Approximation on Simplices with respect to Weighted Sobolev Norms

Dietrich Braess and Christoph Schwab

Abstract

Inequalities of Jackson and Bernstein type are derived for polynomial approximation on simplices with respect to Sobolev norms. Although we cannot use orthogonal polynomials, sharp estimates are obtained from a decomposition into orthogonal subspaces. The formulas reflect the symmetries of simplices, but analogous estimates on rectangles show that we cannot expect rotational invariance of the terms with derivatives.

1 Introduction

The approximation of functions by polynomials with respect to a weighted L_2-norm is strongly related to orthogonal polynomials. This is well known for functions on the real interval $[-1, +1]$. The orthogonal polynomials for constant weights are the Legendre polynomials P_n which satisfy

$$\int_{-1}^{+1} P_n P_m dx = \frac{2}{2n+1}\delta_{nm}.$$

The Legendre polynomials are eigenfunctions of the singular Legendre differential operator,

$$\mathcal{L}P_n = \mu_n P_n, \quad \mu_n = n(n+1)$$

where $\mathcal{L}$ is given by $(\mathcal{L}v)(x) := -((1-x^2)v')'$. We therefore have also orthogonality of the derivatives with respect to a weight function which vanishes at the boundaries

$$\int_{-1}^{+1} (1-x^2) P_n' P_m' dx = \mu_n \frac{2}{2n+1}\delta_{nm}.$$

If we expand an L_2-function with respect to the Legendre polynomials for the natural normalization $v = \sum_{k=0}^{\infty} b_k \left(k+\frac{1}{2}\right)^{1/2} P_k$, then we have obviously,

$$\begin{aligned}
\|v\|_0^2 &:= \int_{-1}^{+1} v^2 dx &&= \sum_{k=0}^{\infty} |b_k|^2, \\
|v|_{1,w}^2 &:= \int_{-1}^{+1} (1-x^2)(v')^2 dx &&= \sum_{k=1}^{\infty} \mu_k |b_k|^2
\end{aligned}$$

and more generally, for any $\ell \in \mathbb{N}_0$,

$$|v|_{\ell,w}^2 := (-1)^\ell \int_{-1}^{+1} v\mathcal{L}^\ell v dx = \sum_{k=1}^{\infty} (\mu_k)^\ell |b_k|^2$$

which is to be understood in the sense that the series converge if and only if $|v|_{\ell,w}$ is finite. Correspondingly, we introduce for $m \in \mathbb{N}_0$ the sets

$$V^m := \left\{ v \in L^2(-1,1);\; |v|_{\ell,w} < \infty \text{ for } \ell = 0, \ldots, \ell. \right\}$$

From the definitions we obtain for $v \in V^m$, $\ell, m \in \mathbb{N}_0$, $m \geq \ell$, the *approximation property* (direct estimate)

$$\inf_{p \in \mathcal{P}_n} |v - p|_{\ell,w} \leq (\mu_{n+1})^{-(m-\ell)/2} |v|_{m,w} \tag{1.1}$$

and the *inverse estimate*

$$|p|_{m,w} \leq (\mu_n)^{(m-\ell)/2} |p|_{\ell,w} \quad \text{for } p \in \mathcal{P}_n. \tag{1.2}$$

Direct and inverse estimates estimates for the rectangle are easily obtained from these results by tensor product arguments. To establish analogous results for triangles and more generally for simplices in $\mathbb{R}^d$ is the purpose of the present paper.

There are two approaches in the literature for obtaining explicit representations of orthogonal polynomials. The first method is based on Appell's polynomials from 1881, see [1, 2]. They provide only a decomposition into finite dimensional subspaces and biorthogonal polynomials, see also [7, 8]. Orthogonal polynomials have been derived from Appell's polynomials in [9, 4], but the expressions are so involved that it seems to be hard to establish approximation properties from those results. Another approach is obtained from a transformation of the triangle to the square [12, 10, 11, 6, 5]. Orthogonal polynomials are expressed in terms of Jacobi polynomials. Unfortunately these polynomials are less suited for our intention since the transformation makes that the derivatives are not directly achieved.

