# Best Approximation of Monomials in Several Variables

Yuan Xu

**Abstract**

For a monomial $x^\alpha$ in $d$ variables, the problem of best approximation to $x^\alpha$ by polynomials of lower degrees is studied on the unit sphere, the unit ball and the standard simplex. For the uniform norm we discuss what is known in $d = 2$, for which complete solutions are known, and in $d \ge 3$ for which only a few cases have been successfully solved. For the $L^2$ norm, we present the complete solution, including explicit formulas for the error of best approximation.

**Mathematical Subject Classification (2000).** 41A10, 41A50, 41A63

**Keywords and phrases.** Chebyshev polynomial, best approximation, several variables

## 1 Introduction

For $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$, we use the standard multi-index notation for the monomial $x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$. The degree of $x^\alpha$ is $|\alpha| = \alpha_1 + \ldots + \alpha_d$. The base $\{x^\alpha : |\alpha| \le n\}$ spans the space $\Pi_n^d$ of polynomials of degree at most $n$ in $d$ variables.

---

This paper is in final form and no version of it will be submitted for publication elsewhere.

Let $\Omega$ be a region in $\mathbb{R}^d$. For $f \in C(\Omega)$, the best approximation of $f$ from $\Pi_n^d$ in the uniform norm is the quantity

$$E_n(f;\Omega) = \inf_{p\in\Pi_{n-1}^d} \|f-p\|_{C(\Omega)}, \tag{1.1}$$

where $\|f\|_{C(\Omega)} = \max_{x\in\Omega} |f(x)|$. If $E_n(f;\Omega) = \|f-p^*\|_{C(\Omega)}$ then we call $p^*$ an extremal polynomial for $f$. We are interested in the case that $f$ is a monomial or some other fixed polynomials. In this case $f-p^*$ is often called a polynomial of least deviation from zero, which we shall call a *least polynomial* for $f$.

For $d=1$ and $\Omega=[-1,1]$ it is well-known that the best approximation to $x^n$ is given in terms of the Chebyshev polynomial of the first kind

$$T_n(x) = \cos n(\arccos x) = 2^{n-1}x^n + q(x), \qquad q \in \Pi_{n-1};$$

in fact, $p^*(x) = x^n - 2^{1-n}T_n(x)$ and $E_n(x^n;[-1,1]) = 2^{1-n}$. In the case of $\Omega$ being a cube in $\mathbb{R}^d$, the best approximation to $x^\alpha$ is given by the product of Chebyshev polynomials. We are interested in the cases that $\Omega$ is either the unit sphere $S^{d-1} = \{x : \|x\| = 1\}$ of $\mathbb{R}^d$, where $\|x\|$ denotes the usual Euclidean norm of $x \in \mathbb{R}^d$, the unit ball $B^d = \{x : \|x\| \le 1\}$ or the standard simplex $T^d = \{x : x_1 \ge 0, \ldots, x_d \ge 0, 1 - x_1 - \ldots - x_d \ge 0\}$ of $\mathbb{R}^d$.

To solve the problem of best approximation to the monomials, we need to complete two steps. The first one is to find the value of $E_n(f;\Omega)$ and the second is to find an explicit formula for the extremal polynomials. For $d=2$ the best approximation to monomials are known explicitly for these regions (see [16, 8, 17]). We explain the result in Section 2 and indicate how the first step can follow from an elementary simpleminded way, which essentially reduce the problem to that of one variable. In the background of the elementary approach, however, is the theory of extremal signature of Rivlin and Shapiro [18], which will play an essential role in the case of $d \ge 3$. The elementary approach is no longer enough for $d \ge 3$ and this is due to an essential difficulty. In some simple cases we will use symmetry of the regions and of the polynomials to simplify the task of finding the extremal polynomials. This and the the extremal signature will be explained in Section 3. For $d > 3$ the extremal polynomials are known only in a few cases (see [1, 2, 19, 25]), these will be given in Section 4 and the method used to derive them will be discussed. Finally, in Section 5, we state the extremal polynomials and the error of the best approximation in $L^2$ norm, which are known explicitly in all three regions.

# 2 Best approximation in two variables

For $d = 2$ the regions are the unit circle $S^1$, the unit disk $B^2$ and the triangle $T^2 = \{(x,y) : x, y \geq 0, x + y \leq 1\}$. We will denote the monomials by $x^k y^{n-k}$ for $0 \leq k \leq n$.

## 2.1 The unit circle $S^1$

The best approximation on the circle can be deduced from that of trigonometric polynomials on $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. Denote by $\mathcal{T}_n$ the space of trigonometric polynomials of degree at most $n$. Using $(x,y) = (\cos\theta, \sin\theta)$ for $\theta \in [0, 2\pi]$, every polynomial $\Pi_n^2(S^1)$ becomes a trigonometric polynomial in $\mathcal{T}_n$. In particular, using $e^{i\theta} = \cos\theta + i\sin\theta$, it is easy to see that

$$x^k y^{n-k} = t_{k,n}^*(\theta) + s_{k,n}(\theta),$$

where $s_{k,n} \in \mathcal{T}_{n-1}$ and

$$t_{k,n}^*(\theta) = \frac{1}{2^{n-1}} \begin{cases} \cos n\theta, & \text{if } n-k \text{ is even}, \\ \sin n\theta, & \text{if } n-k \text{ is odd}. \end{cases} \tag{2.1}$$

Since it is well known that the best approximation to $\cos n\theta$ from $\mathcal{T}_{n-1}$ is zero and $\sin\theta = \cos(\pi - \theta)$, it follows that

$$\inf_{p \in \Pi_{n-1}^2} \|x^k y^{n-k} - p\|_{C(S^1)} = \inf_{t \in \mathcal{T}_{n-1}} \frac{1}{2^{n-1}} \|\cos n\theta - t(\theta)\|_{C(\mathbb{T})} = \frac{1}{2^{n-1}} \tag{2.2}$$

and the infimum is attained by $t_{k,n}^*(\theta)$ in the trigonometric case. The extremal polynomial for $x^k y^{n-k}$ is obtained from rewriting $t_{k,n}(\theta)$ in (2.1) as a polynomial in $x, y$ using $x = \cos\theta$ and $y = \sin\theta$.

## 2.2 The unit disk $B^2$

For the monomial $x^k y^{n-k}$ it is easy to get a lower bound for the value of $E_n(x^k y^{n-k}; B^2)$. In fact, since $S^1$ is a subset of $B^2$, we have

$$\begin{aligned} E_n(x^k y^{n-k}; B^2) &= \inf_{p \in \Pi_{n-1}^2} \|x^k y^{n-k} - p\|_{C(B^2)} \\ &\geq \inf_{p \in \Pi_{n-1}^2} \|x^k y^{n-k} - p\|_{C(S^1)} = \frac{1}{2^{n-1}} \end{aligned} \tag{2.3}$$

Using (2.2). In order to show that the equality holds, we need to find one polynomial $G_{k,n}(x,y) = x^k y^{n-k} - p(x,y)$ with $p \in \Pi_{n-1}^2$ such that

$\|G_{k,n}\|_{C(B^2)} = 2^{1-n}$. The formula in (2.3) also shows that such a polynomial should satisfy $\|G_{k,n}\|_{C(B^2)} = \|G_n\|_{C(S^1)}$, which means that $G_{k,n}$ should satisfy

$$G_{k,n}(\cos\theta, \sin\theta) = 2^{1-n}\varepsilon_n t^*_{k,n}(\theta), \qquad 0 \le \theta \le 2\pi, \tag{2.4}$$

where $\varepsilon = \pm 1$ and $t^*_{k,n}$ is defined in (2.1). Of course the restriction of $G$ on $S^1$ will not determine $G$ completely. Such a polynomial was found in [8] in terms of $U_n$, the Chebyshev polynomial of the second kind,

$$U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}, \qquad x = \cos\theta.$$

We summarize the result in the following:

**Theorem 2.1** *For $n \ge k \ge 0$, $E_n(x^k y^{n-k}; B^2) = 2^{1-n}$ and*

$$G_{k,n}(x,y) = \frac{1}{2^n}\left(U_n(x)U_{n-k}(y) + U_{n-2}(x)U_{n-k-2}(x)\right).$$

*is a least polynomial for $x^k y^{n-k}$ on $B^2$.*

The elementary trigonometric identity shows that $G_{k,n}$ given in the theorem satisfies the equation (2.4). The most difficult part of the proof comes down to show that

$$|G_{k,n}(x,y)| \le 2^{1-n}, \qquad (x,y) \in B^2.$$

In contrast to the case of one variable, polynomials of best approximation in several variables are not unique in general. For the monomial $x^k y^{n-k}$ the uniqueness holds only if $k = 0$ or $k = n$, or $n = 1$ ([8, Theorem 2.2]). Furthermore, it is known that the difference between two least polynomials are necessarily of the form $(1 - x^2 - y^2)q(x,y)$ with $q \in \Pi^2_{n-3}$.

