Low rank approximation of polynomials

Alexander Schrijver[1]

Abstract. Let $k \leq n$. Each polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ can be uniquely written as $p = \sum_\mu \mu p_\mu$, where μ ranges over the set M of all monomials in $\mathbb{R}[x_1, \ldots, x_k]$ and where $p_\mu \in \mathbb{R}[x_{k+1}, \ldots, x_n]$. If p is d-homogeneous and $\varepsilon > 0$, we say that p is *ε-concentrated on the first k variables* if

$$\sum_{\substack{\mu \in M \\ \deg(\mu) < d}} \max_{\substack{x \in \mathbb{R}^{n-k} \\ \|x\|=1}} p_\mu(x)^2 \leq \varepsilon \|p\|^2,$$

where $\|p\|$ is the Bombieri norm of p. We show that for each $d \in \mathbb{N}$ and $\varepsilon > 0$ there exists $k_{d,\varepsilon}$ such that for each n and each d-homogeneous $p \in \mathbb{R}[x_1, \ldots, x_n]$ there exists $k \leq k_{d,\varepsilon}$ such that p is ε-concentrated on the first k variables *after some orthogonal transformation of $\mathbb{R}^n$.* (So $k_{d,\varepsilon}$ is independent of the number n of variables.)

We derive this as a consequence of a more general theorem on low rank approximation of polynomials.

1. Introduction

A polynomial is said to have *rank* 1 if it is nonzero and a product of linear functions. The *rank* of a polynomial p, denoted by rank(p), is the minimum number of rank 1 polynomials that add up to p. Low rank of polynomials helps in computing them. Therefore, we investigate in how much polynomials can be approximated by low rank polynomials. In particular, we consider d-homogeneous polynomials (i.e., homogeneous polynomials of total degree d).

In fact there are a number of different notions of rank function of a polynomial, but for what follows we can take any of them, as long as the function is invariant under orthogonal transformations of the polynomial and it does not increase if we set variables in the polynomial to 0. So one can also take the minimum number of linear functions that generate p, or the *Waring rank*: the minimum number of powers of linear functions that linearly span p. Also one may take the minimum number of rank 1 polynomials of Bombieri norm at most 1 that add up to p (for the definition of Bombieri norm see below). More strongly, one can take any monotonically nondecreasing function of any rank function.

We recall the *Bombieri norm* $\|p\|$ of a d-homogeneous polynomials $p \in \mathbb{R}[x_1, \ldots, x_n]$:

$$(1) \qquad \|p\| := \Big(\sum_{\alpha \in \mathbb{N}^n} \binom{d}{\alpha_1, \ldots, \alpha_n}^{-1} p_\alpha^2 \Big)^{1/2},$$

where, for $\alpha \in \mathbb{N}^n$, p_α be the coefficient of $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.[2]

---

[2]The Bombieri norm is the tensor norm if we view d-homogeneous polynomials as symmetric d-tensors. In our estimates we can replace the Bombieri norm equivalently by the square root of the sum of the squares of the coefficients, as we fix d and as the two values are bounded by each other up to a factor of $d!$. However, the Bombieri norm behaves better algebraically — in particular, it is invariant under orthogonal transformations of $\mathbb{R}^n$.

Let us remark that it is not true that for each $d \in \mathbb{N}$ and $\varepsilon > 0$ there is a $k \in \mathbb{N}$ such that for each d-homogeneous polynomial p there exists a d-homogeneous polynomial q of rank at most k such that $\|p - q\| \leq \varepsilon \|p\|$. This is shown by the polynomials $\sum_{i=1}^{n} x_i^2$.

However, the approximation becomes valid by considering the *operator norm*:

$$(2) \qquad \|p\|_{\mathrm{o}} := \max_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} |p(x)|.$$

Then Fernandez de la Vega, Kannan, Karpinski, and Vempala [1] showed (it is also a special case of the 'weak regularity for Hilbert spaces' in Lovász and Szegedy [2])

(3)   for each $d \in \mathbb{N}$ and $\varepsilon > 0$ there exists $k \in \mathbb{N}$ such that for each d-homogeneous polynomial p (in any number of variables) there exists a d-homogeneous polynomial q of rank $\leq k$ such that $\|p - q\|_{\mathrm{o}} \leq \varepsilon \|p\|$ for each x.