We will choose a different approach and consider subspaces of polynomials as invariant subspaces of suitable differential operators. In contrast to Derriennic

[4] symmetric versions are chosen. After establishing the approximation properties the authors learnt that these operators have already served for the treatment of Bernstein–Durrmeyer operators [3, 13]. Since we can do without the representation of Appell's polynomials, we obtain shorter proofs for the eigenvalue problem. The consequences that we draw are also new.

2 Preliminaries. An Inverse Estimate

We will discuss some facts which provide a motivation for the main results or for the technique of the proofs. A reader who is already familiar with approximation on spheres or simplices may directly pass to the next section.

First we conclude from L_2 approximation on the rectangle $[-1,+1]^2$ that derivatives enter in an anisotropic way. The products of Legendre polynomials $P_k(x)P_\ell(y)$, $k,\ell = 0,1,2,\dots$ are orthogonal polynomials on the square. The same holds for derivatives in the direction of the edges:

$$\begin{aligned} &P_k'(x)P_l(y) &&\text{for the weight function } (1-x^2),\\ &P_k(x)P_l'(y) &&\text{for the weight function } (1-y^2). \end{aligned}$$

The natural norms are now

$$\begin{aligned} \|v\|_0 &:= \int v^2 dxdy,\\ \|v\|_{1,w} &:= \int [v_x^2(1-x^2)+v_y^2(1-y^2)]\,dxdy, \end{aligned}$$

and the corresponding polynomial spaces $\mathcal{Q}_{n,n} := \operatorname{span}\{x^k y^\ell;\ 0\le k,\ell\le n\}$. By elementary calculations we obtain

$$\begin{aligned} \inf_{p\in\mathcal{Q}_{n,n}} \|v-p\|_0 &\le \frac{1}{\sqrt{(n+1)(n+2)}}\,\|v\|_{1,w}\\ \|p\|_{1,w} &\le \sqrt{2n(n+1)}\,\|p\|_0 \quad\text{for } p\in\mathcal{Q}_{n,n}. \end{aligned}$$

Next we derive an inverse estimate for a weighted H^1-norm on triangles by using results from univariate functions. Only a factor smaller than 3 is lost in this way. We refer to the usual reference triangle

$$T := \{(x,y)\in\mathbb{R}^2;\ x\ge 0, y\ge 0, 1-x-y\ge 0\},$$

and the polynomials with fixed total degree

$$\mathcal{P}_n := \operatorname{span}\{x^k y^\ell;\ k+\ell\le n\}.$$

If $p \in \mathcal{P}_n$, then the restriction of p to constant y is a polynomial of degree $\leq n$ in the x-variable, and we conclude from the univariate case (1.2) that

$$\int_0^{1-y} x(1-x-y)p_x^2 dx \leq n(n+1) \int_0^{1-y} p^2 dx \quad \text{for } 0 \leq y < 1.$$

Integration over y yields

$$\int_T x(1-x-y)p_x^2 dxdy \leq n(n+1) \int_T p^2 dxdy, \quad p \in \mathcal{P}_n.$$

We may repeat the process for the directions given by the other two edges of the triangle and obtain the inverse estimate

$$\begin{aligned}\int_T [x(1-x-y)\left(\frac{\partial}{\partial x}p\right)^2 + y(1-x-y)\left(\frac{\partial}{\partial y}p\right)^2 \\ +xy\left(\frac{\partial}{\partial x}p - \frac{\partial}{\partial y}p\right)^2 dxdy \quad &\leq \quad 3n(n+1)\int_T p^2 dxdy.\end{aligned}$$

