Global polynomial optimization by norming sets on sphere and torus *

Marco Vianello[1]

February 12, 2018

Abstract

Using the approximation theoretic notion of norming set, we compute $(1-\varepsilon)$-approximations to the global minimum of arbitrary n-th degree polynomials on the sphere, by discrete minimization on approximately $3.2\,n^2\varepsilon^{-1}$ trigonometric grid points, or $2n^2\varepsilon^{-1}$ quasi-uniform points. The same error size is attained by approximately $6.5\,n^2\varepsilon^{-1}$ trigonometric grid points on the torus.

2010 AMS subject classification: 41A17, 65K05, 90C26.

Keywords: polynomial optimization, sphere, torus, polynomial inequalities, norming sets, polynomial meshes, trigonometric grids, Dubiner distance, quasi-uniform points.

1 Introduction

In this paper, following the approach proposed in [23] for the cube, we make a further step in the use of discrete polynomial inequalities as a tool for global polynomial optimization. Such an approach, based on the approximation theory notion of norming set, places in the framework of approximate continuous optimization by discrete optimization on suitable subsets, see e.g. [8, 9, 10, 25] and the references therein.

Given a compact set (or manifold) $\mathcal{S} \subset \mathbb{R}^d$, we briefly recall that a *norming set* in the infinity norm for $\mathbb{P}_n^d(S)$ (the space of d-variate polynomials of degree not exceeding n, restricted to S), is a subset $X \subset \mathcal{S}$ such that

$$\|p\|_{\mathcal{S}} \le c\,\|p\|_X \ , \ \ \forall p \in \mathbb{P}_n^d(\mathcal{S}) \ , \tag{1}$$

for some constant $c > 1$ independent of p. Observe that X is $\mathbb{P}_n^d(\mathcal{S})$-determining, i.e. the only polynomial in $\mathbb{P}_n^d(S)$ which vanishes on X is the null polynomial on S. This implies that $\mathrm{card}(X) \ge N = \dim(\mathbb{P}_n^d(S))$. It can happen that $\dim(\mathbb{P}_n^d(S)) < \dim(\mathbb{P}_n^d) = \binom{n+d}{d}$, in particular this occurs when S is an algebraic variety. For example on the 2-sphere $S = S^2$ we have that $\dim(\mathbb{P}_n^3(S)) = (n+1)^2 < \dim(\mathbb{P}_n^3) = (n+1)(n+2)(n+3)/6$; cf. [7, 24].

*Work partially supported by the DOR funds and the biennial projects CPDA143275 and BIRD163015 of the University of Padova, and by the GNCS-INdAM.

[1]Department of Mathematics, University of Padova, Italy
e-mail: marcov@math.unipd.it

Finite norming sets are of special interest in approximation theory. In the last 10 years, the study of family of norming sets like (1), called *polynomial meshes* when c is independent of n and $\mathrm{card}(X) = \mathcal{O}(N^s)$, $s \geq 1$, has received a considerable attention in the literature on multivariate polynomial approximation (a polynomial mesh is termed *optimal* when $s = 1$). Indeed, polynomial meshes are good discrete models of a compact set, concerning for example polynomial least square approximation, extraction of interpolation point sets, Bernstein-Markov measures and numerical methods in pluripotential theory. Moreover, they can be constructed in an incremental way by algebraic transformation, finite union and product, and are stable under small perturbations; see e.g. [1, 3, 5, 6, 15, 18, 19, 20, 21, 22] and the references therein.

By the following elementary lemma, we recall the basic inequality that allows to use (1) for approximate polynomial optimization on S, provided that we are able to construct X in such a way that $c \to 1$. By no loss of generality, we focus on the approximation of the minimum.

Lemma 1 *Assume that (1) holds. Then*

$$\min_X p - \min_S p \leq (c-1) \left(\max_S p - \min_S p \right) , \ \ \forall p \in \mathbb{P}_n^d(S) . \tag{2}$$

Proof. Consider the nonnegative polynomial $q(x) = \max_S p - p(x) \in \mathbb{P}_n^d(S)$. We have that $\|q\|_S = \max_S p - \min_S p$ and $\|q\|_X = \max_S p - \min_X p$. Then by (1 applied to q

$$\|q\|_S - \|q\|_X = \min_X p - \min_S p \leq (c-1)\,\|q\|_X \leq (c-1)\,\|q\|_S \ ,$$

that is (2). □

Giving an approximation relative to the range of p, namely to $\max_S p - \min_S p$, is standard in polynomial optimization. The usual way to express an inequality like (2), is to say that $\min_X p$ is a $(1-\varepsilon)$-approximation to $\min_S p$, with $\varepsilon = c - 1$; see, e.g., [8].

