Optimal spline spaces of higher degree for L^2 n-widths

Michael S. Floater* Espen Sande†

May 11, 2017

Abstract

In this paper we derive optimal subspaces for Kolmogorov n-widths in the L^2 norm with respect to sets of functions defined by kernels. This enables us to prove the existence of optimal spline subspaces of arbitrarily high degree for certain classes of functions in Sobolev spaces of importance in finite element methods. We construct these spline spaces explicitly in special cases.

Math Subject Classification: Primary: 41A15, 47G10, Secondary: 41A44

Keywords: n-widths, splines, isogeometric analysis, total positivity, Green's functions

1 Introduction

Let $X = (X, \|\cdot\|)$ be a normed linear space, A a subset of X, and X_n an n-dimensional subspace of X. Let

$$E(A, X_n) = \sup_{u\in A} \inf_{v\in X_n} \|u - v\|$$

be the distance to A from X_n relative to the norm of X. Then the Kolmogorov n-width of A relative to X is defined by

$$d_n(A) = \inf_{X_n} E(A, X_n).$$

A subspace X_n is called an optimal subspace for A provided that

$$d_n(A) = E(A, X_n).$$

*Department of Mathematics, University of Oslo, Moltke Moes vei 35, 0851 Oslo, Norway, *email: michaelf@math.uio.no*

†Same address, *email: espsand@math.uio.no*. The second author was supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement 339643

Recently, Babuska et al. [1] and Evans et al. [2] studied n-widths and optimal subspaces in the space $X = L^2(0,1)$, with $\|\cdot\|$ as the L^2 norm, the goal of [2] being to assess the approximation properties of splines (piecewise polynomials) for use in finite element methods, in the context of isogeometric analysis. One reason for this is that it is already known that certain classes of functions of importance in finite elements have optimal subspaces that indeed consist of splines. In the L^2 case this was first shown by Melkman and Micchelli [7]. Among the function classes they studied, two were based on the Sobolev spaces

$$H^r(0,1) = \{u \in L^2(0,1) : u^{(\alpha)} \in L^2(0,1) \text{ for all } \alpha = 0,1,\dots,r\},$$

for $r \geq 1$. One is the class

$$A^r = \{u \in H^r(0,1) : \quad \|u^{(r)}\| \leq 1\}, \tag{1}$$

for any integer $r \geq 1$. The other is

$$A_0^r = \{u \in A^r : \quad u^{(k)}(0) = u^{(k)}(1) = 0, \quad k = 0,2,4,...,r-2\}, \tag{2}$$

for even $r \geq 2$. Kolmogorov [6] already determined the n-width of A^r. For $n \geq r$ it corresponds to the $(n+1-r)$-th eigenvalue of a boundary value problem and an optimal subspace is the span of the first $n-r$ eigenfunctions of the problem and the first r monomials. In [3] it was conjectured that A^r has an optimal spline subspace. It was Melkman and Micchelli [7] who showed that A^r actually admits two optimal subspaces consisting of splines, of degrees $r-1$ and $2r-1$. From their work it follows that A_0^r similarly admits optimal spline subspaces of degrees $r-1$ and $2r-1$. Further work on these and other n-widths problems can be found in [9].

Numerical tests in [2] were used to see whether A^r and similar function classes have optimal spline subspaces of degrees other than $r-1$ and $2r-1$. Their tests suggest, for example, that for A^1, there may exist optimal spline subspaces of degrees higher than 1. These numerical results motivated us to try to extend the results of [7]. The main purpose of this paper is to report that both A^r and A_0^r have, in fact, optimal spline subspaces of higher degrees, specifically degrees $lr-1$, for any $l = 1,2,3,\dots$. For example, A^1 admits optimal spline subspaces of all degrees $0,1,2,\dots$.

2 Main results

In order to describe the results, it helps to define some spline spaces. Suppose $\boldsymbol{\tau} = (\tau_1,\dots,\tau_m)$ is a knot vector such that

$$0 < \tau_1 < \cdots < \tau_m < 1,$$

and let $I_0 = [0,\tau_1)$, $I_j = [\tau_j,\tau_{j+1})$, $j = 1,\dots,m-1$, and $I_m = [\tau_m,1]$. For any $d \geq 0$, let Π_d be the linear space of polynomials of degree at most d. Then we define

$$S_{d,\boldsymbol{\tau}} = \{s \in C^{d-1}[0,1] : s|_{I_j} \in \Pi_d,\ j = 0,1,\dots,m\}.$$

Thus, $S_{d,\boldsymbol{\tau}}$ is the linear space of splines on $[0,1]$ of degree d with order of continuity $d-1$, knot vector $\boldsymbol{\tau}$ and dimension

$$\dim(S_{d,\boldsymbol{\tau}}) = m + d + 1.$$

We now describe our main results. Firstly, consider the class of functions A_0^r for even r. From the analysis of [7], one can show that the n-width of A_0^r is

$$d_n(A_0^r) = \frac{1}{(n+1)^r \pi^r}, \tag{3}$$

and that an optimal subspace is

$$X_n^0 = [\sin \pi x, \sin 2\pi x, \dots, \sin n\pi x], \tag{4}$$

the span of the functions $\{\sin \pi x, \sin 2\pi x, \dots, \sin n\pi x\}$. Further, let $\boldsymbol{\tau} = (\tau_1, \dots, \tau_n)$ be the uniform knot vector

$$\tau_j = \frac{j}{n+1}, \quad j = 1, \dots, n. \tag{5}$$

We will show

Theorem 1 *For even $r \geq 2$,*

$$X_n^l = \{s \in S_{lr-1,\boldsymbol{\tau}} : s^{(k)}(0) = s^{(k)}(1) = 0, \quad k = 0, 2, 4, \dots, lr-2\}$$

is an optimal subspace for A_0^r for any $l = 1, 2, 3, \dots$.