We will repeatedly meet expressions of this form. In order to present the estimate in a more symmetric form, we recall that x, y, and $1-x-y$ are the barycentric coordinates $\lambda_1, \lambda_2, \lambda_3$. Let $\partial_{k\to j}$ be the derivative in the direction showing from the vertex k to the vertex j. With this we have an estimate for a weighted H^1-norm

$$\sum_{k<j} \int_T \lambda_k \lambda_j (\partial_{k\to j} p)^2 dxdy \leq 3n(n+1) \int_T p^2 dxdy \quad \text{for } p \in \mathcal{P}_n. \tag{2.1}$$

3 Estimates on the Simplex in $\mathbb{R}^d$

Now we are prepared to consider the original approximation problem on a d-simplex S^d. It will be considered as the convex hull of $d+1$ points $a_0, a_1, \ldots, a_d \in \mathbb{R}^d$ which do not lie on a $(d-1)$-dimensional hyperplane. In order to keep the symmetry we refer to the barycentric coordinates $\lambda_0, \lambda_1, \ldots, \lambda_d$ of the points $x = \sum_j \lambda_j a_j \in S^d$. Specifically we have

$$\lambda_j \geq 0, \;\; j = 0, 1, \ldots, d, \quad \sum_j \lambda_j = 1,$$

We will make use of multiindex notation, in particular

$$\lambda^m := \lambda_0^{m_0}\lambda_1^{m_1}\ldots\lambda_d^{m_d}, \qquad \lambda^\alpha = \lambda_0^{\alpha_0}\lambda_1^{\alpha_1}\ldots\lambda_d^{\alpha_d},$$

and $|m| = \sum_j m_j$, $|\alpha| = \sum_j \alpha_j$. We assume that $\alpha_j > -1$ for all j. Hence, $w_\alpha := \lambda^\alpha$ is a weight function for which the inner product

$$(f, g) = \int_{S^d} f g w_\alpha \tag{3.1}$$

and the weighted L_2-norm $\|f\|_{0,w}^2 := (f, f)$ is well defined. As before, we set

$$\mathcal{P}_n := \mathrm{span}\{\lambda^m;\ |m| \le n\} \quad \text{and} \quad \mathcal{R}_n := \mathcal{P}_n \cap \mathcal{P}_{n-1}^\perp.$$

Due to the condition $\sum \lambda_j = 1$, the representation of a function given in terms of barycentric coordinates is not unique. Nevertheless we can write the directional derivative for the direction from a_k to a_j in the form

$$\frac{\partial}{\partial \lambda_j} - \frac{\partial}{\partial \lambda_k} \quad \text{or for short} \ \ \partial_j - \partial_k.$$

Lemma 3.1 *Let $j \ne k$. Then the differential operator of second order*

$$\mathcal{L}_0 := -\lambda^{-\alpha}(\partial_j - \partial_k)\,\lambda_j\lambda_k\lambda^\alpha\,(\partial_j - \partial_k) \tag{3.2}$$

is selfadjoint with respect to the inner product $(\cdot,\cdot)$. It maps $\mathcal{P}_n$ into $\mathcal{P}_n$ and $\mathcal{R}_n$ into $\mathcal{R}_n$.

Proof. Consider a segment on a line parallel to the direction from a_k to a_j. The product $\lambda_j\lambda_k$ vanishes at the two points at which the line intersects the boundary of S^d. Therefore no boundary terms occur when performing partial integration, and we have

$$\begin{aligned}\int_{S^d} f(\mathcal{L}_0 g) w_\alpha &= -\int_{S^d} f(\partial_j - \partial_k)\lambda_j\lambda_k\lambda^\alpha(\partial_j - \partial_k) g \\ &= \int_{S^d} [(\partial_j - \partial_k) f]\lambda_j\lambda_k\lambda^\alpha(\partial_j - \partial_k) g. \end{aligned}\tag{3.3}$$

From the symmetry of the last expression we obtain

$$\int_{S^d} f(\mathcal{L}_0 g) w_\alpha = \int_{S^d} (\mathcal{L}_0 f) g w_\alpha. \tag{3.4}$$

The operator $\mathcal{L}_0$ maps $\mathcal{P}_n$ into $\mathcal{P}_n$ since the differential operators cause a reduction of the degree of the polynomials which compensates the increase of the degree by the multiplication with the factor $\lambda_j\lambda_k$.