Using the generating functions, another construction of the least polynomial for $x^k y^{n-k}$ was given in [17]. Define a family of polynomials $R_\alpha \in \Pi^d_n$, $\alpha \in \mathbb{N}^d_0$, by

$$\|t\|^n T_n\left(\frac{\langle t, x\rangle}{\|t\|}\right) = \sum_{|\alpha|=n} R_\alpha(x) t^\alpha, \qquad t, x \in \mathbb{R}^d,$$

where $\langle t, x\rangle$ is the usual inner product in $\mathbb{R}^d$. Then $R_\alpha(x) = 2^{|\alpha|-1}\binom{|\alpha|}{\alpha}x^\alpha + p_\alpha$, $p_\alpha \in \Pi^d_{n-1}$. For $d = 2$ we write $\alpha = (k, n-k)$. It turns out that

$$\frac{1}{2^{n-1}}\binom{n}{k} R_{k,n-k}(x,y) = x^k y^{n-k} + p_k(x,y), \qquad p_k \in \Pi^2_{n-1},$$

is a least polynomial on $B^2$. This family is essentially different from that of $G_{k,n}$ in Theorem 2.1.

One can also generate further family of least polynomials from the existing ones. It is observed in [4] that if $p_{k,n}$ is a least polynomial for $x^k y^{n-k}$ then the polynomial

$$\frac{1}{2^{(m-1)n}} p_{k,n}\left(T_m(x), yU_{m-1}(x)\right) \in \Pi^2_{mn}$$

is a least polynomial for $x^{mn-n+k}y^{n-k}$, which can be different from either $G_{k,n}$ or the multiple of $R_{k,n-k}$ given above.

Another direction is to generate least polynomials for other homogeneous polynomials from the ones for monomials. As an example, one result [8, Theorem 3.2] states that if $p_{k,n-k}$ is the least polynomial for $x^k y^{n-k}$ on $B^2$ and $k$ is even, then

$$q_{k,n}(x,y) = (-1)^{k/2} p_{k,n}\left((1-x^2-y^2)^{1/2}, y\right)$$

is a least polynomial for $(x^2+y^2)^k y^{n-k}$ on $B^2$. Finally let us mention that the extremal polynomial for $x^k y^{n-k}$ on $B^2$ is also the Kergin interpolant of $x^k y^{n-k}$ ([5, 10, 13]).

**The standard triangle** $T^2$. Just as in the case of $B^2$, it is easy to get a lower bound for the value of $E_n(x^k y^{n-k}; T^2)$. We use the fact that

$$\inf_{p\in\Pi_{n-1}} \|x^n - p\|_{C([0,1])} = \frac{1}{2^{2n-1}} \|T_n(2x-1)\|_{C([0,1])} = \frac{1}{2^{2n-1}}$$

instead of best approximation by trigonometric polynomials. We have

$$\begin{aligned} E_n(x^k y^{n-k}; T^2) &= \inf_{p\in\Pi^2_{n-1}} \|x^k y^{n-k} - p\|_{C(T^2)} \\ &\geq \inf_{p\in\Pi_{n-1}} \|x^k (1-x)^{n-k} - p(x, 1-x)\|_{C([0,1])} \\ &= \inf_{p\in\Pi_{n-1}} \|x^n - p\|_{C([0,1])} = \frac{1}{2^{2n-1}}. \end{aligned} \tag{2.5}$$

In order to show that the equality holds, we need to find one least polynomial $T_{k,n}$ for $x^k y^{n-k}$. Since the least polynomial for $x^n$ on $[0,1]$ is given by $2^{1-2n}T_n(2x-1)$, the polynomial $T_{k,n}$ should agree with the Chebyshev polynomial $T_n(2x-1)$ on $\{(x, 1-x) : 0 \leq x \leq 1\}$ which is part of the boundary of $T^2$. Such a polynomial was found in [16]. We summarize the result as the following:

**Theorem 2.2** *For $m \geq k \geq 0$, define*

$$T_{k,m}(x,y) = \begin{cases} T_{k-m}(2y-1)T_m(8xy-1) & \\ \quad +8xy(2y-1)U_{k-m-1}(2x-1)U_{m-1}(8xy-1), & \text{if } k > m \\ T_k(8xy-1), & \text{if } k = m \\ T_{m-k}(2x-1)T_k(8xy-1) & \\ \quad +8xy(2x-1)U_{m-k-1}(2y-1)U_{k-1}(8xy-1), & \text{if } k < m. \end{cases}$$

*Then $2^{1-2n}T_{k,n-k}$ is a polynomial of least deviation from zero for $x^k y^{n-k}$ on $T^2$ and $E_n(x^k y^{n-k}; T^2) = 2^{1-2n}$.*

The difficult part is to recognize that $T_{k,n-k}$ as given above is the correct function. Once the formula is identified, the only difficult part of the proof is to show that $|T_{k,m}(x,y)| \leq 1$ on $T^2$.

It was pointed out in [4] that there is a close relation between best approximation on $T^2$ and that on $B^2$. In fact, if $p(x,y)$ is a least polynomial for $x^k y^{n-k}$ on the triangle $T^2$ then $p(x^2, y^2)$ is a least polynomial for $x^{2k} y^{2n-2k}$ on $B^2$ and the relation holds conversely. A more precise statement will be given in the following section. This implies, in particular, that we can deduce the least polynomials for $x^k y^{n-k}$ on $T^2$ from those for $x^{2k} y^{2n-2k}$ on $B^2$. The exact relation between the polynomials $T_{k,n-k}$ and the polynomials $G_{k,n}$ is established in [4].

# 3 Extremal signature and invariant polynomials

In this section we discuss some of the tools that can be used to study the best approximation by polynomials.

## 3.1 Extremal signature

What is hidden in the elementary way of getting to $E_n(f; B^2)$ is the characterization of the extremal polynomials in terms of the extremal signature, which is the natural extension of the Chebyshev alternating theorem in one variable. In this section we review the extremal signatures.

The original study in [18] is in the general setting of approximation from a finite dimensional subspace of $C(\Omega)$ on a compact Hausdorff space $\Omega$. We will state the version only in the case that we are interested in. Let $\Omega$ be an infinite compact set in $\mathbb{R}^d$. For our purpose it is either $S^{d-1}$, $B^d$ or $T^d$. A *signature* $\sigma$ on the set $\Omega$ is a function with finite support

$\{x_1, \ldots, x_N\}$ where $x_i \in \Omega$, such that $\sigma(x_i) = \pm 1$ for $1 \le i \le N$ and $\sigma(x) = 0$ if $x \in \Omega \setminus \{x_1, \ldots, x_N\}$. A signature $\sigma$ is *extremal* with respect to $\Pi_n^d$ if there exists a subset $\mathcal{S}$ in the support of $\sigma$ and positive numbers $\lambda_v$, $v \in \mathcal{S}$, such that

$$\sum_{v \in \mathcal{S}} \lambda_v \sigma(v) p(v) = 0, \qquad \text{for all } p \text{ in } \Pi_n^d.$$

Let $r > 0$ be a fixed number. For each $p \in \Pi_{n-1}^d$, we denote by $\mathcal{S}_r(p; f)$ the set

$$\mathcal{S}_r(p; f) = \{x \in \Omega : |f(x) - p(x)| = r\}.$$

In the case of $r = \|f - p\|_{C(\Omega)}$, $\mathcal{S}_r(p; f)$ is the set of extremal points of $f - p$ and we denote it by $\mathcal{S}(p; f)$. The characterization of the best approximation of $f$ from $\Pi_n^d$ is given by the following theorem.

**Theorem 3.1** *1. A polynomial $p^*$ in $\Pi_n^d$ satisfies $\|f - p^*\|_{C(\Omega)} = E_n(f; \Omega)$ if and only if there exists an extremal signature $\sigma$ with support in $\mathcal{S}(p^*; f)$ such that $\sigma(v) = \mathrm{sign}(f - p^*)(v)$ for all $v \in \mathcal{S}(p^*; f)$.*

*2. Suppose there exist a polynomial $p^* \in \Pi_n^d$ and an extremal signature $\sigma$ supported on $\mathcal{S}_r(p^*; f)$. Then $E_n(f; \Omega) \ge r$.*

The theorem states that if $p$ is a polynomial of best approximation on $\Omega$ then it is so on a finite subset of $\Omega$. For the proof of this theorem, see [18]. The part two of the statement is useful for computing a lower bound for $E_n(f; \Omega)$, especially since we often do not know the polynomial of best approximation. For various examples of extremal signatures, see [8, 15, 18, 21, 22].