Secondly, consider the class of functions A^r for any $r \geq 1$. Let $\phi_1, \phi_2, \dots$ be the eigenfunctions of the eigenvalue problem

$$(-1)^r \phi_n^{(2r)} = \mu_n \phi_n, \qquad \phi_n^{(i)}(0) = \phi_n^{(i)}(1) = 0, \quad i = 0, 1, \dots, r-1, \tag{6}$$

with positive eigenvalues $0 < \mu_1 < \mu_2 < \cdots$, and let $\psi_n = \phi_n^{(r)}$, $n \geq 1$. The n-width of A^r, for $n \geq r$, is

$$d_n(A^r) = \mu_{n-r+1}^{-1/2}, \tag{7}$$

and an optimal subspace is

$$X_n^0 = \Pi_{r-1} + [\psi_1, \dots, \psi_{n-r}], \tag{8}$$

as shown by Kolmogorov [6]. Further, ϕ_{n-r+1} has $n-r$ simple zeros in $(0,1)$,

$$\phi_{n-r+1}(\xi_i) = 0, \quad 0 < \xi_1 < \cdots < \xi_{n-r} < 1,$$

and ψ_{n-r+1} has n simple zeros in $(0,1)$,

$$\psi_{n-r+1}(\eta_i) = 0, \quad 0 < \eta_1 < \cdots < \eta_n < 1.$$

Let $\boldsymbol{\xi} = (\xi_1, \dots, \xi_{n-r})$ and $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)$.

Theorem 2 *For any $r \geq 1$,*

$$X_n^l = \{s \in S_{lr-1,\boldsymbol{\tau}} : \ s^{(k)}(0) = s^{(k)}(1) = 0,$$
$$k = (2q+1)r, ..., (2q+2)r - 1, \ q = 0, ..., \lfloor l/2 \rfloor - 1\}$$

is an optimal subspace for A^r for any $l = 1, 2, 3, \dots$, where $\boldsymbol{\tau} = \boldsymbol{\xi}$ for l odd and $\boldsymbol{\tau} = \boldsymbol{\eta}$ for l even.

Here, $\lfloor \cdot \rfloor$ is the floor function. The spaces X_n^1 and X_n^2 were found in [7], but not X_n^l for $l \geq 3$.

3 Sets defined by kernels

We denote the norm and inner product on L^2 by

$$\|f\|^2 = (f, f), \qquad (f, g) = \int_0^1 f(t)g(t)\,dt,$$

for real-valued functions f and g. Let K be the integral operator,

$$Kf(x) = \int_0^1 K(x, y)f(y)\,dy,$$

and as in [7] we use the notation $K(x, y)$ for the kernel of K. We will only consider kernels $K(x, y)$ that are continuous or piecewise continuous for $x, y \in [0, 1]$. If we define the set

$$A = \{Kf : \|f\| \leq 1\}, \tag{9}$$

then, for any n-dimensional subspace X_n of L^2,

$$E(A, X_n) = \sup_{\|f\|\leq 1} \|(I - P_n)Kf\|, \tag{10}$$

where P_n is the orthogonal projection onto X_n.

Similar to matrix multiplication, the kernel of the composition of two integral operators K and L is

$$(KL)(x, y) = (K(x, \cdot), L(\cdot, y)).$$

We will denote by K^* the adjoint of the operator K, defined by

$$(f, K^*g) = (Kf, g).$$

The kernel of K^* is $K^*(x, y) = K(y, x)$, and the two compositions K^*K and KK^* have kernels

$$(K^*K)(x, y) = (K(\cdot, x), K(\cdot, y)), \qquad (KK^*)(x, y) = (K(x, \cdot), K(y, \cdot)).$$

The operator K^*K, being symmetric and positive semi-definite, has eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq \cdots \geq 0,$$

and corresponding orthogonal eigenfunctions

$$K^*K\phi_n = \lambda_n\phi_n, \qquad n = 1, 2, \ldots. \tag{11}$$

If we further define $\psi_n = K\phi_n$, then

$$KK^*\psi_n = \lambda_n\psi_n, \qquad n = 1, 2, \ldots, \tag{12}$$

and the ψ_n are also orthogonal. We note that the square roots of the λ_n are known as the s-numbers of K (or K^*). With these definitions we can state the following well-known theorem for Kolmogorov n-widths [9, p. 65].

Theorem 3 *$d_n(A) = \lambda_{n+1}^{1/2}$ and $X_n^0 = [\psi_1, \ldots, \psi_n]$ is an optimal subspace.*

4 Totally positive kernels

To discuss further optimal subspaces we need the concept of total positivity, and in this section we take some basic definitions from [7] or [9, Chap. IV, Sec. 5.1]. We say that the kernel $K(x,y)$ (or equivalently, the operator K) is totally positive provided that the determinant

$$K\begin{pmatrix} x_1,\ldots,x_n \\ y_1,\ldots,y_n \end{pmatrix} := \det(K(x_i,y_j))_{i,j=1}^n \geq 0, \tag{13}$$

for all $0 \leq x_1 < x_2 < \cdots < x_n \leq 1$, $0 \leq y_1 < y_2 < \cdots < y_n \leq 1$, $n = 1,2,\ldots$.

