APPROXIMATING A NORM BY A POLYNOMIAL

Alexander Barvinok

Abstract. We prove that for any norm $\|\cdot\|$ in the d-dimensional real vector space V and for any odd $n>0$ there is a non-negative polynomial $p(x)$, $x \in V$ of degree $2n$ such that

$$p^{\frac{1}{2n}}(x) \leq \|x\| \leq \binom{n+d-1}{n}^{\frac{1}{2n}} p^{\frac{1}{2n}}(x).$$

Corollaries and polynomial approximations of the Minkowski functional of a convex body are discussed.

1. Introduction and the Main Result

Our main motivation is the following general question. Let us fix a norm $\|\cdot\|$ in a finite-dimensional real vector space V (or, more generally, the Minkowski functional of a convex body in V). Given a point $x \in V$, how fast can one compute or approximate $\|x\|$? For example, various optimization problems can be posed this way. As is well known (see, for example, Lecture 3 of [1]), any norm in V can be approximated by an ℓ_2 norm in V within a factor of $\sqrt{\dim V}$. From the computational complexity point of view, an ℓ_2 norm of x is just the square root of a positive definite quadratic form p in x and hence can be computed "quickly", that is, in time polynomial in $\dim V$ for any $x \in V$ given by its coordinates in some basis of V. Note that we do not count the time required for "preprocessing" the norm to obtain the quadratic form p, as we consider the norm fixed and not a part of the input. It turns out that by employing higher degree forms p, we can improve the approximation: for any $c>0$, given an $x \in V$, one can approximate $\|x\|$ within a factor of $c\sqrt{\dim V}$ in time polynomial in $\dim V$. This, and some other approximation results follow easily from our main theorem.

(1.1) Theorem. *Let V be a d-dimensional real vector space and let $\|\cdot\| : V \longrightarrow \mathbb{R}$ be a norm in V. For any odd integer $n>0$ there exists a homogeneous polynomial*

Key words and phrases. norm, approximation, polynomial, John's ellipsoid, computational complexity.

This research was partially supported by NSF Grant DMS 9734138.

$p : V \longrightarrow \mathbb{R}$ of degree $2n$ such that $p(x) \geq 0$ (in fact, p is the sum of squares of homogeneous polynomials of degree n) and

$$p^{\frac{1}{2n}}(x) \leq \|x\| \leq \binom{n+d-1}{n}^{\frac{1}{2n}} p^{\frac{1}{2n}}(x)$$

for all $x \in V$.

We observe that the approximation factor $\binom{n+d-1}{n}^{\frac{1}{2n}}$ approaches 1 as long as $n/d \longrightarrow \infty$, see also Section 2.2.

We prove Theorem 1.1 in Section 3. In Section 2, we discuss some corollaries of Theorem 1.1. In Section 4, we show how to extend Theorem 1.1 to Minkowski functionals of arbitrary convex bodies.

(1.2) Best approximating polynomials. Let us fix a norm $\|\cdot\|$ in a finite-dimensional vector space V and a positive integer n. One can ask what is the smallest possible constant $C = C\big(\|\cdot\|, n\big)$ for which there exists a polynomial p of degree $2n$ such that

$$p^{\frac{1}{2n}}(x) \leq \|x\| \leq C p^{\frac{1}{2n}}(x) \quad \text{for all} \quad x \in V. \tag{1.2.1}$$

Moreover, what is the value of

$$C(d, n) = \sup_{\substack{\|\cdot\| \text{ is a norm in } V \\ \text{and } \dim V = d}} C\big(\|\cdot\|, n\big).$$

Theorem 1.1 asserts that $C(d, n) \leq \binom{n+d-1}{n}^{\frac{1}{2n}}$ for odd n. It is not known whether the equality holds except in the case of $n = 1$ when indeed $C(d, 1) = \sqrt{d}$ (see, for example, Lecture 3 of [1]). The following simple observation can be useful to determine what a best approximating polynomial may look like.