In the case of best approximation to $x^k y^{n-k}$ on the disk $B^2$, one extremal signature agrees with that of extremal signature on the circle $S^1$. This signature is described as follows:

Let $\theta_1 < \theta_2 < \ldots < \theta_{2n}$ be the $2n$ extremal points of the trigonometric polynomial $t_{k,n}^*$ defined in (2.1). Let $\mathcal{S} = \{(\cos\theta_j, \sin\theta_j) : 1 \le j \le 2n\}$ and define the signature $\sigma$ by $\sigma(\cos\theta_j, \sin\theta_j) = (-1)^j$; then $\sigma$ is an extremal signature for $x^k y^{n-k}$ on $B^2$.

The proof follows from simple trigonometric identities, such as

$$\frac{1}{n} \sum_{k=0}^{n-1} e^{2ijk\pi/n} = \delta_{0,j} \quad \text{and} \quad \frac{1}{n} \sum_{k=0}^{n-1} e^{ij(2k+1)\pi/n} = \delta_{0,j}, \quad 0 \le j \le n-1,$$

where $\delta_{0,j} = 1$ if $j = 0$ and $\delta_{0,j} = 0$ otherwise. Furthermore we notice that the extremal signature is connected to the quadrature formulas. In fact, in the case that $\theta_1, \ldots, \theta_{2n}$ are zeros of $\sin n\theta$, $\theta_k = 2k\pi/n$, it follows from the above two sums that the following quadratures hold,

$$\frac{1}{2\pi}\int_0^{2\pi} f(x)dx = \frac{1}{n}\sum_{k=0}^{n-1} f\left(\frac{2k\pi}{n}\right) := \mathcal{L}_1 f, \qquad f \in \mathcal{T}_{n-1},$$
$$\frac{1}{2\pi}\int_0^{2\pi} f(x)dx = \frac{1}{n}\sum_{k=0}^{n-1} f\left(\frac{(2k+1)\pi}{n}\right) := \mathcal{L}_2 f, \qquad f \in \mathcal{T}_{n-1}.$$

Thus, $\sigma$ is an extremal signature since the sum $\mathcal{L}f := \mathcal{L}_1 f - \mathcal{L}_2 f$ satisfies $\mathcal{L}f = 0$ for $f \in \Pi^2_{n-1}$. The case that $\theta_k$ are zeros of $\cos n\theta$ works similarly.

## 3.2 Invariant polynomials

For $d \geq 3$ one can often use the symmetry of the functions and the set $\Omega$ to gain some knowledge of the least polynomials. In essence, one can state that if $\Omega$ is invariant under a finite group $G$ and $f$ is also invariant under the same group, then the polynomial of best approximation has to be a polynomial invariant under $G$. This ideal is not new (see, for example, [7]) and the idea of making using of symmetry has been widely used in dealing with the cubature formulas (see for example, [11, 14, 20]). In [1] it is used to obtain least polynomials on the unit sphere and on the unit ball. We shall state the result for the general domain.

Let $\Omega$ be a region in $\mathbb{R}^d$ that is invariant under a finite group $G$; that is, $\Omega g := \{xg : g \in G, x \in \Omega\}$ coincides with $\Omega$. For a function $f$ defined on $\Omega$ we defined the action of $g \in G$ on $f$ by $R(g)f(x) = f(xg)$. A function $f$ on $\Omega$ is said to be invariant under $G$ if $R(g)f = f$ for all $g \in G$. Let $G\Pi^d_n$ denote the space of polynomials in $\Pi^d_n$ that are invariant under $G$.

The domains $S^{d-1}$ and $B^d$ are rotation invariant, we can take $G$ as any finite subgroup of the rotation group $O(d)$. For the simplex the largest group under which $T^d$ is invariant is the symmetric group $\mathcal{S}(T^d)$ that contains all permutations of the vertices of the simplex $T^d$. If $f$ is invariant under $\mathcal{S}(T^d)$ then it is a symmetric function of $x_1, \ldots, x_d, 1 - x_1 - \ldots - x_d$. The symmetric group $\mathcal{S}_d$ of $d$ coordinates is a subgroup of $\mathcal{S}(T^d)$.

Since the proof is essentially the same we shall state the result for the $L^p$ norm for $1 \leq p \leq \infty$. For $1 \leq p < \infty$ let $\|f\|_{L^p(\Omega)}$ denote the $L^p$ norm of $f$ on $\Omega$ with respect to an absolutely continuous finite Borel measure $d\mu = \mu' dx$

$$\|f\|_{L^p(\Omega)} = \left(\int_\Omega |f(x)|^p d\mu(x)\right)^{1/p}$$

and we assume that the measure $\mu$ is invariant under $G$. We take the convention that the norm for $p = \infty$ is the $C(\Omega)$ norm.

**Theorem 3.2** *Let $\Omega$ be a region in $\mathbb{R}^d$ invariant under a finite group $G$. If $f$ is invariant under $G$, then*

$$\inf_{p \in \Pi_{n-1}^d} \|f(x) - p(x)\|_{L^p(\Omega)} = \inf_{p \in G\Pi_{n-1}^d} \|f(x) - p(x)\|_{L^p(\Omega)}.$$

*Proof.* Since $G\Pi_{n-1}^d \subset \Pi_{n-1}^d$, we evidently have

$$\inf_{p \in \Pi_{n-1}^d} \|f(x) - p(x)\|_{L^p(\Omega)} \le \inf_{p \in G\Pi_{n-1}^d} \|f(x) - p(x)\|_{L^p(\Omega)}.$$

For a given function $F$ we denote by $F_G$ the symmetric function

$$F_G(x) = \frac{1}{|G|} \sum_{g \in G} R(g) F(x)$$

where $|G|$ denote the number of elements in $G$. Clearly if $F$ is invariant under $G$, then $F_G = F$. Using the triangle inequality and the fact that $f$ is invariant under $G$ we have

$$\begin{aligned} \|f - p\|_{L^p(\Omega)} &= \frac{1}{|G|} \sum_{g \in G} \|R(g)f - R(g)p\|_{L^p(\Omega)} \\ &\ge \|f_G - p_G\|_{L^p(\Omega)} = \|f - p_G\|_{L^p(\Omega)}. \end{aligned}$$

Taking infimum over $p$ completes the proof.

In [1] this theorem is used for the sphere $S^{d-1}$ and the ball $B^d$ with $G$ being a finite subgroup of the rotation group $O(d)$ to obtain the polynomial of best approximation for a number of special polynomials. For example, it is proved there that

$$\inf_{p \in \Pi_{d-1}^d} \|x_1 \cdots x_d - p\|_{L^p(S^{d-1})} = \|x_1 \cdots x_d\|_{L^p(S^{d-1})}, \tag{3.1}$$

where $G$ is the symmetric group $\mathcal{S}_d$, and

$$\inf_{p \in \Pi_3^d} \|x_1^4 + \cdots + x_d^4 - p\|_{L^p(S^{d-1})} = \inf_{c \in \mathbb{R}} \|x_1^4 + \cdots + x_d^4 - c\|_{L^p(S^{d-1})}, \tag{3.2}$$

where $G$ is the hyperoctahedral group $\mathcal{B}_d$, which is the group of symmetry of the unit cube $\{\pm 1, \ldots, \pm 1\}$ in $\mathbb{R}^d$ and it contains permutations of coordinates and the sign changes (the semi-product of the symmetric group $\mathcal{S}_d$

and the abelian group $\mathbb{Z}_2^d$). There are also several other least polynomials in [1], obtained by taking the above or the other finite subgroups of $O(d)$.

Since $\mathbb{Z}_2^d$ is a subgroup of $O(d)$ which contains all possible sign changes of the coordinates, one consequence of the theorem states that if $f$ is even in each of its variables, so is a polynomial $p$ of best approximation to $f$. Let $\mathbb{R}_+ := \{x \in \mathbb{R}^d : x_1 \geq 0, \ldots, x_d \geq 0\}$ and let $\phi$ denote the mapping

$$\phi : \mathbb{R}^d \mapsto \mathbb{R}_+^d : x \mapsto \phi(x) = (x_1^2, \ldots, x_d^2).$$

We note that $\phi$ is one-to-one from $T^d$ to $B_+^d = B^d \cap \mathbb{R}_+^d$. It also induces a one-to-one mapping from $\Pi_n^d$ to $G\Pi_{2n}^d$ with $G = \mathbb{Z}_2^d$, since every polynomial $p \in G\Pi_{2n}^d$ can be written uniquely as $p_e \circ \phi$ for some $p_e \in \Pi_n^d$ and, conversely, $q \circ \phi \in G\Pi_{2n}^d$ for every $q \in \Pi_n^d$. For each $f$ that is even in each of its variables, we also write $f = f_e \circ \phi$.

Let us write in the following $L^p(\Omega, d\mu)$ to emphasis the dependence on the measure $d\mu$. For an invariant measure $d\mu = \mu' dx$ defined on $B^d$ we define another measure $d\nu$ by

$$d\nu(x) = \mu_e'(x)/\sqrt{x_1 \cdots x_d}\, dx, \qquad x \in T^d.$$

The factor $1/\sqrt{x_1 \cdots x_d}$ is the Jacobian of changing variables $y \mapsto x = \phi(y)$.