We will call K nondegenerate if

$$\dim[K(\cdot,y_1),\ldots,K(\cdot,y_n)] = \dim[K(x_1,\cdot),\ldots,K(x_n,\cdot)] = n$$

for all $0 < x_1 < x_2 < \cdots < x_n < 1$ and $0 < y_1 < y_2 < \cdots < y_n < 1$. One can show that K is nondegenerate if and only if for any $0 < y_1 < \cdots < y_n < 1$ the determinant in (13) is non-zero for some choice of the x_i, and for any $0 < x_1 < \cdots < x_n < 1$ it is non-zero for some choice of the y_i.

Suppose next that K and L are two integral operators. The basic composition formula, a generalization of the Cauchy-Binet formula [4, p. 17], applied to the composition of K and L gives

$$(KL)\begin{pmatrix} x_1,\ldots,x_n \\ y_1,\ldots,y_n \end{pmatrix} = \int\cdots\int_{0<s_1<\cdots<s_n<1} K\begin{pmatrix} x_1,\ldots,x_n \\ s_1,\ldots,s_n \end{pmatrix} L\begin{pmatrix} s_1,\ldots,s_n \\ y_1,\ldots,y_n \end{pmatrix} ds_1\cdots ds_n.$$

From this it follows that if K and L are totally positive then so is KL, and if K and L are also nondegenerate then so is KL. Thus if K is NTP (nondegenerate totally positive) then K^*K and KK^* are also NTP.

The basic composition formula shows, moreover, that if K is NTP,

$$(K^*K)\begin{pmatrix} x_1,\ldots,x_n \\ x_1,\ldots,x_n \end{pmatrix} > 0$$

for all $0 < x_1 < x_2 < \cdots < x_n < 1$, and KK^* has the same property. By a theorem of Kellogg [9, p. 109] it then follows that the eigenvalues of K^*K and KK^* in (11) and (12) are positive and simple, $\lambda_1 > \lambda_2 > \cdots > \lambda_n > \cdots > 0$, and the eigenfunctions ϕ_{n+1} and ψ_{n+1} have exactly n simple zeros in $(0,1)$,

$$\phi_{n+1}(\xi_j) = \psi_{n+1}(\eta_j) = 0, \quad j = 1,2,\ldots,n,$$

$$0 < \xi_1 < \xi_2 < \cdots < \xi_n < 1, \qquad 0 < \eta_1 < \eta_2 < \cdots < \eta_n < 1.$$

Using these definitions, Melkman and Micchelli [7] showed

Theorem 4 *If $K(x,y)$ is an NTP kernel,*

$$X_n^1 = [K(\cdot,\xi_1),\ldots,K(\cdot,\xi_n)]$$

and

$$X_n^2 = [(KK^*)(\cdot,\eta_1),\ldots,(KK^*)(\cdot,\eta_n)]$$

are optimal subspaces for A defined by (9).

The proof for X_n^1 is based on first observing that from (10), for any X_n,

$$E(A, X_n) = \sup_{\|f\|\le 1} ((I - P_n)Kf, (I - P_n)Kf)^{1/2}$$
$$= \sup_{\|f\|\le 1} ((I - P_n)Kf, Kf)^{1/2} = \sup_{\|f\|\le 1} (Tf, f)^{1/2},$$

where $T = K^*(I - P_n)K$, and thus $E(A, X_n)$ is the square root of the largest eigenvalue of T. For X_n^1, one can show that $(\lambda_{n+1}, \phi_{n+1})$ is an eigenpair of T, and, finally, using the total positivity of K^*K, that λ_{n+1} is the largest eigenvalue of T, and therefore, by Theorem 3, $E(A, X_n^1) = d_n(A)$. The optimality of X_n^2 follows from the optimality of X_n^1. We will discuss this in the next section.

5 Further optimal subspaces

Let us now consider a way of generating further optimal subspaces for sets A given by (9). Define the set A^* by

$$A^* = \{K^* f : \|f\| \le 1\}.$$

Lemma 1 *For any integral operator K,*

$$E(A, K(X_n)) \le E(A^*, X_n), \tag{14}$$

where

$$K(X_n) := \{Kf : f \in X_n\}.$$

Proof. To prove (14), let P_n be the orthogonal projection onto X_n. Then the image of KP_n is $K(X_n)$ and so

$$E(A, K(X_n)) \le \sup_{\|f\|\le 1} \|(K - KP_n)f\| = \sup_{\|f\|\le 1} \|(K^* - P_n K^*)f\| = E(A^*, X_n).$$

□

Since $d_n(A) = d_n(A^*)$, it follows that if X_n is optimal for the n-width of A^* then $K(X_n)$ is optimal for the n-width of A, provided $K(X_n)$ is n-dimensional. This was essentially the idea used in [7] to prove the optimality of X_n^2 from the optimality of X_n^1. As X_n^1 is optimal for the n-width of A, by reversing the roles of K and K^*,

$$(X_n^1)^* := [(K^*)(\cdot, \eta_1), \ldots, (K^*)(\cdot, \eta_n)]$$

is optimal for the n-width of A^*. Therefore, by Lemma 1, $K((X_n^1)^*) = X_n^2$ is optimal for the n-width of A. It is n-dimensional because KK^* is NTP.