Finally let $p \in \mathcal{R}_n$ and $q \in \mathcal{P}_{n-1}$. Since $\mathcal{L}_0 q \in \mathcal{P}_{n-1}$ and $p \in \mathcal{P}_{n-1}^{\perp}$, we conclude that

$$\int\limits_{S^d} q(\mathcal{L}_0 p) w_\alpha = \int\limits_{S^d} (\mathcal{L}_0 q) p w_\alpha = 0, \tag{3.5}$$

and $\mathcal{L}_0 p$ is orthogonal to $\mathcal{P}_{n-1}$, i.e. $\mathcal{L}_0 p \in \mathcal{R}_n$. □

We are now prepared to introduce a differential operator which due to its symmetry can be regarded as a Laplacian for the simplex

$$\mathcal{L}_w := -\lambda^{-\alpha} \sum_{j<k} (\partial_j - \partial_k) \lambda_j \lambda_k \lambda^\alpha (\partial_j - \partial_k). \tag{3.6}$$

Similar mappings with less symmetries have been already considered by Derriennic [4] for the construction of orthogonal polynomials. After completing the text the authors learnt that the operator was used in [3, 13] for the study of Bernstein–Durrmeyer polynomials and that special cases of the eigenvalue problem (3.7) have already been stated by Appell and Kampé de Fériet in terms of Appell's polynomials. We prefer a direct proof.

Theorem 3.2 *The operator $\mathcal{L}_w$ is selfadjoint and*

$$\mathcal{L}_w p = \mu_n p \quad \text{for all } p \in \mathcal{R}_n. \tag{3.7}$$

with the eigenvalues μ_n explicitly given by

$$\mu_n = \mu_n(d, \alpha) := n(n + d + |\alpha|), \quad n = 1, 2, \dots \tag{3.8}$$

Proof. Let $|m| = n$. First we apply $\mathcal{L}_w$ to the monomials λ^m and use the abbreviation $H_j(\lambda) := \lambda_j^{m_j - 1} \prod_{k \neq j} \lambda_k^{m_k}$. By straight forward calculations mod $\mathcal{P}_{n-1}$ we obtain

$$\begin{aligned}
-\lambda^\alpha \mathcal{L}_w \lambda^m &= \sum_{j<k} (\partial_j - \partial_k) \lambda_j \lambda_k \lambda^\alpha (\partial_j - \partial_k) \lambda^m \\
&= \sum_{j<k} [m_j(m_j + \alpha_j) \lambda_k H_j \lambda^\alpha - m_k(m_j + \alpha_j + 1) \lambda^{m+\alpha} \\
&\qquad - m_j(m_k + \alpha_k + 1) \lambda^{m+\alpha} + m_k(m_k + \alpha_k) \lambda_j H_k \lambda^\alpha] \\
&= \sum_{j \neq k} [m_j(m_j + \alpha_j) \lambda_k H_j \lambda^\alpha - m_k(m_j + \alpha_j + 1) \lambda^{m+\alpha}] \\
&= \sum_j m_j(m_j + \alpha_j)(1 - \lambda_j) H_j \lambda^\alpha - \sum_{j \neq k} m_k(m_j + \alpha_j + 1) \lambda^{m+\alpha} \\
&\equiv -\sum_j m_j(m_j + \alpha_j) \lambda_j H_j \lambda^\alpha - \sum_{j \neq k} m_k(m_j + \alpha_j + 1) \lambda^{m+\alpha} \\
&= -n(n + |\alpha| + d) \lambda^{m+\alpha}.
\end{aligned}$$