**Theorem 3.3** *Let $f \in L^p(B^d, d\mu)$ be invariant under $\mathbb{Z}_2^d$. If $p^*(x)$ is an extremal polynomial for $f$ in $L^p(B^d, d\mu)$, then $p_e^*(x)$ is an extremal polynomial for $f_e$ in $L^p(T^d, d\nu)$. On the other hand, if $g \in L^p(T^d, d\nu)$ and $q^*$ is an extremal polynomial for $f$ in $L^p(T^d, d\nu)$, then $q^* \circ \phi$ is an extremal polynomial for $f \circ \phi$ in $L^p(B^d, d\mu)$.*

*Proof.* If $f \in C(B^d)$ is invariant under $\mathbb{Z}_2^d$ then Theorem 3.2 implies that a polynomial $p^*$ of best approximation to $f$ on $B^d$ is even in each of its variables. Hence, it can be written as $p_e^* \circ \phi$. Using the elementary formula

$$\int_{B^d} f(x_1^2, \ldots, x_d^2) d\mu(x) = \int_{T^d} f(x_1, \ldots, x_d) d\nu(x) \tag{3.3}$$

it follows that $p_e^*$ is the polynomial of best approximation to $f$ on $T^d$. On the other hand, if $q^*$ is an extremal polynomial for $g$ on $T^d$, then $q^* \circ \phi$ is a polynomial in $G\mathcal{P}_{2n}^d$ for $G = \mathbb{Z}_2^d$. Using (3.3) and then Theorem 3.2, we see that $q^* \circ \phi$ is an extremal polynomial for $f \circ \phi$ on $B^d$.

For $d = 2$ this theorem is proved in [4]. The idea of using the correspondence between polynomials on these domains has been used in dealing with orthogonal polynomials and cubature formulas on $B^d$ and $T^d$ ([24]) and it has also been used in [26] to relate the direct and the inverse type theorems for the best approximation on $T^d$ and on $B^d$.

# 4 Best approximation in more than two variables

Comparing to two variables, there is little known in several variables. We discuss what is known and what is involved.

## 4.1 Some basic results

It is easy to see that the following proposition holds (see, for example, [2]).

**Proposition 4.1** *If $F(x_1, \ldots, x_d) = f(x_1, \ldots, x_m, 0, \ldots, 0)$ for some $m < d$, then $E_n(F; \Omega^d) = E_n(f; \Omega^m)$ for $\Omega^d$ being the unit ball $B^d$, the simplex $T^d$ or the sphere $S^{d-1}$.*

*Proof.* Since $B^m$ is a subset of $B^d$, it follows readily that

$$\|F - p\|_{C(B^d)} \ge \|f - p(x_1, \ldots, x_m, 0, \ldots, 0)\|_{C(B^d)} \ge E_n(f; B^m)$$

for any $p \in \Pi_n^d$. Hence, $E_n(F; B^d) \ge E_n(f; B^m)$. On the other hand, if $p^* \in \Pi_n^m$ is an extremal polynomial for $f$ on $B^m$, then $P^*(x_1, \ldots, x_d) = p^*(x_1, \ldots, x_m)$ is a polynomial in $\Pi_n^d$ and $E_n(F; B^d) \le \|F - P^*\|_{C(B^d)} = \|f - p^*\|_{C(B^d)} = E_n(f; B^m)$. The proof for $T^d$ and $S^{d-1}$ is similar.

In particular, using the result in the previous section, this shows that if $\alpha = (\alpha_1, \alpha_2, 0, \ldots, 0)$ then

$$E_n(x^\alpha; B^d) = E_n(x^\alpha; S^{d-1}) = 2^{1-n}, \qquad n = \alpha_1 + \alpha_2 \tag{4.1}$$

and

$$E_n(x^\alpha; T^d) = 2^{1-2n}, \qquad n = \alpha_1 + \alpha_2. \tag{4.2}$$

Note that both these quantities do not depend on the dimension.

The first nontrivial example appears early in [19], which states that

$$\inf_{p \in \Pi_{n-1}^d} \|x_1 \ldots x_d - p\|_{C(S^{d-1})} = \|x_1 \ldots x_d\|_{C(S^{d-1})} = d^{-d/2}.$$

Only in the fairly recent work [1], polynomials of best approximation to other monomials on $B^d$ and on $S^{d-1}$ are found using the invariance of Theorem 3.2. It particular it is shown there that

$$\inf_{p \in \Pi_{n-1}^d} \|x_1 \ldots x_d - p\|_{C(B^d)} = \inf_{p \in \Pi_{n-1}^d} \|x_1 \ldots x_d - p\|_{C(S^{d-1})} = d^{-d/2}. \tag{4.3}$$

Apart from (3.1) and (3.2) there are several other polynomials whose extremal polynomials are found, including $(x_1^4 + \ldots x_d^4)$ among others. However, the only monomial in the list is $x_1 \cdots x_d$.

In [2] the best approximation is found for another monomial $x_1^2 x_2 \ldots x_d$ on the ball and on the sphere,

$$\begin{aligned} E_{d+1}(x_1^2 x_2 \cdots x_d, B^d) &= E_{d+1}(x_1^2 x_2 \cdots x_d, S^{d-1}) \\ &= \|(x_1^2 - a_d) x_2 \cdots x_d\|_{C(S^{d-1})} = \frac{a_d}{(d-1)^{\frac{d-1}{2}}} \end{aligned} \tag{4.4}$$

where $a = a_d$ is the solution of the equation

$$\frac{2}{d+1} \left( \frac{d-1}{d+1} \right)^{\frac{d-1}{2}} (1-a)^{\frac{d+1}{2}} = a.$$

In particular, $a_2 = 1/4$, $a_3 = 3 - \sqrt{8}$, $a_4 = (38 + 5 \times 10^{1/3} - 10^{5/3})/18$.

Recently in [25], we found a family of symmetric polynomials as extremal polynomials for $x_1^2 \cdots x_d^2$ in $B^d$ and in $S^{d-1}$. This work depends heavily on the theorems in the previous section. It starts with several reductions. First, by Theorem 3.3, we only need to consider the best approximation to $x_1 \cdots x_d$ on the simplex $T^d$. Second, by Theorem 3.2, an extremal polynomial can be chosen to be symmetric on $T^d$.

As we mentioned before that there are two steps in dealing with the problem of least polynomials; one is to find a lower bound for the quantity $E_n(x_1 \cdots x_3; T^3)$ and the other is to identify an extremal polynomial that will attain the lower bound. In the case of two variables, we found the lower bound in (2.5) by taking the maximal on a subset that is part of the boundary of $T^2$. This way we reduce the problem to a problem in one variable that is already solved. We can attempt to apply the same method here and take the maximal on the face

$$T_0^d = \{(x_1, \ldots, x_d) \in T^d : x_1 + \ldots + x_d = 1\}$$

of $T^d$. Evidently $x \in T_0^d$ is the same as $(x_1, \ldots, x_{d-1}) \in T^{d-1}$. Let us define $X := (x_1, \ldots, x_{d-1}, 1 - x_1 - \ldots - x_{d-1})$. For $\alpha \in \mathbb{N}_0^d$ define

$$X^\alpha = x_1^{\alpha_1} \cdots x_{d-1}^{\alpha_{d-1}} (1 - x_1 - \cdots - x_{d-1})^{\alpha_d}, \qquad x \in T^d.$$

Then the restriction leads to

$$\begin{aligned} E_n(x^\alpha; T^d) &= \inf_{p \in \Pi_n^d} \|x^\alpha - p\|_{C(T^d)} \\ &\geq \inf_{p \in \Pi_n^{d-1}} \|X^\alpha - p\|_{C(T^{d-1})} = E_n(X^\alpha; T^{d-1}). \end{aligned} \tag{4.5}$$

However, this does not reduce the problem to a lower dimension problem of least polynomial for the monomials, unless $\alpha_d$ (or one of the other component of $\alpha$) is zero. If none of the components of $\alpha$ is zero, then $X^\alpha$ vanishes on the entire boundary of $T^d$. We cannot further reduce the problem by restricting to the boundary of the simplex.

In the simple case of $x_1 x_2 \cdots x_d$ we discuss what to do in the following two subsections. We consider the case $d = 3$ first to illustrate the idea.

## 4.2 Extremal polynomial for $x_1x_2x_3$ on $T^3$

In the case of $x_1x_2x_3$ on $T^3$ the right hand side of the equation (4.5) is a two dimensional problem on $T^2$. Recall that in the previous section we deduced the lower bound of $E_n(x^k y^{n-k}; T^2)$ by further restricting the maximal on the boundary of $T^2$ so that it becomes a problem in one variable. For the function $X_1(x) := x_1x_2(1-x_1-x_2)$, however, such a method will not work as the function is zero on the entire boundary of the triangle $T^2$.

