Set approximation via minimum-volume polynomial sublevel sets
Fabrizio Dabbene, Didier Henrion

To cite this version:
Fabrizio Dabbene, Didier Henrion. Set approximation via minimum-volume polynomial sublevel sets. European Control Conference (ECC), Jul 2013, Zurich, Switzerland. 11p., 2013. <hal-00740794>

HAL Id: hal-00740794
https://hal.archives-ouvertes.fr/hal-00740794
Submitted on 11 Oct 2012

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.
Minimum volume semialgebraic sets for robust estimation

Fabrizio Dabbene¹, Didier Henrion²,³,⁴

October 11, 2012

Abstract

Motivated by problems of uncertainty propagation and robust estimation we are interested in computing a polynomial sublevel set of fixed degree and minimum volume that contains a given semialgebraic set \( K \). At this level of generality this problem is not tractable, even though it becomes convex e.g. when restricted to nonnegative homogeneous polynomials. Our contribution is to describe and justify a tractable \( L^1 \)-norm or trace heuristic for this problem, relying upon hierarchies of linear matrix inequality (LMI) relaxations when \( K \) is semialgebraic, and simplifying to linear programming (LP) when \( K \) is a collection of samples, a discrete union of points.

1 Introduction

In this paper, we consider the problem of computing reliable approximations of a given set \( K \subset \mathbb{R}^n \). The set \( K \) is assumed to have a complicated shape (e.g. nonconvex, non-connected), expressed in terms of semialgebraic conditions, and we seek for approximations which should i) be easy computable and ii) have a simple description.

The problem of deriving reliable approximations of overly complicated sets by means of simpler geometrical shapes has a long history, and it arises in many research fields related to optimization, system identification and control.

In particular, outer bounding sets, i.e. approximations that are guaranteed to contain the set \( K \), are widespread in the technical literature, and they find several applications in robust control and filtering. For instance, set-theoretic state estimators for uncertain discrete-time nonlinear dynamic systems have been proposed in [1, 7, 9, 23]. These strategies adopt a set-membership approach [8, 22], and construct compact sets that are guaranteed to bound the systems states which are consistent with the measured output and the norm-bounded uncertainty. Outer approximation also arise in the context of

¹CNR-IEIT; Corso Duca degli Abruzzi 123, Torino; Italy. fabrizio.dabbene@polito.it
²CNRS, LAAS, 7 avenue du colonel Roche, F-31400 Toulouse; France. henrion@laas.fr
³Université de Toulouse, LAAS, F-31400 Toulouse; France
⁴Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, CZ-16626 Prague, Czech Republic
robust fault detection problems (e.g., see [14]) and of reachability analysis of nonlinear and/or hybrid systems [13, 16].

Similarly, inner approximations are employed in nonlinear programming [21], in the solution of design centering problems [24] and for fixed-order controller design [12].

Recently, the authors of [5] have proposed an approach based on randomization, which constructs convex approximations of generic nonconvex sets which are neither inner nor outer, but they enjoy some specific probabilistic properties. In this context, an approximation is considered to be reliable if it contains “most” of the points in $K$ with prescribed high probability. The key tool in this framework is the generation of random samples inside $K$, and the construction of a convex set containing these samples.

In all the approaches listed above, several geometric figures have been adopted as approximating sets. The application of ellipsoidal sets to the state estimation problem has been introduced in the pioneering work [22] and used by many different authors from then on; see, for example, [7, 9]. The use of polyhedrons was proposed in [15] to obtain an increased estimation accuracy, while zonotopes have been also recently studied in [1, 10].

More recent works, like for instance [4, 12, 19], employ sets defined by semialgebraic conditions. In particular, in [19] the authors use polynomial sum-of-squares (SOS) programming, a particular class of SDP, to address the problem of fitting given data with a convex polynomial, seen as a natural extension of quadratic polynomials and ellipsoids. Convexity of the polynomial is ensured by enforcing that its Hessian is matrix SOS, and volume minimisation is indirectly enforced by increasing the curvature of the polynomial. In [4] the authors propose moment relaxations for the separation and covering problems with semialgebraic sets, thereby also extending the classical ellipsoidal sets used in data fitting problems.

Our contribution is to extend further these works to cope with volume minimization of arbitrary (e.g. non-convex, non-connected, higher degree) semialgebraic sets containing a given semialgebraic set (e.g. a union of points). Since there is no analytic formula for the volume of a semialgebraic set, in terms of the coefficients of the polynomials defining the set, we have no hope to solve this optimization problem globally. Instead, we describe and justify analytically and geometrically a computationally tractable heuristic based on $L^1$-norm or trace minimization.

