arXiv:1012.3271v1 [math.OC] 15 Dec 2010

# BEST $\ell_1$-APPROXIMATION OF NONNEGATIVE POLYNOMIALS BY SUMS OF SQUARES

JEAN B. LASSERRE

Abstract. Given a nonnegative polynomial $f$, we provide an explicit expression for its best $\ell_1$-norm approximation by a sum of squares of given degree.

## 1. Introduction

This note is concerned with the cone of nonnegative polynomials and its subcone of polynomials that are sums of squares (s.o.s.). Understanding the difference between these two cones is of practical importance because if on the one hand nonnegative polynomials are ubiquitous, on the other hand s.o.s. polynomials are much easier to handle. For instance, and in contrast with nonnegative polynomials, checking whether a given polynomial is s.o.s. can be done efficiently by solving a semidefinite program, a powerful technique of convex optimization.

A negative result by Blekherman [3] states that when the degree is fixed, there are much more nonnegative polynomials than sums of squares and the gap between the two corresponding cones increases with the number of variables. On the other hand, if the degree is allowed to vary, it has been known for some time that the cone of s.o.s. polynomials is dense (for the $\ell_1$-norm of coefficients) in the cone of polynomials nonnegative on the box $[-1,1]^n$. See e.g. Berg, Christensen and Ressel [1] and Berg [2]. However, [1] was essentially an existence result and subsequently, Lasserre and Netzer [5] have provided a very simple and explicit sequence of s.o.s. polynomials converging for the $\ell_1$-norm to a given nonnegative polynomial $f$.

In this note we provide an explicit expression for the *best* $\ell_1$-norm approximation of a given nonnegative polynomial $f\in\mathbb{R}[\mathbf{x}]$ by a s.o.s. polynomial $g$ of given degree $2d$ ($\geq \deg f$). It turns out that

$$g \,=\, f+\lambda_0^*+\sum_{i=1}^n \lambda_i^* x_i^{2d},$$

for some nonnegative vector $\lambda^*\in\mathbb{R}^{n+1}$, very much like the approximation already provided in [5] (where the $\lambda_i^*$'s are equal). In addition, the vector $\lambda^*$ is an optimal solution of an explicit semidefinite program, and so can be computed efficiently.

## 2. Main result

2.1. **Notation and definitions.** Let $\mathbb{R}[\mathbf{x}]$ (resp. $\mathbb{R}[\mathbf{x}]_d$) denote the ring of real polynomials in the variables $\mathbf{x}=(x_1,\ldots,x_n)$ (resp. polynomials of degree at most $d$), whereas $\Sigma[\mathbf{x}]$ (resp. $\Sigma[\mathbf{x}]_d$) denotes its subset of sums of squares (s.o.s.) polynomials (resp. of s.o.s. of degree at most $2d$). For every $\alpha\in\mathbb{N}^n$ the notation

1991 *Mathematics Subject Classification.* 11E25 12D15 41A50 90C22.
*Key words and phrases.* Polynomials; sums of squares; semidefinite programming.

$\mathbf{x}^\alpha$ stands for the monomial $x_1^{\alpha_1}\cdots x_n^{\alpha_n}$ and for every $i\in\mathbb{N}$, let $\mathbb{N}^p_d := \{\beta\in\mathbb{N}^n : \sum_j \beta_j \leq d\}$ whose cardinal is $s(d)=\binom{n+d}{n}$. A polynomial $f\in\mathbb{R}[\mathbf{x}]$ is written

$$\mathbf{x}\mapsto f(\mathbf{x}) \,=\, \sum_{\alpha\in\mathbb{N}^n} f_\alpha\,\mathbf{x}^\alpha,$$

and $f$ can be identified with its vector of coefficients $\mathbf{f}=(f_\alpha)$ in the canonical basis $(\mathbf{x}^\alpha)$, $\alpha\in\mathbb{N}^n$. Hence, denote by $\|f\|_1$ the $\ell_1$-norm $\sum_\alpha |f_\alpha|$ of the coefficient vector $\mathbf{f}$ which also defines a norm on $\mathbb{R}[\mathbf{x}]_d$.

Let $\mathcal{S}^p\subset\mathbb{R}^{p\times p}$ denote the space of real $p\times p$ symmetric matrices. For any two matrices $\mathbf{A},\mathbf{B}\in\mathcal{S}^p$, the notation $\mathbf{A}\succeq 0$ (resp. $\succ 0$) stands for $\mathbf{A}$ is positive semidefinite (resp. positive definite), and the notation $\langle\mathbf{A},\mathbf{B}\rangle$ stands for trace $\mathbf{AB}$.