Since $X_1$ is invariant under the symmetric group $\mathcal{S}(T^2)$, Theorem 3.1 shows that the optimal polynomial has to be a symmetric polynomial in $x_1, x_2, 1-x_1-x_2$. Such a polynomial, called $U_3(x_1, x_2)$, is discovered in [25]. It agrees with the Chebyshev polynomial $T_2(2x-1)$ of degree 2 on the boundary of the triangle $T^2$; that is

$$U_3(x,0) = U_3(0,x) = U_3(x,1-x) = T_2(2x-1).$$

**Proposition 4.2** *Let $X_1(x) = x_1x_2(1-x_1-x_2)$. Define the polynomial*

$$U_3(x_1,x_2) = 72x_1x_2(1-x_1-x_2) - 3 + 4(x_1^2 + x_2^2 + (1-x_1-x_2)^2).$$

*Then*

$$E_2(X_1, T^2) = \inf_{p \in T_2^2} \|X_1 - p\|_{C(T^2)} = 72^{-1}\|U_3\|_{C(T^2)} = 72^{-1}.$$

*Proof.* Once the polynomial is identified, we can find the extremal signature. Indeed, define

$$\mathcal{S}_+ = \{x : |U_3(x_1,x_2)| = 1\} = \left\{(0,0),(1,0),(0,1),\left(\frac{1}{3},\frac{1}{3}\right)\right\}$$
$$\mathcal{S}_- = \{x : |U_3(x_1,x_2)| = -1\} = \left\{\left(\frac{1}{2},0\right),\left(0,\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2}\right)\right\}$$

and let $\sigma(v) = 1$ on $\mathcal{S}_+$ and $\sigma(v) = -1$ on $\mathcal{S}_-$. Then $\sigma$ is an extremal signature since the sum $Lf = L_1 f - L_2 f$ defined by

$$L_1 f = \frac{3}{4} f\left(\frac{1}{3}, \frac{1}{3}\right) + \frac{1}{12}(f(0,0) + f(1,0) + f(0,1))$$

and

$$L_2 f = \frac{1}{3}\left(f\left(\frac{1}{2}, 0\right) + f\left(0, \frac{1}{2}\right) + f\left(\frac{1}{2}, \frac{1}{2}\right)\right)$$

satisfies $Lf = 0$ for $f \in \Pi_2^2$. Consequently, since $|U_3(x_1, x_2)| = 72^{-1}$ on $T^2$, it follows from the part 2 of Theorem 3.1 that $E_2(X_1, T^2) \geq 72^{-1}$. To complete the proof we only need to show that $\|U_2\|_{C(T^2)} = 1$, which is an easy exercise.

According to the above proposition and Theorem 3.3, to find the extremal polynomial for $x_1 x_2 x_3$ on $T^3$ amounts to find a symmetric polynomial $R_3(x_1, x_2, x_3)$ such that $R_3$ agrees with $U(x_1, x_2)$ on the face $T_0^3$ of $T_3$. Although this requirement does not determine $R_3$ uniquely, it helps to guide our search. The outcome is given below.

**Theorem 4.3** *Define the polynomial* $R_3(x)$ *by*

$$\begin{aligned} R_3(x) =& 72 x_1 x_2 x_3 - 4(x_1 + x_2 + x_3) + 4(x_1 + x_2 + x_3)^2 \\ & - 8(x_1 x_2 + x_2 x_3 + x_1 x_3) + 1. \end{aligned}$$

*Then* $72^{-1} R_3(x)$ *is the least polynomial for* $x_1 x_2 x_3$ *on* $T^3$ *and*

$$E_2(x_1 x_2 x_3; T^3) = E_2(x_1 x_2 (1 - x_1 - x_2); T^2) = 72^{-1} \|R_3\|_{C(T^3)} = 72^{-1}$$

*Furthermore,* $72^{-1} R_3(x_1^2, x_2^2, x_3^2)$ *is a least polynomial for* $x_1^2 x_2^2 x_3^2$ *on* $B^3$ *and* $S^2$*, and*

$$E_5(x_1^2 x_2^2 x_3^2; B^3) = E_5(x_1^2 x_2^2 x_3^2; S^2) = 72^{-1} \left\|R_3(x_1^2, x_2^2, x_3^2)\right\|_{C(B^3)} = 72^{-1}.$$

By Theorem 3.3 all we need to prove is that $R_3(x_1, x_2, x_3)$ is bounded in $T^3$. The polynomial $R_3$ is bounded on the boundary of $T^3$, since $R_3(x, y, 0) = R_3(x, 0, y) = R_3(0, x, y) = (1 - 2x)^2 + (1 - 2y)^2 - 1$ and $R_3(x, y, 1 - x - y) = U_3(x, y)$. Furthermore, solving critical points shows that $R_3$ attains its maximum and minimum on the boundary.

### 4.3 Extremal polynomial for $x_1 \cdots x_d$ on $T^d$

We now consider the case $d > 3$. By Theorem 3.2 an extremal polynomial can be taken as a symmetric polynomial in $x_1, \dots, x_d$. It is well known that every symmetric polynomial can be written in terms of elementary symmetric polynomials $e_1, e_2, \dots$.

Let $N \in \mathbb{N}$. The elementary symmetric polynomials of degree $k$ in variables $x_1, x_2, \dots, x_N$ are defined by

$$e_k(x) = \sum_{1 \le i_1 < \dots < i_k \le N} x_{i_1} x_{i_2} \cdots x_{i_k}, \qquad 1 \le k \le N.$$

We assume that $N$ is sufficiently large and do not write the dependence of $e_k$ on the number of variables. Using the elementary symmetric polynomials, $R_3$ in Theorem 4.3 takes the form

$$R_3(x) = 72 e_3(x) - 4 e_1(x) + 4 e_1^2(x) - 8 e_2(x) + 1, \qquad x \in \mathbb{R}^3.$$

In the following we will denote the right hand side by $S_3(x)$ for $x \in \mathbb{R}^d$. Notice that for $d > 3$ such a function will be entirely different from $R_3$. We will use the notation $\mathbf{1}^k = (1, 1, \dots, 1) \in \mathbb{R}^k$.

**Definition 4.4** *For $x \in \mathbb{R}^N$ and $N > d$ define $S_3(x)$ by*

$$S_3(x) = 72 e_3(x) - 4 e_1(x) + 4 e_1^2(x) - 8 e_2(x) + 1$$

*and $S_k(x)$ for $k > 3$ by the recursive formula*

$$S_k(x) = r_k e_k(x) - S_{k-1}(x),$$

*where the constant $r_k$ is determined recursively by $S(k^{-1}\mathbf{1}^k) = 1$. Furthermore we denote $S_d(x)$ by $R_d(x)$ if $x \in \mathbb{R}^d$.*

In the above definition the functions $S_k$ are defined for any number of variables. However, the number $r_k$ is determined using the formula for $S_k(x)$ for $x \in \mathbb{R}^k$. Notice that we cannot define $R_d$ without extending the definition of $R_{d-1}$ formally to $d$ variables, which is why we introduce $S_k$ for any number of variables. As an example, we give the explicit formula of $R_4$:

$$\begin{aligned} R_4(x) = {} & 896 x_1 x_2 x_3 x_4 - 72(x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4) \\ & + 8(x_1 x_2 + x_1 x_3 + x_1 x_4 + x_2 x_3 + x_2 x_4 + x_3 x_4) \\ & - 4(x_1 + x_2 + x_3 + x_4)^2 + 4(x_1 + x_2 + x_3 + x_4) - 1. \end{aligned}$$

The definition also shows that

$$R_d(x) = \sum_{k=2}^{d} (-1)^{d-k} r_k e_k(x) + (-1)^{d-3}(4e_1^2(x) - 4e_1(x) + 1)$$

where $r_2 := 8$. For $d \geq 3$, the numbers $r_d$ can be deduced from the following explicit formula ([25])

$$r_d = d \sum_{k=4}^{d} k^{d-3} \binom{d}{k} \left((-1)^k (9k^2 - 32k + 24) + k^2\right).$$

The first few values of $r_d$ and their prime factorization are list below:

$$\begin{aligned}
r_3 &= 72 = 2^3 \cdot 3^2, \\
r_4 &= 896 = 2^7 \cdot 7, \\
r_5 &= 14400 = 2^6 \cdot 3^2 \cdot 5^2, \\
r_6 &= 283392 = 2^8 \cdot 3^3 \cdot 41, \\
r_7 &= 6598144 = 2^9 \cdot 7^2 \cdot 263, \\
r_8 &= 177373184 = 2^{15} \cdot 5413.
\end{aligned}$$

It turns out that $R_d$ play the role of $R_3$ as an extremal polynomial on $T^d$.