2 Optimization by norming sets on the sphere

If one is able to construct X in such a way that c becomes close to 1, estimate (2) gives a guaranteed error bound in the approximation of the polynomial minimum on S. In [23], the case $S = [-1,1]^d$ was treated, taking as X a $(mn+1)^d$ Chebyshev-Lobatto grid, and proving that $c = 1/\cos(\xi)$ in (2) with $\xi = \pi/(2m)$. The proof was based on the uniform spacing of the grid points with respect to a suitable distance related to total degree polynomial spaces, the *Dubiner distance* in the *cube*; see [11] and [4]. Embedding the problem in tensor product polynomial spaces, one would obtain instead $c = 1/\cos^d(\xi)$.

Here, we focus on the unit 2-*sphere* $S = S^2$, and partially on the *torus*. Indeed, global optimization on sphere and torus is a classical problem (see e.g. [8, 13]), but apparently not studied yet in the framework of norming sets.

We begin by a *tensorial* trigonometric approach, which is natural working with angular coordinates (latitude and longitude), where we get $c = 1/\cos^2(\xi)$. Such an approach is not peculiar to the sphere, and can for example be immediately extended to polynomial optimization on the *torus*. Then, we'll turn to an approach tailored to polynomial spaces on the sphere, again by the relevant notion of Dubiner distance and recent results on spherical polynomial meshes, obtaining $c = 1/\cos(\xi)$.

2.1 Optimization on trigonometric grids

On the sphere $S = S^2$ it is natural to work with spherical coordinates, say

$$\sigma(\theta, \phi) = (\cos(\theta)\cos(\phi), \cos(\theta)\sin(\phi), \sin(\theta)) \ , \tag{3}$$

where θ is the latitude and ϕ the longitude. The key observation is that if $p \in \mathbb{P}_n^3(S)$, then $p \circ \sigma \in \mathbb{T}_n \otimes \mathbb{T}_n$, where $\mathbb{T}_n$ denotes the space of univariate *trigonometric polynomials* of degree not exceeding n. We recall that $\dim(\mathbb{P}_n^3(S)) = (n+1)^2$, and $\dim(\mathbb{T}_n \otimes \mathbb{T}_n) = (2n+1)^2$. In order to work with trigonometric polynomial inequalities on the whole period, we take both the angles θ, ϕ in $[-\pi, \pi]$. In such a way the transformation σ is no more injective even eliminating the poles (every not polar point corresponds to two values of the extended latitude).

We can now state and prove the following

Proposition 1 *Let $S = S^2$ be the unit 2-sphere and $X = \sigma(\Theta_{mn} \times \Theta_{mn})$, where $\Theta_s = \{j\pi/s\,,\ -s \le j \le s-1\}$ be $2s$ equally spaced angles in $[-\pi, \pi]$. Then*

$$\min_X p - \min_S p \le \tan^2(\xi) \left(\max_S p - \min_S p\right) \ , \ \ \forall p \in \mathbb{P}_n^3(S) \ \ , \ \ \xi = \frac{\pi}{2m} \ . \tag{4}$$

Proof. It suffices resorting to the well-known trigonometric polynomial inequality proved by Ehlich and Zeller in [12] (se also [5])

$$\|t\|_{[0,2\pi]} \le \frac{1}{\cos(n\pi/2s)} \|t\|_{\Theta_s} \ , \ s > n \ , \ \ \forall t \in \mathbb{T}_n \ . \tag{5}$$

From (5) with $s = mn$, $m > 1$, and the fact that $p \circ \sigma \in \mathbb{T}_n \otimes \mathbb{T}_n$, it is immediate to prove in view of the tensorial structure that

$$\|p\|_S \le \frac{1}{\cos^2(\xi)} \|p\|_X \ , \ \ \forall p \in \mathbb{P}_n^3(S) \ , \ \ \xi = \frac{\pi}{2m} \ , \tag{6}$$

where $X = \sigma(\Theta_{mn} \times \Theta_{mn})$, and then to conclude by Lemma 1. □

Observe that due to the structure of the transformation σ (essentially spherical coordinates) the sampling points are not uniformly distributed on the sphere, but cluster at the north and south pole. Moreover $\mathrm{card}(X) \sim 2m^2n^2$ for $mn \to \infty$ (in particular, for n fixed and $m \to \infty$), with exact equality when the poles are not in X (i.e., for mn odd).