2 Problem statement

Given a compact basic semialgebraic set

$$K := \{ x \in \mathbb{R}^n : g_i(x) \geq 0, \ i = 1, 2, \ldots, m \}$$

where $g_i(x)$ are real multivariate polynomials, we want to compute a polynomial sublevel set

$$V(q) := \{ x \in \mathbb{R}^n : q(x) \leq 1 \} \supset K$$

of minimum volume that contains $K$. Set $V(q)$ is modeled by a polynomial $q$ belonging to $P_d$, the vector space of multivariate real polynomials of degree less than or equal to $d$. 

2
In other words, we want to solve the following optimization problem:

$$\inf_{q \in \mathcal{P}_d} \text{vol } V(q) \quad \text{s.t.} \quad K \subset V(q). \quad (1)$$

In the above problem

$$\text{vol } V(q) := \int_{V(q)} dx = \int_{\mathbb{R}^n} I_{V(q)}(x) dx$$

is the volume or Lebesgue measure of set $V(q)$ and $I_X(x)$ is the indicator function, equal to one when $x \in X$ and zero otherwise.

Note that, typically, the set $K$ has a complicated description in terms of polynomials $g_i$ (e.g. coming from physical measurements and/or estimations) and set $V(q)$ has a simple description (in the sense that the degree of $q$ is small, say less than 10). Minimization of the volume of $V(q)$ means that we want $V(q)$ to capture most of the geometric features of $K$.

If $K$ is convex and $q$ is quadratic, then the infimum of problem (1) is attained, and there is a unique (convex) ellipsoid $V(q)$ of minimum volume that contains $K$, called Löwner-John ellipsoid. It can be computed by convex optimization, see e.g. [3, §4.9].

In general, without convexity assumptions on $K$, a solution to problem (1) is not unique. There is also no guarantee that the computed set $V(q)$ is convex. Optimization problem (1) is nonlinear and semi-infinite, in the sense that we optimize over the finite-dimensional vector space $\mathcal{P}_d$ but subject to an infinite number of constraints (to cope with set inclusions).

If we denote by $\pi_d(x)$ a (column vector) basis of monomials of degree up to $d$, we can write $q(x) = \pi_d^T(x)q$ where $q$ is a vector of coefficients of given size. In vector space $\mathcal{P}_d$, the set $Q := \{q : \pi_d^T(x)q \leq 1\}$ is by definition the polar of the bounded set $\{\pi_d(x) : \|x\| \leq R\}$ where $R > 0$ is a constant chosen sufficiently large so that all vectors $x \in V(q)$ have norm less than $R$. As the polar of a compact set whose interior contains the origin, set $Q$ is compact, see e.g. [18]. It follows that the feasible set of problem (1) is compact, and since the objective function is continuous, the infimum in problem (1) is attained.

3 Convex conic formulation

In this section it is assumed that $q$ is a nonnegative homogeneous polynomial, or form, of even degree $d = 2\delta$ in $n$ variables. Under this restriction, in [17, Lemma 2.4] it is proved that the volume function

$$q \mapsto \text{vol } V(q)$$

is convex in $q$. The proof of this statement relies on the striking observation [20] that

$$\text{vol } V(q) = C_d \int_{\mathbb{R}^n} e^{-q(x)} dx$$

where $C_d$ is a constant depending only on $d$. Note also that boundedness of $V(q)$ implies that $q$ is nonnegative, since if there is a point $x_0 \in \mathbb{R}^n$ such that $q(x_0) < 0$, and hence
$x_0 \in V(q)$, then by homogeneity of $q$ it follows that $q(\lambda x_0) = \lambda^{2d} q(x_0) < 0$ for all $\lambda$ and hence $\lambda x_0 \in V(q)$ for all $\lambda$ which contradicts boundedness of $V(q)$. This implies that problem (1), once restricted to nonnegative forms, is a convex optimization problem.

In [17, Lemma 2.4] explicit expressions are given for the first and second order derivatives of the volume function, in terms of the moments

$$\int_{\mathbb{R}^n} x^\alpha e^{-q(x)} dx$$

(2)

for $\alpha \in \mathbb{N}^n$, $|\alpha| \leq 2d$. In an iterative algorithm solving convex problem (1), one should then be able to compute repeatedly and quickly integrals of this kind, arguably a difficult task.

When $q$ is not homogeneous, we do not know under which conditions on $q$ the volume function $V(q)$ is convex in $q$.

Motivated by these considerations, in the remainder of the paper we propose a simpler approach to solving problem (1) which is not restricted to forms, and which does not require the potentially intricate numerical computation of moments (2). Our approach is however only a heuristic, in the sense that we do not provide guarantees of solving problem (1) globally.

