while the second gives

$$d_n(\mathscr{K}^T_{p'}; L^q[0, 1]) \leqslant E^n_{p,q}(K).$$

Therefore Theorem 4.1 is proven for all n.

It is hardly surprising that this inequality is not always sharp. The basic comparison (4.2) between $\| K \|_{p,q}$ and $| K |_{p,q}$ relies on two applications of

Hölder's inequality which certainly eliminates, for all but special choices of p, q and kernels $K(x, y)$, equality from occurring. A particularly striking example of this occurrence is the case $p = q = 2$. We have already mentioned that E. Schmidt showed that

$$E_{2,2}^{n}(K) = \left(\sum_{n+1}^{\infty} \lambda_j\right)^{1/2}.$$

However, the lower bound from Theorem 4.1 is merely

$$\lambda_{n+1}^{1/2} = d_n(\mathscr{K}_2; L^2[0, 1]) = d_n(\mathscr{K}_2^T; L^2[0, 1]).$$

Neverheless, we have

THEOREM 4.2. *Let K be a nondegenerate totally positive kernel. Then for any $n \geqslant 0$, $1 \leqslant p \leqslant \infty$*

$$d_n(\mathscr{K}_\infty; L^p[0, 1]) = d_n(\mathscr{K}_{p'}^T; L^1[0, 1]) = E_{p,1}^n(K).$$

Moreover,

$$E_{p,1}^n(K) = | E |_{p,1},$$

where

$$E(x, y) = K\begin{pmatrix} x, & \tau_1, \ldots, \tau_n \\ y, & \xi_1, \ldots, \xi_n \end{pmatrix} \Big/ K\begin{pmatrix} \tau_1, \ldots, \tau_n \\ \xi_1, \ldots, \xi_n \end{pmatrix}$$

and $\xi_1, \ldots, \xi_n, \tau_1, \ldots, \tau_n$ are obtained from the function $g_{n,0,p}$ given in Theorem 3.2 where $r = 0$ and q is replaced by p. Furthermore, $\{u_i^0(x)\}_1^n$ and $\{v_i^0(y)\}_1^n$, as defined in Theorem 2.2 with respect to the above $\{\xi_i\}_1^n$ and $\{\tau_i\}_1^n$, are an optimal choice in the solution of (4.1).

Let us observe that for any kernel $\| K \|_{\infty,1} = | K |_{\infty,1}$. Thus when $p = \infty$, the above theorem is proved in [5]. Note however, that for $p < \infty$, $\| K \|_{p,1}$ is not always equal to $| K |_{p,1}$.

Proof. At this point, we have accumulated sufficient information on widths so as to facilitate the proof of this result. We observe that for $1 \leqslant p < \infty$

$$\begin{aligned} E_{p,1}^n(K) &\leqslant | E |_{p,1} \\ &= \left(\int_0^1 \left(\int_0^1 | E(x, y)| \, dy\right)^p dx\right)^{1/p} \\ &= \left(\int_0^1 \left(\left| \int_0^1 E(x, y)\, h_\xi(y)\, dy \right|\right)^p dx\right)^{1/p}. \end{aligned}$$

Furthermore, since $0 = g_{n,0,p}(\tau_i) = \int_0^1 K(\tau_i, y)\, h_\xi(y)\, dy$, $i = 1, 2,..., n$, we have

$$= \left(\int_0^1 \left(\left|\int_0^1 K(x, y)\, h_\xi(y)\, dy\right|\right)^p dx\right)^{1/p}$$

$$= \| g_{n,0,p} \|_p .$$

We now incoke Theorem 3.2 for $r = 0$, and q replaced by p to conclude that $d_n(\mathscr{K}_\infty; L^p[0, 1]) = \| g_{n,0,p} \|_p$. Hence equality is achieved in Theorem 4.1 and, in addition, $d_n(\mathscr{K}_{p'}^T; L^1[0, 1]) \leqslant \| g_{n,0,p} \|_p$.

However, from Theorem 3.5 (with I $= 0$, p replaced by p', and K by K^T), it follows that

$$d_n(\mathscr{K}_{p'}^T; L^1[0, 1]) = \| Kh_\xi \|_p = \| g_{n,0,p} \|_p .$$

This last equality follows from the definition of $\| g_{n,0,p} \|_p$.

References

1. R. Courant and D. Hilbert, "Methods of Mathematical Physics," Vol. 1, Interscience, New York, 1953.
2. C. R. Hobby and R. Rice, A moment problem in L_1 approximation, *Proc. Amer. Math. Soc.* **16** (1965), 665–670.
3. M. A. Krasnosel'ski, "Topological Methods in the Theory of Nonlinear Integral Equations," Pergamon, New York, 1964.
4. A. A. Melkman and C. A. Micchelli, "On Nonuniqueness of Optimal Subspaces for L^2 n-width," to appear *Illinois Journal of Mathematics.*
5. C. A. Micchelli and A. Pinkus, On n-widths in L^∞, *Trans. Amer. Math. Soc.* **234** (1977), 139–174.
6. C. A. Micchelli and A. Pinkus, Total positivity and the exact n-width of certain sets in L^1, *Pacific J. Math.* **71** (1977), 499–515.
7. C. A. Micchelli and A. Pinkus, Best mean approximation to a 2-dimensional kernel by tensor products, *Bull. Amer. Math. Soc.* **83** (1977), 400–402.
8. C. A. Micchelli, Best L^1 approximation by weak Chebyshev systems and the uniqueness of interpolating perfect splines, *J. Approximation Theory* **19** (1977), 1–14.
9. A. Pietsch, S-Numbers of operators in Banach spaces, *Studia Math.* **51** (1974), 201–223.
10. A. Pinkus, A simple proof of the Hobby–Rice theorem, *Proc. Amer. Math. Soc.* **60** (1976), 82–84.
11. E. Schmidt, Zur Theorie der linearen und nichtlinearen Integralgleichungen, I., *Math. Ann.* **63** (1907), 433–476.
12. I. J. Schoenberg and 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.