**Theorem 4.5** *Let $X_1(x) = x_1 \cdots x_{d-1}(1 - x_1 - \cdots - x_{d-1})$. For $d = 3, 4, 5$,*

$$E_{d-1}(x_1 \cdots x_d; T^d) = E_{d-1}(X_1, T^{d-1}) = r_d^{-1} \|R_d\|_{T^{d-1}} = r_d^{-1},$$

*and, furthermore,*

$$E_{2d-1}(x_1^2 \cdots x_d^2; B^d) = E_{2d-1}(x_1^2 \cdots x_d^2; S^{d-1}) = r_d^{-1} \|R_d\|_{T^{d-1}} = r_d^{-1}.$$

In fact we made the conjecture in [25] that the theorem holds for all $d \geq 3$. Recall that we normally have two steps, one is to find a sharp lower bound for $E_n(f; \Omega)$ and the other is to find an extremal polynomial that attains the lower bound. In our case we have the lower bound:

**Theorem 4.6** *For all $d > 3$*

$$E_{d-1}(x_1 \cdots x_d; T^d) = E_{d-1}(X_1, T^{d-1}) \geq r_d^{-1}, \tag{4.6}$$

*and, furthermore,*

$$E_{2d-1}(x_1^2 \cdots x_d^2; B^d) = E_{2d-1}(x_1^2 \cdots x_d^2; S^{d-1}) \geq r_d^{-1}. \tag{4.7}$$

The proof of this step uses the extremal signature and Theorem 3.1. The definition of $R_d$ shows that

$$R_d(x_1, \ldots, x_{d-1}, 0) = -R_{d-1}(x_1, \ldots, x_{d-1}). \tag{4.8}$$

This equation allows us to identify many points on which $|R_d(x)| = 1$. Such points are invariant under the symmetric group $S_d$ of the coordinates. It turns out that these points are enough for the extremal signature.

By the part 2 of Theorem 3.1, the existence of the extremal signature proves Theorem 4.6. The next step is to find an extremal polynomial that attains the lower bound in (4.6). Clearly, a candidate for such a polynomial is $r_d^{-1} R_d$. What we need is a proof that shows the candidate actually works. Such a proof comes down to showing that $|R_d(x)| \le 1$ on $T^d$. This is proved in [25] for $d = 3, 4, 5$ by computing the critical points of $R_d$ (with the help of a computer algebra system).

It turns out that for $d = 3, 4, 5$, $|R_d(x)|$ is attained exactly in the set $\mathcal{S}_+$ and $\mathcal{S}_-$ defined above; in particular, $|R_d(x)|$ attains its maximum on the face $T_0^d = \{x \in T^d : x_1 + \ldots + x_d = 1\}$. By the definition of $R_d(x)$, it is easy to see that $\Delta R_d(x) = (-1)^{d-1} 8d$ where $\Delta$ denote the usual Laplacian operator. This shows that $(-1)^{d-1} R_d(x)$ is a subharmonic function, which implies that

$$(-1)^{d-1} R_d(x) \le \max_{x \in \partial T^d} (-1)^{d-1} R_d(x)$$

by the maximal principle (cf. [9]). However, this is only half of what we need. Because of the definition of $R_d$, induction argument might play a role in a potential proof. If the maximal of $|R_d(x)|$ were attained in the boundary, then the equation (4.8) shows that the maximal of $R_d$ are attained on the face $T_0^d$ of the boundary of $T^d$. Let

$$U_d(x) = R_d(x_1, \ldots, x_{d-1}, 1 - x_1 - \ldots - x_{d-1}), \qquad x \in \mathbb{R}^{d-1}.$$

We then need to prove that $|U_d(x)| \le 1$ on $T^{d-1}$. The computation in the case of $d = 3, 4, 5$ points out that the $|U_d(x)|$ attains its maximal on the boundary of $T^{d-1}$. If this were true, then by (4.8), the inequality $|U_d(x)| \le 1$ would follow from the induction since $U_d(x_1, \ldots, x_{d-2}, 0) = U_{d-1}(x_1, \ldots, x_{d-1})$ and

$$U_d(x_1, \ldots, x_{d-2}, 1 - x_1 - \ldots - x_{d-2}) = U_{d-1}(x_1, \ldots, x_{d-2}).$$

The above discussion justifies the following conjecture.

**Conjecture 4.7** *For $d \ge 6$, $|R_d(x)|$ attains its maximum on the boundary of $T^d$ and $|U_d(x)|$ attains its maximal on the boundary of $T^{d-1}$.*

If the conjecture were proved, then the equality would hold in (4.6) with $R_d$ as one least polynomial and equality would hold in (4.7) with $R_d \circ \phi$ as one least polynomial.

## 4.4 Extremal polynomial for $x_1^2 x_2^2 x_3^2$ on $T^3$

The construction in the previous section suggests that a least polynomial for $(x_1x_2x_3)^n$ should agree with the Chebyshev polynomial $T_{n-1}(2x-1)$ on the boundary of $T^3$. The following result gives the case of $n = 2$.

**Proposition 4.8** *Let $X_1(x) = x_1x_2(1 - x_1 - x_2)$. Define the polynomial*

$$\begin{aligned} U_5(x_1, x_2) =& 27^2 b X_1^2(x_1, x_2) \\ & + 27b\, X_1(x_1, x_2) \left(3a(x_1^2 + x_2^2 + (1 - x_1 - x_2)^2) - c\right) \\ & + 2 \left[-3 + 4(x_1^2 + x_2^2 + (1 - x_1 - x_2)^2)\right]^2 - 1 \end{aligned}$$

*where $a = 28.5926243...$, $b = 21.8935834...$ and $c = 32/9 + a + b$. Then*

$$E_5(X_2, T^2) = \inf_{p \in \Pi_5^2} \|X_2 - p\|_{C(T^2)} = 27^{-2} b^{-1} \|U_5\|_{C(T^2)} = 27^{-2} b^{-1}.$$

The polynomial $U_5$ is symmetric in $x_1, x_2, 1 - x_1 - x_2$ by Theorem 3.2. The form of $U_5$ is chosen so that it agrees with the Chebyshev polynomial $T_4(2t - 1) = 1 - 8(1 - 2x)^2 + 8(1 - 2x)^4$ on the boundary of $T^2$; that is

$$U_5(x, 0) = U_5(0, x) = U_5(x, 1 - x) = T_2(T_2(2x - 1)) = T_4(2x - 1).$$

The constant $c$ is determined by fixing $U_5(1/3, 1/3) = 1$. The values of $a$ and $b$ are chosen so that $|U_5(x)| \leq 1$ inside $T^2$. This is proved by making $1 - U_5(x, y) \geq 0$ and then solving $U_5(x, x) = -1$ and $U_5'(x, x) = 0$, which yields three point in $T^2$, by symmetry, for which $U_5(x, y) = -1$. It turns out that these three points, together with the maximal and minimal points on the boundary of $T^2$, are enough for an extremal signature.

From the best approximation to $x_1^2 x_2^2 (1 - x_1 - x_2)^2$ on $T^2$ we should get best approximation to $x_1^2 x_2^2 x_3^2$ on $T^3$. One only has to find a least polynomial which satisfies $V_5(x)$ such that $V_5(x_1, x_2, 1 - x_1 - x_2) = U_5(x_1, x_2)$ and $|V_5(x, y)| \leq 1$ on $T^3$. This step is not trivial since an additional multiple of $(x_1 + x_2 + x_3)^k$ added to any term in $V_5(x)$ does not change the value of the polynomial on the boundary of $T^3$. Once $V_5$ is found, then $V_5 \circ \phi$ is an least polynomial for $x_1^4 x_2^4 x_3^4$ on $B^3$ and $S^2$ by Theorem 3.3. The result is stated below.

**Theorem 4.9** *Let $a$ and $b$ be as above. Define polynomial $V_5$ by*

$$V_5 = 27^2 b\, e_3^2 - 1 + 2e_1 - 2e_2^2 + 2\left(1 - 4e_1 + 4e_1^2\right)^2 - 27e_3\left((32/9 - 2a + b)e_1^2 + 6ae_2\right), \quad x \in \mathbb{R}^3.$$

*Then $27^{-2}b^{-1}V_5(x)$ is a least polynomial for $x_1^2x_2^2x_3^2$ on $T^3$*

$$E_5(x_1^2x_2^2x_3^2; T^3) = E_5(x_1^2x_2^2(1-x_1-x_2)^2; T^2) = 27^{-2}b^{-1}\|V_5\|_{C(T^3)} = 27^{-2}b^{-1}.$$

*Furthermore, $72^{-2}b^{-1}V_5(x_1^2, x_2^2, x_3^2)$ is a least polynomial for $x_1^4x_2^4x_3^4$ on $B^3$ and $S^2$,*

$$E_5(x_1^2x_2^2x_3^2; B^3) = E_5(x_1^2x_2^2x_3^2; S^2) = 72^{-2}b^{-1}\left\|V_5(x_1^2, x_2^2, x_3^2)\right\|_{C(S^2)} = 27^{-2}b^{-1}.$$

What comes out as a surprise is that there does not seem to be a nice formula for $E_5(X_2, T^2) = 27^{-2}b^{-1}$. This suggests that the best approximation to monomials in several variables is much more complicated than that of two variables.