4 \textit{L}^1\text{-norm minimization}

Let us write $V(q)$ as a polynomial superlevel set

$$U(p) := V(q) = \{x \in \mathbb{R}^n : p(x) := 2 - q(x) \geq 1\}.$$

with polynomial

$$p(x) = \pi_T^T(x) P \pi(x)$$

(3)

expressed as a quadratic form in a given (column vector) basis $\pi(x)$ of monomials of degree up to $\delta := \lceil \frac{d}{2} \rceil$, with symmetric Gram matrix $P$. Then, optimization problem (1) reads

$$v^{\ast}_d := \min_{p \in \mathcal{P}_d} \text{vol} U(p) \quad \text{s.t.} \quad K \subset U(p).$$

(4)

Note that in problem (4) we can indifferently optimize over coefficients of $p$ or coefficients of matrix $P$, since they are related linearly via constraint (3).

Since $K$ is compact by assumption and $U(p)$ is compact for problem (4) to have a finite minimum, we suppose that we are given a compact semialgebraic set

$$B := \{x \in \mathbb{R}^n : b_i(x) \geq 0, \ i = 1, 2, \ldots, m_b\}$$

such that $U(p) \subset B$ and hence

$$U(p) = \{x \in B : p(x) \geq 1\}.$$

The particular choice of polynomials $b_i$ will be specified later on. Now, observe that by definition

$$p(x) \geq I_U(p)(x) \ on \ \mathbb{R}^n.$$
and hence, integrating both sides we get
\[ \int_B p(x)dx \geq \int_B I_{U(p)}(x)dx = \text{vol } U(p), \]
an inequality known as Chebyshev’s inequality, widely used in probability, see e.g. [2, §2.4.9]. If polynomial \( p \) is nonnegative on \( B \) then the above left-hand side is the \( L^1 \)-norm of \( p \), and the inequality becomes
\[ \|p\|_1 \geq \text{vol } U(p). \] (5)

Now consider the following \( L^1 \)-norm minimization problem:
\[ w^*_d := \min_{p \in P_d} \|p\|_1 \text{ s.t. } \begin{cases} p \geq 0 \text{ on } B \\ p \geq 1 \text{ on } K \end{cases}. \] (6)

Lemma 1 The minimum of problem (6) monotonically converges from above to the minimum of problem (4), i.e. \( w^*_{d-1} \geq w^*_d \geq v^*_d \) for all \( d \), and \( \lim_{d \to \infty} w^*_d = \lim_{d \to \infty} v^*_d \).

Proof: The graph of polynomial \( p \) lies above \( I_K \), the indicator function of set \( K \), while being nonnegative on \( B \), so minimizing the \( L^1 \)-norm of \( p \) on \( B \) yields an upper approximation of \( I_K \). Monotonicity of the sequence \( w^*_d \) follows immediately since polynomials of degree \( d + 1 \) include polynomials of degree \( d \). When its degree increases, \( p \) converges in \( L^1 \)-norm to \( I_K \), hence \( \|p\|_1 \) converges to \( \text{vol } K \). The convergence is pointwise almost everywhere, and almost uniform, but not uniform since \( I_K \) is discontinuous on \( B \). \( \square \)

Note that this \( L^1 \)-norm minimization approach was originally proposed in [11] to compute numerically the volume and moments of a semialgebraic set.

5 Trace minimization

In this section we give a geometric interpretation of problem (6). First note that the objective function reads
\[ \|p\|_1 = \int_B p(x)dx = \int_B \pi^T_\delta(x)P\pi_\delta(x)dx = \text{trace } \left( P \int_B \pi_\delta(x)\pi^T_\delta(x)dx \right) = \text{trace } PM \]
where
\[ M := \int_B \pi_\delta(x)\pi^T_\delta(x)dx \]
is the matrix of moments of the Lebesgue measure on \( B \) in basis \( \pi_\delta(x) \). In equation (3) if the basis is chosen such that its entries are orthonormal w.r.t. the (scalar product induced by the) Lebesgue measure on \( B \), then \( M \) is the identity matrix and inequality (5) becomes
\[ \text{trace } P \geq \text{vol } U(p) \] (7)
which indicates that, under the above constraints, minimizing the trace of the Gram matrix $P$ entails minimizing the volume of $U(p)$.

The choice of polynomials $b_i$ in the definition of the bounding set $B$ should be such that the objective function in problem (6) is easy to compute. If

$$p(x) = \pi_d^T(x)p = \sum_{\alpha} p_\alpha \pi_d(x)_{\alpha}$$

then

$$\int_B p(x) dx = \sum_{\alpha} p_\alpha \int_B [\pi_d(x)]_{\alpha} dx = \sum_{\alpha} p_\alpha y_\alpha$$

and we should be able to compute easily the moments

$$y_\alpha := \int_B [\pi_d(x)]_{\alpha} dx$$

of the Lebesgue measure on $B$ w.r.t. basis $\pi_d(x)$. This is the case e.g. if $B$ is a box.