It is worth stressing that this construction is not peculiar to the sphere. Indeed, we can apply it to the *torus*, by resorting to the standard toroidal coordinates

$$\sigma(\theta, \phi) = ((R + r\cos(\theta))\cos(\phi), (R + r\cos(\theta))\sin(\phi), r\sin(\theta)) \ , \tag{7}$$

$(\theta, \phi) \in [-\pi, \pi]^2$, where R is the major and r the minor radius of the torus, since also in this case $p \circ \sigma \in \mathbb{T}_n \otimes \mathbb{T}_n$. We may only observe that here $\mathrm{card}(X) = 4m^2n^2$ for any m, n, the transformation σ being now injective.

2.2 Optimization on quasi-uniform points

In view of the highly nonuniform distribution of the sampling points in the trigonometric grid approach, we may think to look for better distributed norming sets on the sphere. The main aid here comes from the notion of *Dubiner distance* on a multidimensional compact set (or manifold) $S \subset \mathbb{R}^d$, that is

$$\delta(x,y) = \sup_{p \in \mathbb{P}_n^d(S)} \left\{ \frac{|\arccos(p(x)) - \arccos(p(y))|}{\deg(p)} \,,\, \deg(p) \geq 1 \,,\, \|p\|_S \leq 1 \right\} . \tag{8}$$

Such a distance was introduced by M. Dubiner in the seminal paper [11], and together with the Markov distance and the Baran distance plays an important role in multivariate polynomial approximation; see e.g. [4] and the references therein for relevant properties and results.

Its role in the construction of norming sets for polynomial spaces can be summarized by the following lemma (cf. [23])

Lemma 2 *Let X be a compact subset of a compact set $S \subset \mathbb{R}^d$ whose covering radius $\rho_\delta(X)$ with respect to the Dubiner distance does not exceed ξ/n, where $\xi \in (0, \pi/2)$ and $n \geq 1$, i.e.*

$$\rho_\delta(X) = \max_{x \in S} \delta(x, X) = \max_{x \in S} \min_{y \in X} \delta(x,y) \leq \xi/n \,. \tag{9}$$

Then, X is a norming set like (1) for $\mathbb{P}_n^d(S)$ with constant $c = 1/\cos(\xi)$.

Proof. We sketch the proof, that appears essentially in [2], for the only purpose of clarifying the role of the Dubiner distance in the construction of norming sets. Assume by no loss of generality that $\|p\|_S = p(x_0) = 1$ for some $x_0 \in S$ (possibly normalizing and/or multiplying p by -1). By definition of Dubiner distance, there exists $y \in X$ such that

$$|\arccos(p(x_0)) - \arccos(p(y))| = |\arccos(p(y))| \leq \frac{\xi\, deg(p)}{n} \leq \xi < \frac{\pi}{2} \,.$$

Then, since the inverse cosine function is monotonically decreasing, we get $p(y) \geq \cos(\xi) > 0$, and finally

$$\|p\|_S = 1 \leq \frac{p(y)}{\cos\xi} \leq \frac{1}{\cos\xi} \|p\|_X \,. \quad \square$$

We notice that Lemma 2 improves a similar inequality in [18], where $c = 1/(1-\xi)$ with $\xi \in (0,1)$ appears instead of $c = 1/\cos(\xi)$. This result was used in [23], to show that a $(mn+1)^d$ Chebyshev-Lobatto grid is a norming set for the cube $[-1,1]^d$ with $c = 1/\cos(\pi/2m)$. Indeed, the Dubiner distance for the cube is $\delta(x,y) = \max\{|\arccos(x_1) - \arccos(y_1)|, \ldots, |\arccos(x_d) - \arccos(y_d)|\}$, and this is one of the few instances where it is explicitly known (the other ones are the sphere, the ball and the simplex, see [4]).