Important in (3) is that k is independent of the number of variables. In fact, in [1] it is shown that one can take $k = \lfloor \varepsilon^{-2} \rfloor$.

In this paper we give an extension of (3), using the compactness result of [3] as main tool.

2. Stronger low rank approximation

To describe the extension, let for any n-variable polynomial p and any subspace V of $\mathbb{R}^n$, π_V denote the orthogonal projection onto V, and $p_V := p \circ \pi_V$. Then $\|p\|_{\mathrm{o}}$ may be alternatively described as the supremum of $\|p_V\|$ taken over all 1-dimensional subspaces of $\mathbb{R}^n$. As an extension, define for any $k \in \mathbb{N}$:

$$(4) \qquad \|p\|_{(k)} := \sup_{\substack{V \text{ subspace} \\ \dim V \leq k}} \|p_V\|.$$

So $\|p\|_{\mathrm{o}} = \|p\|_{(1)}$. Moreover, as $\|p_V\| \leq \|p\|$ for any subspace V, we have $\|p\|_{(k)} \leq \|p\|$.

Let P be the set of all d-homogeneous polynomials in $\mathbb{R}[x_1, x_2, \ldots]$, each using only a finite number of variables. Call a function $r : P \to \mathbb{R}$ *monotone* if it is invariant under orthogonal transformations of the space and if it does not increase if we set variables in a polynomial to 0. Each of the above rank functions is monotone, and also each of the above norms. Moreover, if r is monotone and $f : \mathbb{R} \to \mathbb{R}$ is monotonically increasing, then also $f \circ r$ is monotone.

Then in (3), one may replace $\|p-q\|_{\mathrm{o}}$ by $\|p-q\|_{f(\mathrm{rank}(q))}$, for any fixed function $f : \mathbb{N} \to \mathbb{N}$ given in advance. In fact:

Theorem 1. *For each $d \in \mathbb{N}$, $\varepsilon > 0$, and monotone $r : P \to \mathbb{N}$, there exists $k_{d,\varepsilon,r}$ such that for each $p \in P$ there exists $q \in P$ with $r(q) \leq k_{d,\varepsilon,r}$ and $\|p - q\|_{(r(q))} \leq \varepsilon \|p\|$.*

Proof. Let $d \in \mathbb{N}$, $\varepsilon > 0$, and $r : P \to \mathbb{N}$ be given. As P has an inner product associated with the Bombieri norm $\|.\|$, the completion H of P is a Hilbert space. As for each k,

$\|.\|_{(k)} \leq \|.\|$, the norm $\|.\|_{(k)}$ extends uniquely to H. Let $B(P)$ and $B(H)$ be the closed unit balls of P and H respectively.

Then for each fixed k, the norm $\|.\|_{(k)}$ is continuous with respect to the $\|.\|_o$-topology on H. This follows from the fact that there exists a c such that $\|p\|_{(k)} \leq c\|p\|_o$ for each $p \in P$. Indeed, consider any subspace V of dimension k, which we may assume to be $\mathbb{R}^k$. Now the collection C of all d-homogeneous polynomials $p \in \mathbb{R}^k$ with $\|p\|_o \leq 1$ is bounded. Otherwise, as C is a compact convex set in finite dimensions, there would be a nonzero p with $\|p\|_o = 0$. However, if $p \neq 0$, then $p(x) \neq 0$ for some x. Hence there exists a c such that $\|p\| \leq c$ for all $p \in C$, which is c as required.

Define, for each $q \in P$,

$$(5) \qquad U_q := \{p \in B(H) \mid \|p - q\|_{(r(q))} < \varepsilon\}.$$

So U_q is open in the $\|.\|_o$-topology. Moreover, the U_q for $q \in B(P)$ cover $B(H)$. Indeed, for any $p \in B(H)$ there exists $q \in B(P)$ with $\|p - q\| < \varepsilon$. Then $\|p - q\|_{(r(q))} \leq \|p - q\| < \varepsilon$, so $p \in U_q$.