Let $\mathbf{v}_d(\mathbf{x})=(\mathbf{x}^\alpha)$, $\alpha\in\mathbb{N}^n_d$, and let $\mathbf{B}_\alpha\in\mathbb{R}^{s(d)\times s(d)}$ be real symmetric matrices such that

$$\mathbf{v}_d(\mathbf{x})\,\mathbf{v}_d(\mathbf{x})^T \,=\, \sum_{\alpha\in\mathbb{N}^n_{2d}} \mathbf{x}^\alpha\,\mathbf{B}_\alpha. \tag{2.1}$$

Recall that a polynomial $g\in\mathbb{R}[\mathbf{x}]_{2d}$ is a s.o.s. if and only if there exists a real positive semidefinite matrix $\mathbf{X}\in\mathbb{R}^{s(d)\times s(d)}$ such that

$$g_\alpha \,=\, \langle\mathbf{X},\mathbf{B}_\alpha\rangle,\qquad \forall\alpha\in\mathbb{N}^n_{2d}.$$

**$d$-moment matrix.** With a sequence $\mathbf{y}=(y_\alpha)$, $\alpha\in\mathbb{N}^n$, let $L_\mathbf{y}:\mathbb{R}[\mathbf{x}]\to\mathbb{R}$ be the linear functional

$$f\quad(=\sum_\alpha f_\alpha\,\mathbf{x}^\alpha)\quad\mapsto\quad L_\mathbf{y}(f)\,=\,\sum_\alpha f_\alpha\,y_\alpha,\quad f\in\mathbb{R}[\mathbf{x}].$$

With $d\in\mathbb{N}$, the $d$-moment matrix associated with $\mathbf{y}$ is the real symmetric matrix $\mathbf{M}_d(\mathbf{y})$ with rows and columns indexed in $\mathbb{N}^n_d$, and defined by:

$$\mathbf{M}_d(\mathbf{y})(\alpha,\beta) \,:=\, L_\mathbf{y}(\mathbf{x}^{\alpha+\beta}) \,=\, y_{\alpha+\beta},\qquad \forall\alpha,\beta\in\mathbb{N}^n_d. \tag{2.2}$$

It is straightforward to check that

$$\left\{\, L_\mathbf{y}(g^2)\geq 0\quad\forall g\in\mathbb{R}[\mathbf{x}]_d\,\right\}\quad\Leftrightarrow\quad \mathbf{M}_d(\mathbf{y})\,\succeq\,0,\quad d=0,1,\ldots.$$

**Semidefinite programming.** A semidefinite program is a convex (more precisely convex conic) optimization problem of the form $\min_\mathbf{X}\{\langle\mathbf{C},\mathbf{X}\rangle : \mathcal{A}\,\mathbf{X}=\mathbf{b};\,\mathbf{X}\succeq 0\}$, for some real symmetric matrices $\mathbf{C},\mathbf{X}\in\mathcal{S}^p$, vector $\mathbf{b}\in\mathbb{R}^m$, and some linear mapping $\mathcal{A}:\mathcal{S}^p\to\mathbb{R}^m$. Semidefinite programming is a powerful technique of convex optimization, ubiquitous in many areas. A semidefinite program can be solved efficiently and even in time polynomial in the input size of the problem, for fixed arbitrary precision. For more details the interested reader is referred to e.g. [7].

2.2. **The result.** Consider the following optimization problem:

$$\rho_d := \min_g \{\, \|f-g\|_1 \,:\, g\in\Sigma[\mathbf{x}]_d\,\}, \tag{2.3}$$

that is, one searches for the best $\ell_1$-approximation of $f$ by a s.o.s. polynomial of degree at most $2d\,(\geq \deg f)$. Of course, and even though (2.3) is well defined for an arbitrary $f\in\mathbb{R}[\mathbf{x}]$, such a problem is of particular interest when $f$ is nonnegative but not a s.o.s.