Here we use the fact that the *Dubiner distance on the sphere* $S = S^2$ is nothing but the classical *geodesic distance*

$$\delta(x,y) = \arccos(\langle x, y \rangle) \,, \quad \forall x, y \in S^2 \,, \tag{10}$$

where $\langle x, y\rangle$ denotes the Euclidean scalar product in $\mathbb{R}^3$; cf. [11]. This allows to show that good covering configurations for the sphere are norming sets for $\mathbb{P}_n^3(S)$, and can be used for polynomial optimization in the spirit of Lemma 1.

We recall that a sequence of finite point configurations $X_M \subset S^2$, with cardinality $M \geq 2$, is termed a "good covering" of the sphere if its covering radius with respect to the Euclidean distance in $\mathbb{R}^3$ satisfies the inequality

$$\rho(X_M) = \max_{x\in S} \min_{y\in X_M} |x-y| \leq \frac{\alpha}{\sqrt{M}} \;, \tag{11}$$

for some $\alpha > 0$ (see, e.g., the excellent survey paper [14]). We can now give the following

Proposition 2 *Let $S = S^2$ be the unit 2-sphere and let $\{X_M\}$ be a good covering of S. Take $X = X_M$ with $M = \lceil g(mn)\rceil$, where*

$$g(u) = \left(\frac{2\alpha}{\pi}\,\frac{4u^2}{4u-1}\right)^2 = \left(\frac{2\alpha}{\pi}\right)^2 u^2\,(1+\mathcal{O}(u^{-1}))\;,\;\; u\to\infty\;. \tag{12}$$

Then

$$\min_X p - \min_S p \leq \frac{1-\cos(\xi)}{\cos(\xi)}\left(\max_S p - \min_S p\right)\;,\;\;\forall p\in\mathbb{P}_n^3(S)\;\;,\;\;\xi = \frac{\pi}{2m}\;. \tag{13}$$

Proof. By (11) and elementary considerations of spherical geometry, for every $x \in S$ there exists $y(x) \in X$ such that the estimate

$$\delta(x, y(x)) = 2\arcsin\left(\frac{|x-y(x)|}{2}\right) \leq 2\arcsin\left(\frac{\alpha}{2\sqrt{M}}\right)$$

holds, provided that $\alpha/(2\sqrt{M}) \leq 1$ that is $\sqrt{M} \geq \alpha/2$, where δ is the geodesic distance, which coincides with the Dubiner distance. In view of Lemma 2, to get a norming set with constant $c = 1/\cos(\xi)$ it is sufficient to fulfill the inequality $\delta(x, y(x)) \leq \xi/n$, that is

$$\frac{\alpha}{2\sqrt{M}} \leq \sin\left(\frac{\xi}{2n}\right)\;.$$

Now, by the elementary trigonometric inequality $\sin(t) \geq t(1 - t/\pi)$ we have that the former is satisfied if

$$\frac{\alpha}{2\sqrt{M}} \leq \frac{\xi}{2n}\left(1-\frac{\xi}{2\pi n}\right)\;,$$

which gives $M \geq g(mn)$ by easy calculations. □

Observe that $\mathrm{card}(X) \sim (2\alpha/\pi)^2 m^2 n^2$ for $mn \to \infty$ (in particular, for n fixed and $m \to \infty$). In order to get a good point distribution for polynomial optimization on the sphere, we can consider "quasi-uniform" coverings, that are configurations with bounded ratio between the covering radius and the point separation (mesh ratio). Indeed, these configurations provide a low information redundancy. In [14] several quasi-uniform configurations are discussed, in particular the *zonal equal area configurations* (the centers of zonal equal area partitions of the sphere), that turn out to be both, quasi-uniform and equidistributed in the sense of the surface area measure. They turn out to be theoretically good covering with $\alpha = 3.5$ (but the numerical experiments suggest $\alpha = 2.5$, cf. [16]). Moreover, the Matlab toolbox [17] provides efficient algorithms for their computation.