Let G be the group of all transformations that consist of an orthogonal transformation of $\mathbb{R}^n$ for some n, leaving the other coordinates invariant. Then G acts naturally on H. By [3], the orbit space $(B(H), \|.\|_o)/G$ is compact. (This is the quotient topological space of the topological space $(B(H), \|.\|_o)$, taking the orbits of G as quotient classes.) Hence there is a finite set $Q \subseteq B(P)$ such that voor each $p \in B(P)$ there exist $q \in Q$ and $\psi \in G$ such that $p^\psi \in U_q$. Let $k_{d,\varepsilon,f} := \max\{r(q) \mid q \in Q\}$. We show that $k_{d,\varepsilon,f}$ is as required.

Let $p \in P$, say $p \in \mathbb{R}[x_1, \ldots, x_n]$. We may assume that $\|p\| = 1$. Then there exist $q \in Q$ and $\psi \in G$ such that $p^\psi \in U_q$. So $\|p^\psi - q\|_{(r(q))} < \varepsilon$. As $\|.\|_{(r(q))}$ is G-invariant, this gives, setting $\varphi := \psi^{-1}$, $\|p - q^\varphi\|_{(r(q))} < \varepsilon$. Let q' be the orthogonal projection of q^φ onto $\mathbb{R}[x_1, \ldots, x_n]$. So $r(q') \leq r(q^\varphi) = r(q) \leq k_{d,\varepsilon,f}$. Then $\|p - q'\|_{(r(q'))} < \varepsilon$, for let U be a subspace of $\mathbb{R}^n$ of dimension $\leq r(q')$. Then $(q')_U = (q^\varphi)_U$, and hence $\|(p - q')_U\| = \|(p - q^\varphi)_U\| \leq \|p - q^\varphi\|_{(r(q'))} < \varepsilon$. ∎

3. ε-concentration on first k variables

Let $p \in \mathbb{R}[x_1, \ldots, x_n]$ be d-homogeneous, and let $k \leq n$. For $\alpha \in \mathbb{N}^k$, denote $x^\alpha := x_1^{\alpha_1} \cdots x_k^{\alpha_k}$. Then p can be uniquely written as

$$(6) \qquad p = \sum_{\substack{\alpha \in \mathbb{N}^k \\ |\alpha| = d}} x^\alpha p_\alpha,$$

where $p_\alpha \in \mathbb{R}^{n-k}$ and $|\alpha| := \alpha_1 + \cdots + \alpha_k$. We say that p is ε-concentrated on $x_1, \ldots, x_k$ if

$$(7) \qquad \sum_{\substack{\alpha \in \mathbb{N}^k \\ |\alpha| < d}} \|p_\alpha\|_o^2 \leq \varepsilon\|p\|.$$

Corollary 1a. *For each d and $\varepsilon > 0$ there exists $k_{d,\varepsilon}$ such that for each n and each d-homogeneous $p \in \mathbb{R}[x_1,\ldots,x_n]$ there exist $k \leq k_{d,\varepsilon}$ and an orthogonal transformation φ of $\mathbb{R}^n$ such that p^φ is ε-concentrated on the first k variables.*

Proof. Let $f : \mathbb{N} \to \mathbb{N}$ be defined by $f(k) := k+$ the number of k-variable monomials of degree $< d$. For any d-homogeneous polynomial q, let $w(q)$ be its Waring rank (the minimum number of d-powers of linear functions that span q), and Let $r(q) := f(w(q))$. Let $k_{d,\varepsilon} := k_{d,\varepsilon/d!,r}$, where the latter is taken from Theorem 1. We show that $k_{d,\varepsilon}$ is as required.