Suppose that there is a set of finitely many non-negative polynomials p_i with $\deg p_i = 2n$ which satisfy (1.2.1). Thus we have

$$p_i(x) \leq \|x\|^{2n} \leq C^{2n} p_i(x) \quad \text{for all} \quad x \in V \quad \text{and for every} \quad p_i. \tag{1.2.2}$$

Then any convex combination p of polynomials p_i satisfies (1.2.2) and hence (1.2.1). Suppose now that the normed space V possesses a compact group G of linear isometries. If a polynomial p satisfies (1.2.1) then, for any $g \in G$, the polynomial $p_g(x) = p(gx)$ satisfies (1.2.1) with the same constant C and hence, by averaging over G, we can choose a G-invariant polynomial p which satisfies (1.2.1). Hence if a norm $\|\cdot\|$ is invariant under the action of a compact group, we can always choose an invariant best approximating polynomial.

2. Corollaries and Remarks

(2.1) Approximation by polynomials of a fixed degree. Let us fix an n in Theorem 1.1. Then, as d grows, the value of $p^{\frac{1}{2n}}(x)$ approximates $\|\cdot\|$ within a factor of $c_n\sqrt{d}$, where $c_n \approx (n!)^{-\frac{1}{2n}} \approx \sqrt{e/n}$. Since for any fixed n, computation of $p(x)$ takes a $d^{O(n)}$ time, for any $c > 0$ we obtain a polynomial time algorithm to approximate $\|x\|$ within a factor of $c\sqrt{d}$ (again, we do not count the time required for preprocessing, that is, to find the polynomial p).

One may wonder whether a significantly better approximation factor, for example $C(d,n) = O\big(d^{\frac{1}{2n}}\big)$, can be achieved in Theorem 1.1 (that would also agree with the $\sqrt{d}$ bound in the classical case of $n = 1$). This, however, does not seem to be the case as the following example shows. Let $\|\cdot\|$ be the ℓ_1 norm in $\mathbb{R}^d$, that is

$$\|x\| = \sum_{i=1}^{d} |\xi_i| \quad \text{for} \quad x = (\xi_1, \ldots, \xi_d).$$

The norm $\|\cdot\|$ is invariant under signed permutations of the coordinates

$$(\xi_1, \ldots, \xi_d) \longmapsto (\pm\xi_{i_1}, \ldots \pm \xi_{i_d})$$

and hence, as discussed in Section 1.2, we can choose an invariant best approximating polynomial p. Hence p is a symmetric polynomial in $\xi_1^2, \ldots, \xi_d^2$. In particular, for $n = 3$, we have

$$p(x) = \alpha_d \sum_{i=1}^{d} \xi_i^6 + \beta_d \sum_{1 \leq i \neq j \leq d} \xi_i^2 \xi_j^4 + \gamma_d \sum_{1 \leq i < j < k \leq d} \xi_i^2 \xi_j^2 \xi_k^2$$

for some real α_d, β_d and γ_d. Since we must have $0 \leq p(x) \leq \|x\|^6$, by substituting $x = (1, 0, \ldots, 0)$, $x = (1, 1, 0, \ldots, 0)$ and $x = (1, 1, 1, 0, \ldots, 0)$ we get

$$\begin{gathered} 0 \leq \alpha_d \leq 1, \quad 0 \leq 2\alpha_d + 2\beta_d \leq 64 \quad \text{and} \quad 0 \leq 3\alpha_d + 6\beta_d + \gamma_d \leq 729, \\ \text{which implies that} \quad \alpha_d \leq 1, \quad \beta_d \leq 32 \quad \text{and} \quad \gamma_d \leq 735. \end{gathered}$$

Substituting $x = (1, \ldots, 1)$, we observe that $\|x\| = d$ and that $p(x) = O(d^3)$. Therefore, we must have $C(d, 3) \geq C(\ell_1, 3) \geq c\sqrt{d}$ for some absolute constant $c > 0$.