Applying Lemma 1 twice we find that

$$E(A, KK^*(X_n)) \le E(A^*, K^*(X_n)) \le E(A, X_n). \tag{15}$$

Hence, if X_n is an optimal subspace for the n-width of A, then so is $KK^*(X_n)$, as long as it is n-dimensional, and the same is true of $(KK^*)^i(X_n)$ for any $i = 1, 2, \ldots$. For X_n^0 in Theorem 3 we find that $KK^*(X_n^0) = X_n^0$. However, applying KK^* to the two spaces X_n^1 and X_n^2 derived in Theorem 4, and since all compositions of K and K^* are NTP if K is NTP we obtain

Theorem 5 *If K is NTP,*

$$X_n^l = \begin{cases} (KK^*)^i(X_n^1) = [(KK^*)^i K(\cdot,\xi_1),\dots,(KK^*)^i K(\cdot,\xi_n)], & l = 2i+1, \\ (KK^*)^i(X_n^2) = [(KK^*)^{i+1}(\cdot,\eta_1),\dots,(KK^*)^{i+1}(\cdot,\eta_n)], & l = 2i+2, \end{cases}$$

is an optimal subspace for all $l = 1,2,3,\dots$.

We remark that Theorem 5 could also be proved directly from Lemma 1 by iteratively applying the same argument used to prove the optimality of X_n^2 from the optimality of X_n^1. The optimality of X_n^3 follows from the optimality of X_n^2, and so on.

We further note that if K is symmetric as well as NTP, then if X_n is an optimal subspace for A so is $K(X_n)$. In this case $KK^* = K^*K = K^2$, and thus $\xi_i = \eta_i$, $i = 1,\dots,n$.

Example 1. For any nonnegative numbers $t_1,\dots,t_m$ we define the polynomial of degree $r = 2m$ by $q_r(x) = \prod_{j=1}^m(-x^2+t_j^2)$. Let $D = d/dx$ and consider the set

$$A = \{u \in H^r(0,1): \quad \|q_r(D)u\| \le 1 \quad u^{(k)}(0) = u^{(k)}(1) = 0, \quad k = 0,2,...,r-2\}.$$

Then A is of the form (9) where $K(x,y)$ is the Green's function for the differential equation

$$\prod_{j=1}^m \left(-\frac{d^2}{dx^2} + t_j^2\right) u = f, \qquad u^{(k)}(0) = u^{(k)}(1) = 0, \quad k = 0,2,...,2m-2.$$

The kernel $K(x,y)$ is NTP [8]. It is also symmetric with eigenvalues $q_r(ik\pi)^{-1}$ and corresponding eigenfunctions $\sin k\pi x$ for $k = 1,2,\dots$, and so by Mercer's theorem we can express $K(x,y)$ as

$$K(x,y) = 2\sum_{k=1}^{\infty} \frac{\sin k\pi x \sin k\pi y}{q_r(ik\pi)}. \tag{16}$$

Hence, K^2 has the same eigenfunctions as K, but with eigenvalues $q_r(ik\pi)^{-2}$ for $k = 1,2,\dots$. By Theorem 3 the n-width of A is equal to

$$d_n(A) = q_r(i(n+1)\pi)^{-1},$$

and X_n^0 in (4) is an optimal subspace. Since the zeros of $\sin(n+1)\pi x$ are the τ_j of (5), Theorem 5 shows that

$$X_n^l = [(K^l)(\cdot,\tau_1),\dots,(K^l)(\cdot,\tau_n)] \tag{17}$$

is an optimal subspace for all $l = 1,2,3,\dots$, where, from (16),

$$(K^l)(x,y) = 2\sum_{k=1}^{\infty} \frac{\sin k\pi x \sin k\pi y}{q_r(ik\pi)^l}, \qquad l = 1,2,\dots.$$

Example 2. Again, let $r = 2m$, and consider the special case of Example 1 in which $t_1, \ldots, t_m = 0$. Then $A = A_0^r$, as defined by equation (2), and $K(x,y)$ is the Green's function for the differential equation

$$(-1)^m u^{(2m)} = f, \qquad u^{(k)}(0) = u^{(k)}(1) = 0, \quad k = 0, 2, ..., 2m-2.$$

This kernel is also NTP and symmetric, and its eigenvalues are $1/(k\pi)^{2m}$ with corresponding eigenfunctions $\sin k\pi x$ for $k = 1, 2, \ldots$. Hence, the eigenvalues of K^2 are $1/(k\pi)^{4m}$ for $k = 1, 2, \ldots$, and so by Theorem 3 the n-width of A_0^r is given by (3) and X_n^0 in (4) is again an optimal subspace. Moreover, in this special case, we have the explicit representation $K = J^m$, where

$$J(x,y) = \begin{cases} x(1-y) & x < y, \\ (1-x)y & x \geq y, \end{cases}$$

see, e.g. [5, p. 108]. Observe that $J(\cdot, y)$ is a linear spline with a breakpoint at y with the end conditions

$$J(0,y) = J(1,y) = 0.$$

For $m \geq 2$, the kernel $J^m(x,y)$ of J^m is the unique solution to the boundary value problem

$$-\frac{\partial^2}{\partial x^2}(J^m)(x,y) = (J^{m-1})(x,y), \qquad J^m(0,y) = J^m(1,y) = 0.$$

It follows by recursion on m that for any $m \geq 1$, $J^m(\cdot, y)$ is a C^{r-2} spline of degree $r-1$, with a breakpoint at y, with end conditions

$$\frac{\partial^k}{\partial x^k} J^m(0,y) = \frac{\partial^k}{\partial x^k} J^m(1,y) = 0, \qquad k = 0, 2, 4, \ldots, r-2.$$

Since $K^l = J^{lm}$, $K^l(\cdot, y)$ has the same properties as $J^m(\cdot, y)$ but with r replaced by lr. Applying this to (17) proves Theorem 1.

