Best $\ell_1$-approximation of nonnegative polynomials by sums of squares
Jean-Bernard Lasserre

To cite this version:
Jean-Bernard Lasserre. Best $\ell_1$-approximation of nonnegative polynomials by sums of squares. 2010. hal-00546660

HAL Id: hal-00546660
https://hal.science/hal-00546660
Preprint submitted on 14 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
BEST $\ell_1$-APPROXIMATION OF NONNEGATIVE POLYNOMIALS
BY SUMS OF SQUARES

JEAN B. LASSEURRE

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 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 and Berg. However, was essentially an existence result and subsequently, Lasserre and Netzer 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 $f \in \mathbb{R}[x]$ 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 (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 $\mathbb{R}[x]$ (resp. $\mathbb{R}[x]_d$) denote the ring of real polynomials in the variables $x = (x_1, \ldots, x_n)$ (resp. polynomials of degree at most $d$), whereas $\Sigma[x]$ (resp. $\Sigma[x]_d$) 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.
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 \( f \in \mathbb{R}[x] \), such a problem is of particular interest when \( f \) is nonnegative but not a s.o.s.
Theorem 2.1. Let $f \in \mathbb{R}[x]$ 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
\begin{equation}
(2.4) \quad x \mapsto g(x) = f(x) + (\lambda_0^* + \sum_{i=1}^n \lambda_i^* x_i^{2d}),
\end{equation}
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:
\begin{equation}
(2.5) \quad \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[\mathbb{R}^d] \right\}.
\end{equation}

Proof. Consider $f$ as an element of $\mathbb{R}[x]_{2d}$ by setting $f_\alpha = 0$ whenever $|\alpha| > \deg f$ (where $|\alpha| = \sum \alpha_i$), and rewrite (2.3) as the semidefinite program:
\begin{equation}
(2.6) \quad \min \sum_{\alpha \in \mathbb{N}_{2d}^n} \lambda_\alpha \quad \text{s.t.} \quad \lambda_\alpha + g_\alpha \geq f_\alpha, \quad \forall \alpha \in \mathbb{N}_{2d}^n
\end{equation}
\begin{equation}
\lambda_\alpha - g_\alpha \geq -f_\alpha, \quad \forall \alpha \in \mathbb{N}_{2d}^n
\end{equation}
\begin{equation}
g_\alpha \leq \langle X, B_\alpha \rangle = 0, \quad \forall \alpha \in \mathbb{N}_{2d}^n.
\end{equation}
The dual semidefinite program of (2.6) reads:
\begin{equation}
(2.7) \quad \max_{u_\alpha, v_\alpha \geq 0, y} \sum_{\alpha \in \mathbb{N}_{2d}^n} f_\alpha(u_\alpha - v_\alpha) \quad \text{s.t.} \quad \begin{cases}

u_\alpha + v_\alpha & \leq 1, \quad \forall \alpha \in \mathbb{N}_{2d}^n \\
u_\alpha - v_\alpha + y_\alpha & = 0, \quad \forall \alpha \in \mathbb{N}_{2d}^n \\
M_d(y) & \succeq 0,
\end{cases}
\end{equation}
or, equivalently,
\begin{equation}
(2.8) \quad \max_{y} \quad -L_y(f) \quad \text{s.t.} \quad M_d(y) \succeq 0, \quad |y_\alpha| \leq 1, \quad \forall \alpha \in \mathbb{N}_{2d}^n.
\end{equation}
The semidefinite program (2.8) has an optimal solution $y^*$ because the feasible set is compact. In addition, let $y = (y_\alpha)$ be the moment sequence of the measure $d\mu = e^{-\|x\|^2}dx$, scaled so that $|y_\alpha| < 1$ for all $\alpha \in \mathbb{N}_{2d}^n$. Then $(y, u, v)$ with $u = -\min[0, y]$ and $v = \max[0, y]$, is strictly feasible in (2.7) because $M_d(y) > 0$, and so Slater’s condition holds for (2.3). Therefore, by a standard duality result in convex optimization, there is no duality gap between (2.4) and (2.7) (or (2.8)), and (2.4) has an optimal solution $(\lambda^*, X^*, y^*)$. Hence $\rho_d = -L_{y^*}(f)$ for any optimal solution $y^*$ of (2.8).
Now by \([6, \text{Lemma 1}]\), \(M_d(y) \succeq 0\) implies that \(|y_\alpha| \leq \max[L_y(1), \max_i L_y(x_i^{2d})]\), for every \(\alpha \in \mathbb{N}^n_{2d}\). Therefore, \((2.8)\) has the equivalent formulation

\[
\begin{cases}
\rho_d = -\min_y L_y(f) \\
\text{s.t. } M_d(y) \succeq 0 \\
L_y(1) \leq 1 \\
L_y(x_i^{2d}) \leq 1, \quad i = 1, \ldots, n,
\end{cases}
\]

(2.9)

whose dual is exactly \((2.3)\). Again Slater’s condition holds for \((2.9)\) and it has an optimal solution \(y^*\). Therefore \((2.5)\) also has an optimal solution \(\lambda^* \in \mathbb{R}_{n+1}^n\) 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 \(x \mapsto f(x) = x_1^2x_2^2(x_1^2 + x_2^2 - 1) + 1/27\) of degree 6, which is nonnegative but not a s.o.s., and with a global minimum \(f^* = 0\) attained at 4 global minimizers \(x^* = (\pm 1/3)^{1/2}, (\pm 1/3)^{1/2}\). The results are displayed in Table 1 for \(d = 3, 4, 5\).

| \(d\) | \(\lambda^*\) | \(\rho_d\) |
|------|-------------|----------|
| 3    | \(\approx 10^{-3}(5.445, 5.367, 5.367)\) | \(\approx 1.6 \times 10^{-2}\) |
| 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.