Notice that with such configurations by Proposition 2 and the fact that $(1-\cos(\xi))/\cos(\xi) \sim \xi^2/2$, $\xi \to 0$, we get a relative error bound on the polynomial minimum $\varepsilon_m \sim \pi^2/(8m^2)$, $m \to \infty$, with $\mathrm{card}(X) \sim (7/\pi)^2 m^2 n^2 \approx 5m^2n^2$ quasi-uniform points (but in practice we can take $\alpha = 2.5$ and thus approximately $(5/\pi)^2 m^2 n^2 \approx 2.5\, m^2 n^2$ points).

In order to compare with the trigonometric grid approach of Proposition 1, we may observe that there $\tan^2(\xi) \sim \xi^2$, $\xi \to 0$, and hence we have to replace m by $\lceil \sqrt{2}\, m \rceil$ to get an error bound $\varepsilon_m \sim \pi^2/(8m^2)$, $m \to \infty$, with a corresponding cardinality $\mathrm{card}(X) \sim 2(\sqrt{2}\, m)^2 n^2 = 4m^2n^2$.

We may summarize the considerations above under a slightly different point of view. For a fixed error tolerance $\varepsilon > 0$, in order to get a $(1-\varepsilon)$-approximation to the minimum of a trivariate polynomial of degree n on the sphere, setting $\varepsilon = \pi^2/(8m^2)$, it is sufficient to sample at approximately $4(8/\pi^2)n^2\varepsilon^{-1} \approx 3.24\, n^2\varepsilon^{-1}$ trigonometric grid points (that is $\approx 1.8\, n\varepsilon^{-1/2}$ equispaced nodes per spherical coordinate), or alternatively $(7/\pi)^2(8/\pi^2)n^2\varepsilon^{-1} \approx 5n^2\varepsilon^{-1}$ quasi-uniform points, the centers of a zonal equal area partition (but in practice we can take $(5/\pi)^2(8/\pi^2)n^2\varepsilon^{-1} \approx 2n^2\varepsilon^{-1}$ points). On the other hand, on the torus the same error is attained with approximately $8(8/\pi^2)n^2\varepsilon^{-1} \approx 6.48\, n^2\varepsilon^{-1}$ trigonometric grid points (that is $\approx 2.55\, n\varepsilon^{-1/2}$ equispaced nodes per toroidal coordinate).

For the purpose of illustration, we compute the minimum of the 4-th degree trivariate polynomial

$$p(x_1,x_2,x_3) = (x_1-a)^2(x_2-b)^2 + (x_3-c)^2 \; , \tag{14}$$

on the unit sphere and on the torus with $R = 1$ and $r = 1/3$, where (a,b,c) is a random point on each surface (obtained via a uniform bivariate random variable in the corresponding angular coordinates); such a minimum is clearly zero. In Table 1 we display, for some values of ε, the average over 100 trials for (a,b,c) of the relative errors

$$E = \left(\min_X p - \min_S p\right) / \left(\max_S p - \min_S p\right) \; , \tag{15}$$

where X is either a trigonometric grid with $\lceil 1.8\, n\varepsilon^{-1/2} \rceil = \lceil 7.2\, \varepsilon^{-1/2} \rceil$ equispaced nodes per spherical coordinate, or a quasi-uniform mesh of the sphere (the centers of a zonal equal area partition computed by [17]) with $2n^2\varepsilon^{-1} = 32\varepsilon^{-1}$ points. In the last row we report the average errors of minimization on a trigonometric grid of the torus, with $\lceil 2.55\, n\varepsilon^{-1/2} \rceil = \lceil 10.2\, \varepsilon^{-1/2} \rceil$ equispaced nodes per toroidal coordinate; see the discussion above. Notice that in all instances ε turns out to be an overestimate of the actual errors.

We may conclude by observing that the norming set approach, which is in some sense a "brute force" optimization method, could be useful not only by its direct application, but also to provide starting guesses for more sophisticated optimization procedures, like for example that discussed in [13].

Table 1: Average errors over 100 trials for (a, b, c) in the global minimization of the 4th-degree polynomial (14) by norming sets on sphere and torus.

	ε	10^{-1}	10^{-2}	10^{-3}	10^{-4}
sphere	E_{av}^{trig}	1e-3	4e-5	3e-6	3e-7
	$E_{av}^{q.u.}$	5e-4	6e-5	5e-6	5e-7
torus	E_{av}^{trig}	6e-4	2e-5	8e-6	2e-7

References

[1] T. Bloom, N. Levenberg, F. Piazzon and F. Wielonsky, Bernstein-Markov: a survey, Dolomites Res. Notes Approx. DRNA 8 (2015), 75–91.