6 Adding additional functions

We now extend the above analysis to sets of the form

$$A = \left\{ \sum_{j=1}^{r} a_j k_j(\cdot) + Kf : \|f\| \leq 1, \, a_i \in \mathbb{R} \right\}, \tag{18}$$

for certain additional functions $k_1, ..., k_r$. As in [7] or [9, p. 118], we assume that $K(x,y)$ and the $\{k_i\}_{i=1}^r$ satisfy the following three properties:

1. $\{k_1, \ldots, k_r\}$ is a Chebyshev system on $(0,1)$, i.e., for all $0 < x_1 < \cdots < x_r < 1$,

$$k\begin{pmatrix} x_1, \ldots, x_r \\ 1, \ldots, r \end{pmatrix} = \det(k_j(x_i))_{i,j=1}^r > 0.$$

2. For every choice of points $0 \leq y_1 < \cdots < y_m \leq 1$ and $0 \leq x_1 < \cdots < x_{r+m} \leq 1$, $m \geq 0$, the determinant

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

is nonnegative.

3. Furthermore, for any $0 < y_1 < \cdots < y_m < 1$ the above determinant is non-zero for some choice of the x_i, and for any given $0 < x_1 < \cdots < x_{r+m} < 1$ it is non-zero for some choice of the y_i.

Next, let

$$K_r = (I - Q_r)K, \tag{19}$$

where Q_r is the orthogonal projection onto $Z_r = [k_1, \ldots, k_r]$. Then $A = Z_r \oplus \tilde{A}$ where $\tilde{A} = \{K_r f : \|f\| \leq 1\}$, and $\oplus$ is an orthogonal sum. Furthermore, for $n \geq r$,

$$d_n(A) = d_{n-r}(\tilde{A}),$$

and if $[u_1, \ldots, u_{n-r}]$ is optimal for the $(n-r)$-width of $\tilde{A}$ then $[k_1, \ldots, k_r, u_1, \ldots, u_{n-r}]$ is optimal for the n-width of A.

Further following [7] and [9, p. 118], we have from the three properties above that the kernel of $K_r^* K_r$ is totally positive and that

$$(K_r^* K_r)\begin{pmatrix} x_1, \ldots, x_n \\ x_1, \ldots, x_n \end{pmatrix} > 0$$

for all $0 < x_1 < x_2 < \cdots < x_n < 1$, even though K_r is not totally positive. Thus $K_r^* K_r$ has distinct, positive eigenvalues $\lambda_1 > \lambda_2 > \cdots > 0$ and corresponding orthogonal eigenfunctions

$$K_r^* K_r \phi_n = \lambda_n \phi_n, \quad n = 1, 2, \ldots. \tag{20}$$

Moreover, ϕ_{n-r+1} has exactly $n - r$ simple zeros in $(0, 1)$,

$$\phi_{n-r+1}(\xi_i) = 0, \quad 0 < \xi_1 < \cdots < \xi_{n-r} < 1.$$

Further, letting $\psi_n = K_r \phi_n$,

$$K_r K_r^* \psi_n = \lambda_n \psi_n, \quad n = 1, 2, \ldots, \tag{21}$$

and, although $K_r K_r^*$ is not totally positive, its eigenfunction ψ_{n-r+1} has exactly n simple zeros in $(0, 1)$,

$$\psi_{n-r+1}(\eta_i) = 0, \quad 0 < \eta_1 < \cdots < \eta_n < 1.$$

We now let J_r be the interpolation operator from $C[0,1]$ to Z_r determined by $J_r h(\eta_i) = h(\eta_i)$, $i = 1, \ldots, r$, and define $\overline{K}_r = (I - J_r)K$. Then for $n \geq r$ we have from [7] that

$$d_{n-r}(\tilde{A}) = \lambda_{n-r+1}^{1/2},$$

and

$$\begin{aligned}
\tilde{X}_{n-r}^0 &= [\psi_1, \ldots, \psi_{n-r}], \\
\tilde{X}_{n-r}^1 &= [K_r(\cdot, \xi_1), \ldots, K_r(\cdot, \xi_{n-r})], \\
\tilde{X}_{n-r}^2 &= [(\overline{K}_r \overline{K}_r^*)(\cdot, \eta_{r+1}), \ldots, (\overline{K}_r \overline{K}_r^*)(\cdot, \eta_n)]
\end{aligned}$$

are optimal subspaces for the $(n-r)$-width of $\tilde{A}$, and thus

$$X_n^k = Z_r + \tilde{X}_{n-r}^k, \qquad k = 0, 1, 2, \tag{22}$$

are optimal subspaces for the n-width of A.