The above result suggests many questions that one may want to explore. Some questions are if least polynomials for other monomials, say $x_1^k x_2^l x_3^j$, on $T^3$ also attains their maximum and minimum on the boundary of $T^3$, and if a least polynomial $p$ satisfies that $p(x_1, x_2, 1-x_1-x_2)$ agrees with the Chebyshev polynomials of appropriate degree on the boundary of $T^2$. One can also ask the question that to what extend a polynomial is determined by its values on the boundary of the triangle.

## 5 Extremal polynomials in $L^2$ norm

In this section we report results on the best $L^2$ approximation on the unit sphere $S^{d-1}$, the unit ball $B^d$ and the standard simplex $T^d$. Here the $L^2$ norm can be taken with respect to the measure $d\mu = W(x)dx$ for some weight function $W$ on the domain $\Omega$. We shall denote the error of best approximation to $f$ by $E_n(x^\alpha; \Omega)_{L^2(W)}$ to indicate the dependence on $W$. We call the polynomial $R_\alpha$ a least polynomial for $x^\alpha$ if $R_\alpha = x^\alpha - Q_\alpha$, $Q_\alpha \in \Pi_{|\alpha|-1}^d$, and

$$E_n(x^\alpha; \Omega)_{L^2(W)} = \inf_{p \in \Pi_{|\alpha|-1}^d} \|x^\alpha - p\|_{L^2(\Omega, W)} = \|R_\alpha\|_{L^2(\Omega, W)}.$$

According to the standard Hilbert space theory, the extremal polynomial $R_\alpha$ is orthogonal to all polynomials of degree lower than $|\alpha|$ and it is unique. Such a polynomial is called a monomial orthogonal polynomial in [27] and their explicit expression and their $L^2$ norms are computed explicitly. For earlier study of $L^2$ norm of these polynomials, we refer to [4] and the references therein. The results in [27] are derived for a family of weight functions on these domains, which includes the classical weight functions (see below). It should be pointed out that many of the monomial orthogonal polynomials are already known in [3] for the classical weight functions.

Throughout this section we use the standard multiindex notation. For $\alpha, \beta \in \mathbb{N}_0^m$ we also write $\alpha! = \alpha_1! \cdots \alpha_m!$ and $(\alpha)_\beta = (\alpha_1)_{\beta_1} \cdots (\alpha_m)_{\beta_m}$, where $(a)_n = a(a+1)\ldots(a+n-1)$ is the Pochhammer symbol. Furthermore, for $\alpha \in \mathbb{N}^m$ and $a, b \in \mathbb{R}$, we write $a\alpha + b = (a\alpha_1 + b, \ldots, a\alpha_m + b)$.

## 5.1 Best $L^2$ approximation on $S^{d-1}$

The weight function on $S^{d-1}$ is

$$W_\kappa(x) = c_\kappa \prod_{i=1}^{d} |x_i|^{2\kappa_i}, \qquad \kappa_i \geq 0,$$

where $c_\kappa$ is a normalization constant so that the integral of $W_\kappa$ over $S^{d-1}$ is 1. In the following we write $|\kappa| = \kappa_1 + \ldots + \kappa_d$. Let $\lfloor x \rfloor$ denote the integer part of $x$. For a multiindex $\alpha \in \mathbb{N}_0^d$ we use $\lfloor \alpha/2 \rfloor$ to denote $(\lfloor \alpha_1/2 \rfloor, \ldots, \lfloor \alpha_d/2 \rfloor)$.

Let $F_B$ be the Lauricella hypergeometric series of type $B$, which generalizes the hypergeometric function ${}_2F_1$ to several variables (cf. [6]),

$$F_B(\alpha, \beta; c; x) = \sum_\gamma \frac{(\alpha)_\gamma (\beta)_\gamma}{(c)_{|\gamma|} \gamma!} x^\gamma, \quad \alpha, \beta \in \mathbb{N}_0^d, \quad c \in \mathbb{R}, \quad \max_{1 \leq i \leq d} |x_i| < 1,$$

where the summation is taken over $\gamma \in \mathbb{N}_0^d$. We will also need the generalized Gegenbauer polynomials $C_n^{(\lambda,\mu)}(t)$ defined by

$$C_n^{(\lambda,\mu)}(x) = c_\mu \int_{-1}^{1} C_n^\lambda(xt)(1+t)(1-t^2)^{\mu-1} dt.$$

These polynomials are orthogonal with respect to the weight function

$$w_{\lambda,\mu}(t) = |t|^{2\mu}(1-t^2)^{\lambda-1/2}, \qquad -1 \leq t \leq 1,$$

and they become Gegenbauer polynomials when $\mu = 0$; that is, $C_n^{(\lambda,0)}(t) = C_n^{\lambda}(t)$. We denote by $k_n^{(\lambda,\mu)}$ the leading coefficient of $C_n^{(\lambda,\mu)}(x)$. The explicit formula of the least polynomial $R_\alpha$ for $x^\alpha$ is given as follows.

**Theorem 5.1** *Let $\rho = |\kappa| + (d-2)/2$, $\beta = \alpha - \lfloor \frac{\alpha+1}{2} \rfloor$ and $n = |\alpha|$. Then*

1. $R_\alpha(x) = x^\alpha F_B\left(-\beta, -\alpha+\beta-\kappa+\frac{1}{2}; -n-\rho+1; \frac{\|x\|^2}{x_1^2}, \ldots, \frac{\|x\|^2}{x_d^2}\right).$

2. *The best approximation to $x^\alpha$ is given by*

$$\inf_{p\in\Pi_{n-1}^d} \int_{S^{d-1}} |x^\alpha - p(x)|^2 W_\kappa(x) d\omega = \int_{S^{d-1}} |R_\alpha(x)|^2 W_\kappa(x) d\omega$$

$$= \frac{2\rho\left(\kappa+\frac{1}{2}\right)_\alpha}{(\rho)_{|\alpha|}} \int_0^1 \prod_{i=1}^d \frac{C_{\alpha_i}^{(\frac{1}{2},\kappa_i)}(t)}{k_{\alpha_i}^{(\frac{1}{2},\kappa_i)}} t^{|\alpha|+2\rho-1} dt.$$

In the case of $\kappa_1 = \ldots = \kappa_d = 0$, this is the classical case of $L^2$ approximation with respect to the Lebesgue measure on $S^{d-1}$. The polynomial $R_\alpha$ is defined using a generating function. It also satisfies several other properties, including a recurrence relation. The norm of $R_\alpha$ is written as an integral of one variable. This integral, however, is difficult to evaluate. It can be written as an multiple sum of hypergeometric type. For a detail study and several special cases, see [27].

## 5.2 Best $L^2$ approximation on $B^d$

The weight function on $B^d$ is

$$W_\kappa^B(x) = c_\kappa^B \prod_{i=1}^d |x_i|^{2\kappa_i}(1-\|x\|)^{\kappa_{d+1}-1/2}, \qquad x \in B^d, \quad \kappa_i \geq 0,$$

where $c_\kappa^B$ is a normalization constant so that the integral of $W_\kappa^B$ over $B^d$ is 1. The explicit formula of the least polynomial $R_\alpha^B$ for $x^\alpha$ is given as follows.

**Theorem 5.2** *Let $\rho = |\kappa| + (d-1)/2$. For $\alpha \in \mathbb{N}_0^d$ and $x \in \mathbb{R}^d$,*

1. $R_\alpha^B(x) = x^\alpha F_B\left(-\beta, -\alpha+\beta-\kappa+\frac{1}{2}; -|\alpha|-\rho+1; \frac{1}{x_1^2}, \ldots, \frac{1}{x_d^2}\right).$

*2. The best approximation to $x^\alpha$ is given by*

$$\inf_{p\in\Pi^d_{n-1}} \int_{B^d} |x^\alpha - p(x)|^2 W^B_\kappa(x)dx = \int_{B^d} |R^B_\alpha(x)|^2 W^B_\kappa(x)dx$$
$$= \frac{2\rho\left(\kappa+\frac{1}{2}\right)_\alpha}{(\rho)_{|\alpha|}} \int_0^1 \prod_{i=1}^d \frac{C^{(\frac{1}{2},\kappa_i)}_{\alpha_i}(t)}{k^{(\frac{1}{2},\kappa_i)}_{\alpha_i}} t^{|\alpha|+2\rho-1}dt.$$

This theorem more or less follows from a close relation between orthogonal polynomials on $B^d$ and those on $S^d$ (see [23, 27]). The polynomials correspond to the special case $W(x) = (1-\|x\|^2)^{\mu-1/2}$ are the classical orthogonal polynomials. Let us mention the special case of Lebesgue measure on $B^d$. Let $P_n(t)$ denote the Legendre polynomial of degree $n$. Then

$$\min_{Q\in\Pi^d_{n-1}} \frac{1}{\operatorname{vol} B^d} \int_{B^d} |x^\alpha - Q(x)|^2 dx = \frac{d\,\alpha!}{2^n(d/2)_n} \int_0^1 \prod_{i=1}^d P_{\alpha_i}(t) t^{n+d-1}dt, \tag{5.1}$$

where $\operatorname{vol} B^d = \pi^{d/2}/\Gamma(d/2+1)$ is the volume of $B^d$.