[2] L. Bos, A Simple Recipe for Modelling a d-cube by Lissajous curves, Dolomites Res. Notes Approx. DRNA 10 (2017), 1–4.

[3] L. Bos, S. De Marchi, A. Sommariva and M. Vianello, Weakly Admissible Meshes and Discrete Extremal Sets, Numer. Math. Theory Methods Appl. 4 (2011), 1–12.

[4] L. Bos, N. Levenberg and S. Waldron, Pseudometrics, distances and multivariate polynomial inequalities, J. Approx. Theory 153 (2008), 80–96.

[5] L. Bos and M. Vianello, Low cardinality admissible meshes on quadrangles, triangles and disks, Math. Inequal. Appl. 15 (2012), 229–235.

[6] J.P. Calvi and N. Levenberg, Uniform approximation by discrete least squares polynomials, J. Approx. Theory 152 (2008), 82–100.

[7] D.A. Cox, J. Little and D. O'Shea, Ideals, varieties, and algorithms, fourth edition, Springer, 2015.

[8] E. de Klerk, The complexity of optimizing over a simplex, hypercube or sphere: a short survey, CEJOR Cent. Eur. J. Oper. Res. 16 (2008), 111–125.

[9] E. de Klerk, J.B. Lasserre, M. Laurent and Z. Sun, Bound-constrained polynomial optimization using only elementary calculations, Math. Oper. Res., published online on March 15, 2017.

[10] E. de Klerk, M. Laurent, Z. Sun and J.C. Vera, On the convergence rate of grid search for polynomial optimization over the simplex, Optim. Lett. 11 (2017), 597–608.

[11] M. Dubiner, The theory of multidimensional polynomial approximation, J. Anal. Math. 67 (1995), 39–116.

[12] H. Ehlich and K. Zeller, Schwankung von Polynomen zwischen Gitter punkten, Math. Z. 86 (1964), 41–44.

[13] M. Gräf and R. Hielscher, Fast Global Optimization on the Torus, the Sphere, and the Rotation Group, SIAM J. Optim. 25 (2015), 540–563.

[14] D.P. Hardin, T. Michaels and E.B. Saff, A Comparison of Popular Point Configurations on S^2, Dolomites Res. Notes Approx. DRNA 9 (2016), 16–49.

[15] A. Kroó, On optimal polynomial meshes, J. Approx. Theory 163 (2011), 1107–1124.

[16] P. Leopardi, A partition of the unit sphere into regions of equal area and small diameter, Electron. Trans. on Numer. Anal. 25 (2006), 309–327.

[17] P. Leopardi, Recursive Zonal Equal Area Sphere Partitioning Toolbox, `http://eqsp.sourceforge.net`.

[18] P. Leopardi, A. Sommariva and M. Vianello, Optimal polynomial meshes and Caratheodory-Tchakaloff submeshes on the sphere, Dolomites Res. Notes Approx. DRNA 10 (2017), 18–24.

[19] F. Piazzon, Optimal polynomial admissible meshes on some classes of compact subsets of $\mathbb{R}^d$, J. Approx. Theory 207 (2016), 241–264.

[20] F. Piazzon, Pluripotential Numerics, arXiv preprint 1704.03411, Constr. Approx., under revision.

[21] F. Piazzon and M. Vianello, Small perturbations of polynomial meshes, Appl. Anal. 92 (2013), 1063–1073.

[22] F. Piazzon and M. Vianello, Sub-optimal polynomial meshes on planar Lipschitz domains, Numer. Funct. Anal. Optim. 35 (2014), 1467–1475.

[23] F. Piazzon and M. Vianello, A note on total degree polynomial optimization by Chebyshev grids, Optim. Lett. 12 (2018), 63–71.

[24] I.H. Sloan and R.S. Womersley, Constructive Polynomial Approximation on the Sphere, J. Approx. Theory 103 (2000), 91–118.

[25] J.F. Zhang and C.P. Kwong, Some applications of a polynomial inequality to global optimization, J. Optim. Theory Appl. 127 (2005), 193–205.