7 Further optimal subspaces

Let us now express L^2 as the orthogonal sum $L^2 = Z_r \oplus \tilde{L}^2$, where

$$\tilde{L}^2 = \{u \in L^2 : u \perp Z_r\}.$$

Suppose that $\tilde{X}_{n-r}$ is some $(n-r)$-dimensional subspace of $\tilde{L}^2$, and let

$$X_n = Z_r \oplus \tilde{X}_{n-r}.$$

Then,

$$E(A, X_n) = E(\tilde{A}, \tilde{X}_{n-r}).$$

Now, similar to Section 5 we have

Lemma 2 *Suppose that $n \geq r$, and that $\tilde{X}_{n-r}$ is some $(n-r)$-dimensional subspace of $\tilde{L}^2$. Then*

$$E(A, Z_r \oplus K_r K_r^*(\tilde{X}_{n-r})) \leq E(A, Z_r \oplus \tilde{X}_{n-r}).$$

Proof. Applying (15) to $\tilde{A}$ gives $E(\tilde{A}, K_r K_r^*(\tilde{X}_{n-r})) \leq E(\tilde{A}, \tilde{X}_{n-r})$. □

Thus, if $\tilde{X}_{n-r}$ is an optimal subspace for the $(n-r)$-width of $\tilde{A}$, $Z_r \oplus K_r K_r^*(\tilde{X}_{n-r})$ is an optimal subspace for the n-width of A as long as it is n-dimensional.

8 Spline spaces

We now return to the problem of approximating A^r in (1). By considering the Taylor expansion

$$u(x) = \sum_{i=0}^{r-1} \frac{u^{(i)}(0)}{i!} x^i + \frac{1}{(r-1)!} \int_0^x (x-y)^{r-1} u^{(r)}(y)\, dy,$$

we can express A^r in the form of (18) where

$$k_i(x) = x^i, \quad i = 0, 1, \ldots, r-1, \qquad K(x,y) = \frac{1}{(r-1)!}(x-y)_+^{r-1}, \tag{23}$$

in which case $Z_r = \Pi_{r-1}$. In this case the eigenfunctions of $K_r^* K_r$ and $K_r K_r^*$, given by (20) and (21), are equally well determined by the equations (6), used by [6]. To see this, let us start by showing that $K_r^* K_r$, with K_r given by (19) and (23), solves a boundary value problem.

Lemma 3 *If $u(x) = K_r^* K_r f(x)$ then u is the unique solution to the boundary value problem*

$$\begin{aligned} (-1)^r u^{(2r)}(x) &= f(x), \qquad && x \in (0,1), \\ u^{(i)}(0) = u^{(i)}(1) &= 0, && i = 0, 1, \ldots, r-1. \end{aligned} \tag{24}$$

Proof. We have

$$u = K^*(I - Q_r)^2 K f = K^*(I - Q_r) K f,$$

and

$$Kh(x) = \frac{1}{(r-1)!} \int_0^x (x-y)^{r-1} h(y)\, dy, \qquad K^* h(x) = \frac{1}{(r-1)!} \int_x^1 (y-x)^{r-1} h(y)\, dy.$$

Differentiating these gives

$$\begin{aligned} (Kh)^{(i)}(x) &= \frac{1}{(r-1-i)!} \int_0^x (x-y)^{r-1-i} h(y)\, dy, \qquad && i = 0, 1, \ldots, r-1, \\ (K^* h)^{(i)}(x) &= \frac{(-1)^r}{(r-1-i)!} \int_x^1 (y-x)^{r-1-i} h(y)\, dy, && i = 0, 1, \ldots, r-1, \end{aligned}$$

and

$$(Kh)^{(r)}(x) = h(x), \qquad (K^* h)^{(r)}(x) = (-1)^r h(x).$$

From the r-th derivatives we have

$$u^{(2r)}(x) = (-1)^r ((I - Q_r) K f)^{(r)}(x) = (-1)^r (Kf)^{(r)}(x) = (-1)^r f(x),$$

which is the differential equation in (24). Regarding the boundary conditions, for any h,

$$(K^* h)^{(i)}(1) = 0, \qquad i = 0, 1, \ldots, r-1.$$

On the other hand,

$$(K^* h)^{(i)}(0) = \frac{1}{(r-1-i)!} \int_0^1 y^{r-1-i} h(y)\, dy, \qquad i = 0, 1, \ldots, r-1,$$

and so, if h is orthogonal to Π_{r-1}, we also have

$$(K^* h)^{(i)}(0) = 0, \qquad i = 0, 1, \ldots, r-1.$$

Since $h = (I - Q_r) K f$ is indeed orthogonal to Π_{r-1}, we thus obtain the boundary conditions.

To see that u is unique, suppose $f = 0$ in (24). Then u must be a polynomial of degree at most $2r - 1$. But then, to satisfy the (Hermite) boundary conditions, we must have $u = 0$. □

Next, we show that $K_r K_r^*$ also solves a boundary value problem.

Lemma 4 *If f is orthogonal to Π_{r-1} and $u(x) = K_r K_r^* f(x)$, then u is the unique solution, orthogonal to Π_{r-1}, of the boundary value problem*

$$\begin{aligned} (-1)^r u^{(2r)}(x) &= f(x), && x \in (0,1), \\ u^{(i)}(0) = u^{(i)}(1) &= 0, && i = r, r+1, \ldots, 2r-1. \end{aligned} \tag{25}$$

Proof. We have

$$u = (I - Q_r) K K^* (I - Q_r) f = (I - Q_r) K K^* f.$$

Therefore,

$$u^{(2r)}(x) = (K^* f)^{(r)}(x) = (-1)^r f(x),$$

which is the differential equation in (25). Regarding the boundary conditions, we find that for $i = r, r+1, \ldots, 2r-1$,

$$u^{(i)}(x) = (K^* f)^{(i-r)}(x),$$

which is zero when $x = 1$ and since f is orthogonal to Π_{r-1}, it is also zero when $x = 0$.