**Theorem 2.1.** *Let $f \in \mathbb{R}[\mathbf{x}]$ and let $2d \geq \deg f$. The best $\ell_1$-norm approximation of $f$ by a s.o.s. polynomial of degree at most $2d$ is given by*

$$\mathbf{x} \mapsto g(\mathbf{x}) \,=\, f(\mathbf{x}) + \,(\lambda_0^* + \sum_{i=1}^n \lambda_i^* \, x_i^{2d}), \tag{2.4}$$

*for some nonnegative vector $\lambda^* \in \mathbb{R}^{n+1}$. Hence $\rho_d = \sum_{i=0}^n \lambda_i^*$, and in addition, $\lambda^*$ is an optimal solution of the semidefinite program:*

$$\min_{\lambda \geq 0} \left\{ \sum_{i=0}^n \lambda_i \,:\, f + \lambda_0 + \sum_{i=1}^n \lambda_i \, x_i^{2d} \in \Sigma[\mathbf{x}]_d \right\}. \tag{2.5}$$

*Proof.* Consider $f$ as an element of $\mathbb{R}[\mathbf{x}]_{2d}$ by setting $f_\alpha = 0$ whenever $|\alpha| > \deg f$ (where $|\alpha| = \sum \alpha_i$), and rewrite (2.3) as the semidefinite program:

$$\begin{array}{rlll} \rho_d := \displaystyle\min_{\lambda \geq 0, \mathbf{X} \succeq 0, g} & \displaystyle\sum_{\alpha \in \mathbb{N}^n_{2d}} \lambda_\alpha & & \\ \text{s.t.} & \lambda_\alpha + g_\alpha & \geq f_\alpha, & \forall \alpha \in \mathbb{N}^n_{2d} \\ & \lambda_\alpha - g_\alpha & \geq -f_\alpha, & \forall \alpha \in \mathbb{N}^n_{2d} \\ & g_\alpha - \langle \mathbf{X}, \mathbf{B}_\alpha \rangle & = 0, & \forall \alpha \in \mathbb{N}^n_{2d}. \end{array} \tag{2.6}$$

The dual semidefinite program of (2.6) reads:

$$\left\{ \begin{array}{rrll} \displaystyle\max_{u_\alpha, v_\alpha \geq 0, \mathbf{y}} & \displaystyle\sum_{\alpha \in \mathbb{N}^n_d} f_\alpha (u_\alpha - v_\alpha) & & \\ \text{s.t.} & u_\alpha + v_\alpha & \leq 1 & \forall \alpha \in \mathbb{N}^n_{2d} \\ & u_\alpha - v_\alpha + y_\alpha & = 0 & \forall \alpha \in \mathbb{N}^n_{2d}, \\ & \mathbf{M}_d(\mathbf{y}) & \succeq 0, & \end{array} \right. \tag{2.7}$$

or, equivalently,

$$\left\{ \begin{array}{ll} \displaystyle\max_{\mathbf{y}} & -L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_d(\mathbf{y}) \succeq 0 \\ & |y_\alpha| \leq 1, \quad \forall \alpha \in \mathbb{N}^n_{2d}. \end{array} \right. \tag{2.8}$$

The semidefinite program (2.8) has an optimal solution $\mathbf{y}^*$ because the feasible set is compact. In addition, let $\mathbf{y} = (y_\alpha)$ be the moment sequence of the measure $d\mu = \mathrm{e}^{-\|\mathbf{x}\|^2} d\mathbf{x}$, scaled so that $|y_\alpha| < 1$ for all $\alpha \in \mathbb{N}^n_{2d}$. Then $(\mathbf{y}, \mathbf{u}, \mathbf{v})$ with $\mathbf{u} = -\min[\mathbf{y}, 0]$ and $\mathbf{v} = \max[\mathbf{y}, 0]$, is strictly feasible in (2.7) because $\mathbf{M}_d(\mathbf{y}) \succ 0$, and so Slater's condition[1] holds for (2.7). Therefore, by a standard duality result in convex optimization, there is no duality gap between (2.6) and (2.7) (or (2.8)), and (2.6) has an optimal solution $(\lambda^*, \mathbf{X}^*, g^*)$. Hence $\rho_d = -L_{\mathbf{y}^*}(f)$ for any optimal solution $\mathbf{y}^*$ of (2.8).