(2.2) Linear growth of the degree. If we allow n to grow linearly with d, we can get a constant factor approximation. Indeed, if we choose $n = \gamma d$ for some $\gamma > 0$ in Theorem 1.1, for large d we have

$$C_0(\gamma) = \binom{n+d-1}{n}^{\frac{1}{2n}} \approx \exp\left\{\frac{1}{2} \ln \frac{\gamma + 1}{\gamma} + \frac{1}{2\gamma} \ln(\gamma + 1)\right\}. \tag{2.2.1}$$

Thus $p^{\frac{1}{2n}}(x)$ approximates $\|\cdot\|$ within a factor of $C_0(\gamma)$ depending on γ alone. In particular, if $\gamma \longrightarrow \infty$, the approximation factor approaches 1. Since for any fixed $\gamma > 0$, computation of $p(x)$ takes $2^{O(d)}$ time, for any constant $c > 1$ we can get an algorithm of $2^{O(d)}$ complexity approximating the value of $\|x\|$ within a factor of c.

The anonymous referee noticed that a different constant factor approximation can be achieved via the following construction. Let V^* be the dual space and let

$$B^* = \Big\{ f \in V^* : \quad f(x) \leq 1 \quad \text{for all} \quad x \in V \quad \text{such that} \quad \|x\| \leq 1 \Big\}$$

be the unit ball of the dual norm in V^*. As is known, (see, for example, Lemma 4.10 of [3]), for any $0 < \delta < 1$ one can choose a set N of $|N| \leq (1 + 2/\delta)^d$ points in B^* which form a δ-net (in the dual norm). Given an integer n, let us define the polynomial p by

$$p(x) = \frac{1}{|N|} \sum_{f \in N} f^{2n}(x).$$

Let us fix a $\gamma > 0$. It is not hard to show that if $\delta = \delta(\gamma)$ is chosen in the optimal way, as long as $n = \gamma d$, the value of $p^{\frac{1}{2n}}(x)$ approximates $\|x\|$ within a constant factor $C_1(\gamma)$. Interestingly, for $\gamma \longrightarrow \infty$, the asymptotics of (2.2.1) and $C_1(\gamma)$ coincide:

$$C_0(\gamma), C_1(\gamma) = 1 + \frac{\ln \gamma}{2\gamma}\Big(1 + o(1)\Big) \quad \text{as} \quad \gamma \longrightarrow \infty.$$

However, for small γ, the bound $C_0(\gamma)$ of (2.2.1) is substantially better than the one obtained for $C_1(\gamma)$ using this construction. We have

$$\begin{aligned} &C_0(\gamma) = \sqrt{e/\gamma}\big(1 + o(1)\big) \quad \text{for} \quad \gamma \longrightarrow 0 \quad \text{as opposed to} \\ &C_1(\gamma) = 3^{\frac{1}{2\gamma}\big(1+o(1)\big)} \quad \text{for} \quad \gamma \longrightarrow 0. \end{aligned}$$

(2.3) Approximating by other computable functions. It is possible that one can achieve a better approximation by employing a wider class of computable functions. For example, the ℓ_1 norm which appears to be resistant to polynomial approximations (cf. Section 2.1) is very easy to compute. A natural candidate would be the class of functions which are sums of $p_i^{\frac{1}{2n}}$ for different polynomials p_i. In particular, the ℓ_1 norm itself is a function of this type.

3. Proof of Theorem 1.1

Let B be the unit ball of $\|\cdot\|$, so

$$B = \big\{ x \in V : \|x\| \leq 1 \big\}.$$

Hence B is a convex compact set containing the origin in its interior and symmetric about the origin.