Regarding uniqueness, suppose $f = 0$ in (25). Then u must be a polynomial of degree at most $2r - 1$,

$$u(x) = \sum_{i=0}^{2r-1} c_i x^i.$$

Then, by either derivative boundary condition of highest order, we see that $c_{2r-1} = 0$. Then u has degree at most $2r - 2$. Then by the boundary conditions of next highest order, $c_{2r-2} = 0$. We continue in this way to deduce that $u \in \Pi_{r-1}$. But in that case, assuming that u is orthogonal to Π_{r-1}, we must have $u = 0$. □

From these two lemmas, the eigenfunctions of $K_r^* K_r$ and $K_r K_r^*$ are indeed also determined by equation (6). Therefore the n-width of A^r in (1) is given by (7) and an optimal subspace is X_n^0 of (8).

Next, consider the optimal subspace X_n^1 of (22). This was identified as a spline space in [7] as follows. First observe that

$$X_n^1 = [k_1, \ldots, k_r, K(\cdot, \xi_1), \ldots, K(\cdot, \xi_{n-r})],$$

since the function $K_r(\cdot, \xi) - K(\cdot, \xi)$ belongs to $[k_1, \ldots, k_r]$ for any ξ. Therefore, from the truncated power form of K in (23), we see that

$$X_n^1 = S_{r-1, \boldsymbol{\xi}},$$

as in Theorem 2.

Finally, we turn to X_n^2 in (22). This was also identified as a spline space in [7], but some of the details of the derivation were omitted and we include them here. We start by computing the r-th derivative of $(\overline{K}_r \overline{K}_r^*)(\cdot, y)$ appearing in X_n^2.

Lemma 5 *Let $B_{r-1}(x;\eta_1,\ldots,\eta_r,y)$ be the B-spline in x of degree $r-1$ with respect to the knots $\eta_1,\ldots,\eta_r,y$. Then*

$$\frac{\partial^r}{\partial x^r}(\overline{K}_r\overline{K}_r^*)(x,y) = c_y B_{r-1}(x;\eta_1,\ldots,\eta_r,y),$$

where c_y is a constant independent of x.

Proof. By definition,

$$\overline{K}_r(x,y) = K(x,y) - p(x,y),$$

where, for each y, $p(\cdot,y)$ is the interpolating polynomial of degree $\le r-1$ such that $p(\eta_i,y) = K(\eta_i,y)$, $i=1,\ldots,r$. The kernel of $\overline{K}_r\overline{K}_r^*$ is therefore

$$(\overline{K}_r\overline{K}_r^*)(x,y) = (\overline{K}_r(x,\cdot),\overline{K}_r(y,\cdot)) = (K(x,\cdot),\overline{K}_r(y,\cdot)) - (p(x,\cdot),\overline{K}_r(y,\cdot)).$$

Consider now the r-th derivative of $(\overline{K}_r\overline{K}_r^*)(x,y)$ with respect to x. Since $p(x,y)$ is a polynomial in x of degree $\le r-1$,

$$\frac{\partial^r}{\partial x^r}(\overline{K}_r\overline{K}_r^*)(x,y) = \frac{\partial^r}{\partial x^r}(K(x,\cdot),\overline{K}_r(y,\cdot)).$$

Since

$$(K(x,\cdot),\overline{K}_r(y,\cdot)) = \int_0^x \frac{(x-z)^{r-1}}{(r-1)!}\overline{K}_r(y,z)\,dz,$$

repeated differentiation with respect to x gives

$$\frac{\partial^{r-1}}{\partial x^{r-1}}(K(x,\cdot),\overline{K}_r(y,\cdot)) = \int_0^x \overline{K}_r(y,z)\,dz,$$

and therefore,

$$\frac{\partial^r}{\partial x^r}(\overline{K}_r\overline{K}_r^*)(x,y) = \overline{K}_r(y,x).$$

Now, since $K(\cdot,y)-p(\cdot,y)$ is the error of the polynomial interpolant $p(\cdot,y)$ to $K(\cdot,y)$ at the points $\eta_1,\ldots,\eta_r$, Newton's error formula implies

$$\overline{K}_r(x,y) = (x-\eta_1)\cdots(x-\eta_r)[\eta_1,\ldots,\eta_r,x]K(\cdot,y).$$

Therefore, swapping the variables x and y, we have

$$\frac{\partial^r}{\partial x^r}(\overline{K}_r\overline{K}_r^*)(x,y) = (y-\eta_1)\cdots(y-\eta_r)[\eta_1,\ldots,\eta_r,y]\frac{(\cdot-x)_+^{r-1}}{(r-1)!},$$

and so the result follows from the divided difference definition of a B-spline. □

By Lemma 5, the $(n-r)$-dimensional space

$$\left[\frac{\partial^r}{\partial x^r}(\overline{K}_r\overline{K}_r^*)(x,\eta_{r+1}),\ldots,\frac{\partial^r}{\partial x^r}(\overline{K}_r\overline{K}_r^*)(x,\eta_n)\right]$$

is, with the knot vector $\boldsymbol{\eta}$ of Theorem 2, the spline space

$$\{s \in S_{r-1,\boldsymbol{\eta}} : s|_{[0,\eta_1]} = s|_{[\eta_n,1]} = 0\}.$$

By integrating the splines in this space r times with respect to x, and since $X_n^2 = \Pi_{r-1} + \tilde{X}_n^2$, we find that

$$\begin{aligned} X_n^2 &= \{s \in S_{2r-1,\boldsymbol{\eta}} : s|_{[0,\eta_1]}, s|_{[\eta_n,1]} \in \Pi_{r-1}\} \\ &= \{s \in S_{2r-1,\boldsymbol{\eta}} :\ s^{(k)}(0) = s^{(k)}(1) = 0,\ k = r, ..., 2r-1\}, \end{aligned}$$

which agrees with Theorem 2.