Since $\mathcal{L}_w$ is linear, we have

$$\mathcal{L}_w p \equiv \mu_n\, p \pmod{\mathcal{P}_{n-1}}$$

for each $p \in \mathcal{P}_n$. From the preceding lemma we know that we have even equality if $p \in \mathcal{R}_n$. □

In accordance with (2.1) we now define a weighted H^1-seminorm which will form an appropriate pair together with $\|\cdot\|_{0,w}$

$$|f|_{1,w}^2 := \sum_{j<k} \int_{S^d} |(\partial_j - \partial_k) f|^2 \lambda_j \lambda_k w_\alpha.$$

From (3.4) we obtain our essential tool

$$|f|_{1,w}^2 = \int_{S^d} f(\mathcal{L}_w f) w_\alpha. \tag{3.9}$$

In particular assume that f is expanded into polynomials from the orthogonal subspaces

$$f = \sum_{k=0}^{\infty} p_k \quad \text{with } p_k \in \mathcal{R}_k.$$

From the orthogonality relations (3.5) and Theorem 3.2 we conclude that

$$\begin{aligned} \|f\|_{0,w}^2 &= \sum_{k=0}^{\infty} \|p_k\|_{0,w}^2\,, \\ |f|_{1,w}^2 &= \sum_{k=0}^{\infty} \int_{S^d} p_k(\mathcal{L}_w p_k) w_\alpha = \sum_{k=0}^{\infty} \mu_k \|p_k\|_{0,w}^2\,, \end{aligned}$$

and, more generally, for any $\ell \in \mathbb{N}_0$,

$$|f|_{\ell,w}^2 := \sum_{k=0}^{\infty} \int_{S^d} p_k(\mathcal{L}_w^\ell p_k) w_\alpha = \sum_{k=0}^{\infty} (\mu_n)^\ell \|p_k\|_{0,w}^2.$$

The last equality is understood in the sense that the infinite series converges if and only if $|f|_{\ell,w}$ is finite. Similar to $|f|_{1,w}$, the seminorm $|f|_{\ell,w}$ admits the following representation in terms of f and its derivatives:

$$|f|_{\ell,w}^2 = \begin{cases} \displaystyle\int_{S^d} (\mathcal{L}_w^m f)^2 w_\alpha & \text{if } \ell = 2m, \\ \displaystyle\int_{S^d} (\mathcal{L}_w^m f)\mathcal{L}_w(\mathcal{L}_w^m f) w_\alpha & \text{if } \ell = 2m+1. \end{cases} \tag{3.10}$$

Accordingly, for $m \in \mathbb{N}_0$, we define the weighted spaces

$$V_w^m(S^d) := \left\{ v \in L^2(S^d);\ |f|_{\ell,w} < \infty \text{ for } \ell = 0, 1, \ldots, m \right\}.$$

Then the main result is immediate and there is no gap between the direct and the inverse estimate.

Theorem 3.3 *Let ℓ, m be nonnegative integers and $m \geq \ell$ and denote by $\mu_n = n(n+d+|\alpha|)$ the eigenvalues of $\mathcal{L}_w$. Then, for any $v \in V_w^m(S^d)$, the approximation property*

$$\inf_{p \in \mathcal{P}_n} |v - p|_{\ell,w} \leq (\mu_{n+1})^{-(m-\ell)/2}\, |v|_{m,w} \quad n = 0, 1, 2, \ldots$$

holds, and for any $p \in \mathcal{P}_n$ we have the inverse estimate

$$|p|_{m,w} \leq (\mu_n)^{(m-\ell)/2}\, |p|_{\ell,w}.$$

Both inequalities are sharp.

From Theorem 3.3 we conclude that the factor $3n(n+1)$ in the estimate (2.1) may be replaced by $n(n+2)$. To this end we set $d = 2, \alpha = 0, m = 1$ and $\ell = 0$. The norm $|\cdot|_{1,w}$ refers to a weighted integral of first order derivatives.