Let V^* be the dual space of all linear functions $f: V \longrightarrow \mathbb{R}$ and let $B^* \subset V^*$ be the polar of B:

$$B^* = \Big\{ f \in V^* : f(x) \leq 1 \quad \text{for all} \quad x \in B \Big\}.$$

Hence B^* is a convex compact set symmetric about the origin. Using the standard duality argument, we can write

$$\|x\| = \max_{f \in B^*} f(x). \tag{3.1}$$

Let

$$W = V^{\otimes n} = \underbrace{V \otimes \ldots \otimes V}_{n \text{ times}} \quad \text{and} \quad W^* = \left(V^{\otimes n}\right)^* = \underbrace{V^* \otimes \ldots \otimes V^*}_{n \text{ times}}$$

be the n-th tensor powers of V and V^* respectively.

For vectors $x \in V$ and $f \in V^*$ let

$$x^{\otimes n} = \underbrace{x \otimes \ldots \otimes x}_{n \text{ times}} \quad \text{and} \quad f^{\otimes n} = \underbrace{f \otimes \ldots \otimes f}_{n \text{ times}}$$

denote the n-th tensor power $x^{\otimes n} \in W$ and $f^{\otimes n} \in W^*$ respectively.

By (3.1), we can write

$$\|x\|^n = \max_{f \in B^*} \big(f(x)\big)^n = \max_{f \in B^*} f^{\otimes n}(x^{\otimes n}). \tag{3.2}$$

Let D be the convex hull of $f^{\otimes n}$ for $f \in B^*$:

$$D = \operatorname{conv}\big\{ f^{\otimes n} : f \in B^* \big\}.$$

Then D is a convex compact subset of W^*, symmetric about the origin (we use that n is odd). From (3.2) we can write

$$\|x\|^n = \max_{f \in B^*} f^{\otimes n}(x^{\otimes n}) = \max_{g \in D} g(x^{\otimes n}). \tag{3.3}$$

Let us estimate the dimension of D. There is a natural action of the symmetric group S_n in W^* which permutes the factors V^*, so that

$$\sigma(f_1 \otimes \ldots \otimes f_n) = f_{\sigma^{-1}(1)} \otimes \ldots \otimes f_{\sigma^{-1}(n)}.$$

Let $\operatorname{Sym}(W^*) \subset W^*$ be the *symmetric part* of W^*, that is, the invariant subspace of that action. As is known, the dimension of $\operatorname{Sym}(W^*)$ is that of the space of homogeneous polynomials of degree n in d real variables (cf., for example, Lecture 6 of [2]). Next, we observe that $f^{\otimes n} \in \operatorname{Sym}(W^*)$ for all $f \in V^*$ and, therefore,

$$\dim D \leq \dim \operatorname{Sym}(W^*) = \binom{n+d-1}{n} \tag{3.4}$$

Let E be the John ellipsoid of D in the affine hull of D, that is the (unique) ellipsoid of the maximum volume inscribed in D. As is known, (see, for example, Lecture 3 of [1])

$$E \subset D \subset \big(\sqrt{\dim D}\big) E.$$

Combining this with (3.3), we write

$$\max_{g \in E} g(x^{\otimes n}) \leq \|x\|^n \leq \big(\sqrt{\dim D}\big) \max_{g \in E} g(x^{\otimes n})$$

and, by (3.4),

$$\max_{g \in E} g(x^{\otimes n}) \leq \|x\|^n \leq \binom{n+d-1}{n}^{\frac{1}{2}} \max_{g \in E} g(x^{\otimes n}). \tag{3.5}$$

Let

$$q(x) = \max_{g \in E} g(x^{\otimes n}).$$

We claim that $p(x) = q^2(x)$ is a homogeneous polynomial in x of degree $2n$. Indeed, let us choose a basis $e_1, \ldots, e_d$ in V. Then W acquires the basis

$$e_{i_1 \ldots i_n} = e_{i_1} \otimes \ldots \otimes e_{i_n} \quad \text{for} \quad 1 \leq i_1, \ldots, i_n \leq d.$$

Geometrically, V and V^* are identified with $\mathbb{R}^d$ and W and W^* are identified with $\mathbb{R}^{d^n}$. Let $K \subset W^*$ be the Euclidean unit ball defined by the inequality

$$K = \Big\{ h \in W^* : \sum_{1 \leq i_1, \ldots, i_n \leq d} h^2(e_{i_1 \ldots i_n}) \leq 1 \Big\}.$$