## 5.3 Best $L^2$ approximation on $T^d$

The weight function on $T^d$ is

$$W^T_\kappa(x) = c^T_\kappa \prod_{i=1}^d |x_i|^{\kappa_i-1/2}(1-|x|)^{\kappa_{d+1}-1/2}, \qquad x\in T^d, \quad \kappa_i \ge 0$$

where $c^T_\kappa$ is a normalization constant so that the integral of $W^T_\kappa$ over $T^d$ is 1. We need another generalization of the hypergeometric series ${}_2F_1$ to several variables, the Lauricella function of type $A$, which is defined by (cf. [6])

$$F_A(c,\alpha;\beta;x) = \sum_\gamma \frac{(c)_{|\gamma|}(\alpha)_\gamma}{(\beta)_\gamma \gamma!} x^\gamma, \qquad \alpha,\beta\in\mathbb{N}_0^{d+1}, \quad c\in\mathbb{R},$$

where the summation is taken over $\gamma\in\mathbb{N}_0^{d+1}$. For the simplex $T^d$ we work with the homogeneous coordinates $X = (x_1,\ldots,x_d, 1-x_1-\ldots-x_d)$. For $\alpha\in\mathbb{N}_0^{d+1}$ we can derive the polynomial of best approximation for $X^\alpha$. The explicit formula of the least polynomial $R^T_\alpha$ for $X^\alpha$ is given as follows.

**Theorem 5.3** *Let $\rho = |\kappa| + (d-1)/2$. For $\alpha\in\mathbb{N}_0^d$ and $x\in\mathbb{R}^d$,*

1. *the least polynomial for $X^\alpha$ in the $L^2$ norm is*

$$R_\alpha^T(x) = X^\alpha F_B\Big(-\alpha, -\alpha-\kappa+\frac{1}{2}; -2|\alpha|-\rho+1; \frac{1}{x_1}, \ldots, \frac{1}{x_{d+1}}\Big)$$
$$= (-1)^n \frac{(\kappa+\mathbf{1})_\alpha}{(n+|\kappa|+d)_n} F_A(|\alpha|+|\kappa|+d, -\alpha; \kappa+\mathbf{1}; X).$$

2. *The best approximation to $X^\alpha$ is given by*

$$\inf_{p\in\Pi_{n-1}^d} \int_{T^d} |X^\alpha - p(x)|^2 W_\kappa^T(x)dx = \int_{T^d} |R_\alpha^T(x)|^2 W_\kappa^T(x)dx$$
$$= \frac{\rho\,\alpha!\,\big(\kappa+\frac{1}{2}\big)_\alpha}{(\rho)_{2|\alpha|}} \int_0^1 \prod_{i=1}^{d+1} P_{\alpha_i}^{(0,\kappa_i-1/2)}(2r-1) r^{|\alpha|+\rho-1} dr.$$

This theorem essentially follows from the result on $B^d$, using either Theorem 3.3 or the close relation between orthogonal polynomials on these two domains. Putting $\alpha_{d+1}=0$ gives the best approximation to $x^\alpha$ on $T^d$. It is interesting to note that the additional term $(1-x_1-\ldots-x_d)^k$ causes many problems for the uniform norm but none for the $L^2$ norm. Let us mention the special case of Lebesgue measure on $T^d$. Again $P_n(t)$ denotes the Legendre polynomial of degree $n$. Then for $\alpha\in\mathbb{N}_0^d$ and $n=|\alpha|$,

$$\min_{Q\in\Pi_{n-1}^d} \frac{1}{d!}\int_{T^d} |x^\alpha - Q(x)|^2 dx = \frac{d\,\alpha!^2}{(d)_{2n}} \int_0^1 \prod_{i=1}^{d} P_{\alpha_i}(2r-1) r^{n+d-1} dr. \quad (5.2)$$

The formulas of the $L^2$ norm of the least polynomials in all three cases look simple; for example, see the formulas (5.1) and (5.2). However, these formulas are not easy to evaluate exactly. On the other hand, they seem to provide a way to obtain sharp estimates for these quantities, which remains to be done.

*Acknowledgement*: The author is partially supported by the NSF Grant DMS-0201669.

## References

[1] N. N. Andreev and V. A. Yudin, Polynomials of least deviation from zero, and Chebyshev-type cubature formulas, *Proc. Steklov Inst. Math.* **232** (2001), no. 1, 39–51.

[2] N. N. Andreev and V. A. Yudin, Best approximation of polynomials on the sphere and on the ball, in *Recent progress in multivariate approximation (Witten-Bommerholz, 2000)*, 23–30, Internat. Ser. Numer. Math., 137, Birkhuser, Basel, 2001.

[3] P. Appell and J. K. de Fériet. *Fonctions hypergéométriques et hypersphériques, Polynomes d'Hermite*, Gauthier-Villars, Paris, 1926.

[4] B. D. Bojanov, W. Haussmann, and G. P. Nikolov, Bivariate polynomials of least deviation from zero, *Canad. J. Math.* **53** (2001), 489–505.

[5] L. Bos, On Kergin-interpolation in the disk, *J. Approx. Theory* **37** (1983), 251–261.

[6] H. Exton, *Multiple hypergeometric functions and applications*, Halsted, New York, 1976.

[7] M. I. Ganzburg and S. A. Pichugov, Invariance of the elements of the best approximation and a theorem of Gleaser, *Ukrainian Math. Journal*, **33**, no. 5 (1981), 508–510.

[8] W. B. Gearhart, Some Chebyshev Approximations by polynomials in two variables, *J. Approx. Theory*, **8** (1973), 195-209.

[9] F. John. *Partial Differential Equations*, 4th ed. Springer-Verlag, New York, 1982.

[10] X. Z. Liang, Y. M. Ye, Kergin interpolation at the points which are zeros of the bivariate polynomial of least deviation from zero on the disk, *Northeast. Math. J.* **2** (1986), no. 4, 408–414

[11] V.I. Lebedev Quadrature on a sphere, *USSR Comp. Math. and Math. Phys.* **16** (1977), 10–24.

[12] I. G. Macdonald, *Symmetric functions and Hall polynomials*, 2ed ed. Oxford Mathematical Monographs, Clarendon Press, New York, 1995.

[13] U. Maier, On best approximation of the monomials on the unit ball of $\mathbb{R}^r$, *J. Approx. Theory*, **92** (1998), 74–81.

[14] I. P. Mysovskikh, Interpolatory cubature formulas (in Russian), "Nauka", Moscow, 1981.

[15] D. J. Newman and H. S. Shapiro, Some theorems on Cebysev approximation, *Duke Math. J.* **30** (1963), 673–682.

[16] D. J. Newman and Yuan Xu, Tchebycheff polynomials on a triangular region, *Constr. Approx.* **9** (1993), 543–546.

[17] M. Reimer, On multivariate polynomials of least deviation from zero on the unit ball, *Math. Z*, **1**53 (1977), 51-58.

[18] T. J. Rivlin and H. S. Shapiro, A unified approach to certain problems of approximation and minimization, *J. Soc. Indust. Appl. Math.* **9** (1961), 670–699.

[19] H. S. Shapiro, Some theorems on Cebysev approximation II, *J. Math. Anal. Appl.* **17** (1967), 262–268.

[20] S. L. Sobolev Cubature formulas on the sphere invariant under finite groups of rotations, *Sov. Math. Dokl.* **3** (1962), 1307–1310.

[21] A. Stroud, *Approximate calculation of multiple integrals*, Prentice-Hall, Englewood Cliffs, NJ, 1971.

[22] J. P. Thiran and C. Detaille, On real and complex valued bivariate Chebyshev polynomials, *J. Approx. Theory*, **59** (1989), p 321–337.

[23] Yuan Xu, Orthogonal polynomials and cubature formulae on spheres and on balls, *SIAM J. Math. Anal.* **29** (1998), 779–793.

[24] Yuan Xu, Orthogonal polynomials and cubature formulae on spheres and on simplices, *Methods Anal. and Appl.* **5** (1998), 169–184.

[25] Yuan Xu, On polynomials of least deviation from zero in several variables, *Experimental Math.*, **13** (2004), p.103 -112.

[26] Yuan Xu, Weighted approximation of functions on the unit sphere, *Const. Approx.* **21** (2005), 1-28.

[27] Yuan Xu, Monomial orthogonal polynomials of several variables, J. Approx. Theory, to appear.
`http://www.math.uoregon.edu/~yuan` (Monomial.ps.gz).

Yuan Xu
Department of Mathematics
University of Oregon
Eugene, Oregon 97403-1222.
email: `yuan@math.uoregon.edu`