The situation is more involved for $|\cdot|_{m,w}$ if $m > 1$ although the case $m = 2$ is still transparent in view of (3.10). Obviously the commutation rule $\partial_j \lambda_j = \lambda_j \partial_j + 1$ implies

$$(\partial_j - \partial_k)\lambda_j \lambda_k \lambda^\alpha (\partial_j - \partial_k) = \lambda_j \lambda_k (\partial_j - \partial_k)^2 + (\lambda_k - \lambda_j)(\partial_j - \partial_k)$$

From (3.10) and Young's inequality we obtain

$$\inf_{p \in \mathcal{P}_n} \|v - p\|_0^2 \leq 2\mu_{n+1}^{-2} \sum_{j<k} \Big\{ \int_{S^d} \lambda_j^2 \lambda_k^2 [(\partial_j - \partial_k)^2 v]^2 + \int_{S^d} (\lambda_k - \lambda_j)^2 [(\partial_j - \partial_k) v]^2 \Big\}.$$

References

[1] P. Appell (1881), Sur des polynômes de deux variables analogues aux polynômes de Jacobi. Arch. Math. Phys. 66, 238–245

[2] P. Appell and J. Kampé de Fériet (1926), Fonctions hypergéométriques et hypersphériques. Polynômes d'Hermite, Gauthier-Villars, Paris

[3] H. Berens, H.J. Schmid, and Yuan Xu (1992), Bernstein–Durrmeyer polynomials on a simplex. J. Approximation Theory 68, 247–261

[4] M.-M. Derriennic (1985), Polynômes orthogonaux de type Jacobi sur un triangle. C. R. Acad. Sci., Paris, Ser. I 300, 471–474

[5] M. Dubiner (1991), Spectral methods on triangles and other domains. J. Sci. Comput. 6, No.4, 345–390

[6] C.F. Dunkl (1984), Orthogonal polynomials with symmetry of order three. Can. J. Math. 36, 685–717

[7] A. Erdélyi, W. Magnus, F. Oberhettinger, and F.G. Tricomi (eds.) (1953) Higher transcendental functions. Vol. II. pp 269–273 (Bateman Manuscript Project.) New York–Toronto–London: McGraw-Hill

[8] E.D. Fackerell and R.A. Littler (1974), Polynomials biorthogonal to Appell's polynomials. Bull. Austral. Math. Soc. 11, 181–195

[9] W. Gröbner (1948), Über die Konstruktion von Systemen orthogonaler Polynome in ein- und zweidimensionalen Bereichen. (German). Monatsh. Math. 52, 38–54

[10] S. Karlin and J. McGregor (1964), On some stochastic models in genetics. Stochastic models in medicine and biology, 245–279 (Proc. Sympos. University of Wisconsin, Madison, June 1963. University of Wisconsin Press, Madison)

[11] T. Koornwinder (1975), Two-variable analogues of the classical orthogonal polynomials. Theory Appl. spec. Funct., Proc. adv. Semin., (D. Askey, ed.) Madison 1975, pp 435–495

[12] J. Proriol (1957), Sur une famille de polynômes a deux variables orthogonaux dans un triangle. C. R. Acad. Sci., Paris 245, 2459–2461

[13] Th. Sauer (1994), The genuine Bernstein–Durrmeyer operator on a simplex. Results in Math. 26, 99–130

[14] G. Szegö (1974), Orthogonal Polynomials. (Third edition). Amer. Math. Soc., Providence, RI

Dietrich Braess,
Faculty of Mathematics, Ruhr-University
D-44780 Bochum, Germany
e–mail: `braess@num.ruhr-uni-bochum.de`

Christoph Schwab,
Seminar für Angewandte Mathematik, ETH Zürich
CH-8092 Zürich, Switzerland
e–mail: `schwab@sam.math.ethz.ch`