Since E is an ellipsoid, there is a linear transformation $T : W^* \longrightarrow W^*$ such that $T(K) = E$. Let $T^* : W \longrightarrow W$ be the conjugate linear transformation and let $y = T^*\big(x^{\otimes n}\big)$. Hence the coordinates $y_{i_1 \ldots i_n}$ of y with respect to the basis $\big\{e_{i_1 \ldots i_n}\big\}$ are polynomials in x of degree n. Then

$$\begin{aligned} q(x) &= \max_{g \in E} g(x^{\otimes n}) = \max_{h \in K} T(h)\big(x^{\otimes n}\big) = \max_{h \in K} h\big(T^*(x^{\otimes n})\big) \\ &= \max_{h \in K} h(y) = \sqrt{\sum_{1 \leq i_1, \ldots, i_n \leq d} y^2_{i_1 \ldots i_n}}. \end{aligned}$$

Hence we conclude that $p(x) = q^2(x)$ is a homogeneous polynomial in x of degree $2n$, which is non-negative for all $x \in V$ (moreover, $p(x)$ is a sum of squares). From (3.5), we conclude that

$$p^{\frac{1}{2n}}(x) \leq \|x\| \leq \binom{n+d-1}{n}^{\frac{1}{2n}} p^{\frac{1}{2n}}(x),$$

as claimed.

4. An Extension to Minkowski Functionals

There is a version of Theorem 1.1 for Minkowski functionals of convex bodies which are not necessarily symmetric about the origin.

(4.1) Theorem. *Let V be a d-dimensional real vector space, let $B \subset V$ be a convex compact set containing the origin in its interior and let $\|x\| = \inf\{\lambda > 0 : x \in \lambda B\}$ be its Minkowski functional. For any odd integer $n > 0$ there exist a homogeneous polynomial $p : V \longrightarrow \mathbb{R}$ of degree $2n$ and a homogeneous polynomial $r : V \longrightarrow \mathbb{R}$ of degree n such that $p(x) \geq 0$ and*

$$\left(r(x) + \sqrt{p(x)}\right)^{\frac{1}{n}} \leq \|x\| \leq \left(r(x) + \binom{n+d-1}{n}\sqrt{p(x)}\right)^{\frac{1}{n}}$$

for all $x \in V$.

Proof. The proof follows the proof of Theorem 1.1 with some modifications. Up to (3.4) no essential changes are needed (note, however, that now we have to use that n is odd in (3.2)). Then, since the set D is not necessarily symmetric about the origin, we can only find an ellipsoid E (centered at the origin) of W^* and a point $w \in D$, such that

$$E \subset D - w \subset (\dim D)E,$$

see, for example, Lecture 3 of [1]. Then (3.5) transforms into

$$\max_{g \in E} g(x^{\otimes n}) \leq \|x\|^n - w(x^{\otimes n}) \leq \binom{n+d-1}{n} \max_{g \in E} g(x^{\otimes n}).$$

Denoting

$$p(x) = \left(\max_{g \in E} g(x^{\otimes n})\right)^2 \quad \text{and} \quad r(x) = w(x^{\otimes n})$$

we proceed as in the proof of Theorem 1.1. $\square$

Acknowledgment

The author is grateful to the anonymous referee for useful comments and interesting questions.

References

1. K. Ball, *An elementary introduction to modern convex geometry*, Flavors of Geometry, Math. Sci. Res. Inst. Publ., vol. 31, Cambridge Univ. Press, Cambridge, 1997, pp. 1–58.
2. W. Fulton and J. Harris, *Representation Theory. A First Course*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991.
3. G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry*, Cambridge tracts in mathematics, vol. 94, Cambridge University Press, Cambridge, 1989.

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1109
E-mail address: barvinok@umich.edu