9 Higher degree spline spaces

Using Lemma 2, it follows that the subspaces

$$X_n^l = \begin{cases} \Pi_{r-1} + (K_r K_r^*)^i \tilde{X}_n^1, & l = 2i+1, \\ \Pi_{r-1} + (K_r K_r^*)^i \tilde{X}_n^2, & l = 2i+2, \end{cases}$$

are optimal for A^r for all $l = 1, 2, 3, \ldots$. They are n-dimensional because, by the Green's function property of $K_r K_r^*$ of Lemma 4, if $K_r K_r^* f = 0$ for any f orthogonal to Π_{r-1}, then $f = 0$. Since $K_r^* = K^*(I - Q_r)$, we can express these spaces more simply as

$$X_n^l = \begin{cases} \Pi_{r-1} + (K_r K_r^*)^i X_n^1, & l = 2i+1, \\ \Pi_{r-1} + (K_r K_r^*)^i X_n^2, & l = 2i+2. \end{cases}$$

Now applying Lemma 4 recursively to X_n^1 and X_n^2 shows that the X_n^l are the spline spaces of Theorem 2 as claimed.

10 Approximating functions in H^1

Let us consider the special case of the theory for A^r in (1) when $r = 1$. Then the eigenvalues and eigenfunctions in (6) are

$$\mu_n = n^2\pi^2, \qquad \phi_n(x) = \sin n\pi x, \qquad \psi_n(x) = \cos n\pi x, \qquad n = 1, 2, \ldots,$$

and therefore the n-width of A^1 is

$$d_n(A^1) = \frac{1}{n\pi}, \tag{26}$$

and an optimal subspace is

$$X_n^0 = [1, \cos \pi x, \cos 2\pi x, \ldots, \cos(n-1)\pi x],$$

as identified in [6]. As regards the spline subspaces, observe that the zeros of $\sin n\pi x$ and $\cos n\pi x$ are

$$\begin{aligned} \xi_j &= j/n, & j &= 1, \ldots, n-1, \\ \eta_j &= (j - 1/2)/n, & j &= 1, \ldots, n, \end{aligned}$$

and so the knot vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ of Theorem 2 are uniform, and

$$X_n^l = \{s \in S_{l-1,\boldsymbol{\tau}} :\ s^{(k)}(0) = s^{(k)}(1) = 0,\ k = 1, 3, 5, \ldots, 2\lfloor l/2 \rfloor - 1\} \tag{27}$$

with $\boldsymbol{\tau} = \boldsymbol{\xi}$ for l odd and $\boldsymbol{\tau} = \boldsymbol{\eta}$ for l even. So for A^1 there is an optimal spline space of every degree.

From (26) we obtain the optimal error estimate

$$\|u - P_{n,l}u\| \leq \frac{1}{n\pi}\|u'\|, \tag{28}$$

for $u \in H^1$, where $P_{n,l}$ is the orthogonal projection onto the spline space X_n^l in (27). We remark that very similar spline spaces to the X_n^l in (27) are defined in [10]. In fact, when l is odd the X_n^l coincide with the spline spaces of even degree in their paper and in this case inequality (28) improves on their Theorem 1.

References

[1] I. Babuska, U. Banerjee, and J. E. Osborne, *On principles for the selection of shape functions for the generalized finite element method*, Comput. Methods Appl. Mech. Engrg. **191** (2002), 5595–5629.

[2] J. A. Evans, Y. Bazilevs, I. Babuska, and T. J. R. Hughes, *n-Widths, sup-infs, and optimality ratios for the k-version of the isogeometric finite element method*, Comput. Methods Appl. Mech. Engrg. **198** (2009), 1726–1741.

[3] A. Ioffe and V. M. Tihomirov, *Duality of convex functions and extremum problems*, Russian Math. Surveys **23** (1968), 53–124.

[4] S. Karlin, *Total positivity, Vol. I*, Stanford University Press, Stanford, California, 1968.

[5] H. B. Keller, *Numerical methods for two-point boundary-value problems*, Dover Publications, New York, 1992.

[6] A. Kolmogorov, *Über die beste Annäherubg von Funktionen einer gegebenen Funktionenklass*, Ann. of Math. **37** (1936), 107–110.

[7] A. A. Melkman and C. A. Micchelli, *Spline spaces are optimal for L2 n-widths*, Illinois J. Math. **22** (1978), 541–564.

[8] C. A. Micchelli and A. Pinkus, *On n-widths in L-infinity*, Trans. Amer. Math. Soc. **234** (1977), 139–174.

[9] A. Pinkus, *n-Widths in approximation theory*, Springer-Verlag, Berlin, 1985.

[10] S. Takacs and T. Takacs, *Approximation error estimates and inverse inequalities for B-splines of maximum smoothness*, Math. Models Methods Appl. Sci. **26** (2016), no. 7, 1411–1445.