[1]Slater's condition holds the conic optimization problem $\mathbf{P}:\ \min_{\mathbf{x}} \{\mathbf{c}'\mathbf{x} \,:\, \mathbf{A}\mathbf{x} = \mathbf{b};\ \mathbf{x} \in \mathbf{K}\}$, where $\mathbf{K} \subset \mathbb{R}^n$ is a convex cone, if there exists a strictly feasible solution $\mathbf{x}_0 \in \operatorname{int} \mathbf{K}$. In this case, there is no duality gap between $\mathbf{P}$ and its dual $\mathbf{P}^*:\ \max_{\mathbf{z}} \{\mathbf{b}'\mathbf{z} \,:\, \mathbf{c} - \mathbf{A}'\mathbf{z} \in \mathbf{K}^*\}$. In addition, if the optimal value is bounded then $\mathbf{P}^*$ has an optimal solution.

Now by [6, Lemma 1], $\mathbf{M}_d(\mathbf{y}) \succeq 0$ implies that $|y_\alpha| \leq \max[L_\mathbf{y}(1), \max_i L_\mathbf{y}(x_i^{2d})]$, for every $\alpha \in \mathbb{N}^n_{2d}$. Therefore, (2.8) has the equivalent formulation

$$(2.9) \qquad \left\{ \begin{array}{rlrl} \rho_d = -\min\limits_{\mathbf{y}} & L_\mathbf{y}(f)) & & \\ \text{s.t.} & \mathbf{M}_d(\mathbf{y}) & \succeq 0 & \\ & L_\mathbf{y}(1) & \leq 1 & \\ & L_\mathbf{y}(x_i^{2d}) & \leq 1, & i = 1, \ldots, n, \end{array} \right.$$

whose dual is eaxctly (2.5). Again Slater's condition holds for (2.9) and it has an optimal solution $\mathbf{y}^*$. Therefore (2.5) also has an optimal solution $\lambda^* \in \mathbb{R}^{n+1}_+$ with $\rho_d = \sum_i \lambda_i^*$, the desired result. $\square$

So the best $\ell_1$-norm s.o.s. approximation $g$ in Theorem (2.1) is very much the same as the $\ell_1$-approximation provided in Lasserre and Netzer [5] where all coefficients $\lambda_j^*$ were identical.

**Example 1.** Consider the Motzkin-like polynomial[2] $\mathbf{x} \mapsto f(\mathbf{x}) = x_1^2 x_2^2 (x_1^2 + x_2^2 - 1) + 1/27$ of degree 6, which is nonnegative but not a s.o.s., and with a global minimum $f^* = 0$ attained at 4 global minimizers $\mathbf{x}^* = (\pm(1/3)^{1/2}, \pm(1/3)^{1/2})$. The results are displayed in Table 1 for $d = 3, 4, 5$.

| $d$ | $\lambda^*$ | $\rho_d$ |
|---|---|---|
| 3 | $\approx 10^{-3}\,(5.445, 5.367, 5.367)$ | $\approx 1.6\,10^{-2}$ |
| 4 | $\approx 10^{-4}\,(2.4, 9.36, 9.36)$ | $\approx 2.\,10^{-3}$ |
| 5 | $\approx 10^{-5}\,(0.04, 4.34, 4.34)$ | $\approx 8.\,10^{-5}$ |

Table 1. Best $\ell_1$-approximation for the Motzkin polynomial.

## References

[1] C. Berg, J.P.R. Christensen and P. Ressel, Positive definite functions on Abelian semigroups. Math. Ann. **223**, 253–274 (1976)

[2] C. Berg, The multidimensional moment problem and semigroups. Proc. Symp. Appl. Math. **37**, 110–124 (1987).

[3] G. Blekherman, There are significantly more nonnegative polynomials than sums of squares, Isr. J. Math. **153**, 355-380 (2006)

[4] D. Henrion, J.B. Lasserre and J. Lofberg, Gloptipoly 3: moments, optimization and semidefinite programming, Optim. Methods and Software **24**, 761–779 (2009)

[5] J.B. Lasserre and T. Netzer, SOS approximations of nonnegative polynomials via simple high degree perturbations, Math. Z. **256**, 99–112 (2006)

[6] J.B. Lasserre, Sufficient conditions for a real polynomial to be a sum of squares, Arch. Math. **89**, 390–398 (2007)

[7] L. Vandenberghe and S. Boyd, Semidefinite programming, SIAM Rev. **38**, 49–95 (1996)

LAAS-CNRS and Institute of Mathematics, University of Toulouse, LAAS, 7 avenue du Colonel Roche, 31077 Toulouse Cédex 4,France

*E-mail address*: `lasserre@laas.fr`

[2] Computation was made by running the GloptiPoly software [4] dedicated to solving the generalized problem of moments.