Let $p \in \mathbb{R}[x_1,\ldots,x_n]$, and let $q \in P$ be as given by Theorem 1. Then $k := w(q) \leq r(q) \leq k_{d,\varepsilon} = k_{d,\varepsilon/d!,r}$. As q has Waring rank k, there exists a k-dimensional subspace U of $\mathbb{R}^n$ such that $q = q_U$. By applying an orthogonal transformation of $\mathbb{R}^n$, we can assume that $U = \mathbb{R}^k$. So $q \in \mathbb{R}[x_1,\ldots,x_k]$. We prove that p satisfies (7).

For each $\alpha \in \mathbb{N}^k$ with $|\alpha| < d$, choose $z_\alpha \in \mathbb{R}^{n-k}$ with $\|z_\alpha\| = 1$ maximizing $|p_\alpha(z_\alpha)|$. Let V be the space spanned by $\mathbb{R}^k$ and by the z_α. Then $\|p_\alpha\|_\mathrm{o} = \|p_\alpha(z_\alpha)\| \leq |(p_V)_\alpha\| = \|(p_V - p_U)_\alpha\|$ for each such α. Moreover, $\dim(V) \leq f(k) = f(w(q)) = r(q)$. So

$$
\tag{8}
\begin{aligned}
\sum_{\substack{\alpha \in \mathbb{N}^k \\ |\alpha| < d}} \|p_\alpha\|_\mathrm{o}^2 &\leq \sum_{\substack{\alpha \in \mathbb{N}^k \\ |\alpha| \leq d}} \|(p_V - p_U)_\alpha\|^2 = \sum_{\substack{\alpha \in \mathbb{N}^k \\ |\alpha| \leq d}} \sum_{\substack{\beta \in \mathbb{N}^{n-k} \\ |\beta| = d-|\alpha|}} \binom{d-|\alpha|}{\beta_1,\ldots,\beta_{n-k}}^{-1} (p_V - p_U)_{\alpha,\beta}^2 \leq \\
d! \sum_{\substack{\alpha \in \mathbb{N}^k \\ |\alpha| \leq d}} \binom{d}{\alpha_1,\ldots,\alpha_k}^{-1} &\sum_{\substack{\beta \in \mathbb{N}^{n-k} \\ |\beta| = d-|\alpha|}} \binom{d-|\alpha|}{\beta_1,\ldots,\beta_{n-k}}^{-1} (p_V - p_U)_{\alpha,\beta}^2 = d! \|p_V - p_U\|^2 \leq \\
d! \|p_V - q\|^2 &= d! \|(p-q)_V\|^2 \leq d! \|p-q\|_{(r(q))}^2 \leq \varepsilon^2 \|p\|^2.
\end{aligned}
$$

Here $\|p_V - p_U\| \leq \|p_V - q\|$ follows from the fact that $q = q_U$ and $p_U = (p_V)_U$, so p_U is the polynomial defined on U closest to p_V. ∎

Let $\|p\|_\infty$ be the maximum absolute value of the coefficients of p. (This norm is not invariant under orthogonal transformations.) We note that for any d there is a c such that $\|p\|_\infty \leq c\|p\|$ for each d-homogeneous polynomial. This follows from the facts that $\|p\|_\infty = \|p_V\|_\infty$ for some d-dimensional subspace V of $\mathbb{R}^n$ and that the set C of d-homogeneous polynomials in $\mathbb{R}[x_1,\ldots,x_d]$ with $\|p\| \leq 1$ is bounded. This implies that in Corollary 1a one may replace $\|p_\alpha\|_\mathrm{o}$ by $\|p_\alpha\|_\infty$.

References

[1] W. Fernandez de la Vega, R. Kannan, M. Karpinski, S. Vempala, Tensor decomposition and approximation schemes for constraint satisfaction problems, in: *Proceedings of the 37th Annual ACM Symposium on Theory of Computing* (STOC'05), pp. 747–754, ACM, New York, 2005.

[2] L. Lovász, B. Szegedy, Szemerédi's lemma for the analyst, *Geometric and Functional Analysis* 17 (2007) 252–270.

[3] G. Regts, A. Schrijver, Compact orbit spaces in Hilbert spaces and limits of edge-colouring models, preprint, 2012. ArXiv http://arxiv.org/abs/1210.2204