**Remark 1 (Minimum trace heuristic for ellipsoids)** Note that, in the case of quadratic polynomials, i.e. $d = 2$, we retrieve the classical trace heuristic used for volume minimization, see e.g. [6]. If $B = [-1,1]^n$ then the basis $\pi_1(x) = \frac{\sqrt{d}}{2} x$ is orthonormal w.r.t. the Lebesgue measure on $B$ and $\|p\|_1 = \frac{3}{2} \text{trace} P$. The constraint that $p$ is nonnegative on $B$ implies that the curvature of the boundary of $U(p)$ is nonnegative, hence that $U(p)$ is convex.

In [19], the authors restricted the search to convex polynomial sublevels $U(p) = V(q) = \{x : q(x) \leq 1\}$ by enforcing positive semidefiniteness of the Gram matrix of the quadratic Hessian form of $q$. They proposed to maximize (the logarithm of) the determinant of the Gram matrix, justifying this choice by explaining that this increases the curvature of the polynomial sublevel set along all directions, and hence minimizes the volume. Supported by Lemma 1 and the above discussion, we came to the consistent conclusion that minimizing the trace of the Gram matrix of $p$, that is, maximizing the trace of the Gram matrix of $q$, is a relevant heuristic for volume minimization. Note however that in our approach we do not enforce convexity of $U(p)$.

### 6 Handling constraints

If $K = \{x : g_i(x) \geq 0, i = 1,2,\ldots,m\}$ is a general semialgebraic set then we must ensure that polynomial $p - 1$ is nonnegative on $K$, and for this we can use Putinar’s Positivstellensatz and a hierarchy of finite-dimensional convex LMI relaxations which are linear in the coefficients of $p$. More specifically, we write $p - 1 = r_0 + \sum_i r_i g_i$ where $r_0, r_1, \ldots, r_m$ are polynomial sum-of-squares of given degree, to be found. For each fixed degree, the problem of finding such polynomials is an LMI, see e.g. [17, Section 3.2]. The constraint that $p$ is nonnegative on $B$ can be ensured similarly.

A particularly interesting case is when $K$ is a discrete set, i.e. a union of points $x_i \in \mathbb{R}^n, i = 1,\ldots,N$. Indeed, in this case the inclusion constraint $K \subset U(p)$ is equivalent to a
finite number of inequalities \( p(x_i) \geq 1, \ i = 1, \ldots, N \) which are linear in the coefficients of \( p \). Similarly, the constraint that \( p \) is nonnegative on \( B \) can be handled by linear inequalities \( p(x_j) \geq 0 \) enforced at a dense grid of points \( x_j \in B, \ j = 1, \ldots, M \), for \( M \) sufficiently large. Note that, with this pure linear programming (LP) approach, it is not guaranteed that \( p \) is nonnegative on \( B \), but what matters primarily is that \( K \subset U(p) \), which is indeed guaranteed. This purely LP formulation allows to deal with problems with rather large \( N \).

7 Examples

We illustrate the proposed approach for the case when \( K \) is a discrete set, a union of points of \( \mathbb{R}^n \). For our numerical examples, we have used the YALMIP interface for Matlab to model the LMI optimization problem (1), and the SDP solver SeDuMi to solve numerically the problem. Since the degrees of the semialgebraic sets we compute are typically low (say less than 20), we did not attempt to use appropriate polynomial bases (e.g. Chebyshev polynomials) to improve the quality and resolution of the optimization problems, see however [11] for a discussion on these numerical matters.

7.1 Line \((n = 1)\)

To illustrate the behavior of the proposed optimization procedure, we first consider \( B = [-1, 1] \) and \( K = \{-\frac{1}{2}, 0, \frac{1}{2} \} \). On Figure 1 we represent the solutions \( p \) of degrees 2, 7, 17 and 26 of minimization problem (6). We observe that the superlevel set \( U(p) = \{ x \in B : p(x) \geq 1 \} \) is simply connected for degree 2, doubly connected for degree 7, and triply connected for degrees 17 and 26. Note that, as the degree increases of \( p \), the length of the intervals for which \( p(x) > 1 \) tends to zero. This is consistent with the fact that the volume of a finite set is zero. We also observe that polynomial \( p \) shows increasing oscillations for increasing degrees, a typical feature of such discontinuous function approximations.

7.2 Plane \((n = 2)\)

In \( B = [-1, 1]^2 \) we consider two clouds of 50 points each, i.e. \( N = 100 \). On Figure 2 we represent the solutions \( p \) of degrees 2, 5, 9 and 14 of minimization problem (6). Here too we observe that increasing the degree of \( p \) allows to disconnect set \( U(p) \). The side effects near the border of \( B \) on the top right figure can be removed by enlarging the bounding set \( B \).

7.3 Space \((n = 3)\)

In \( B = [-1, 1]^3 \) we consider \( N = 10 \) points. The solutions \( p \) of degrees 4, 10, and 14 of minimization problem (6) is depicted in Figure 3. Here too we observe that increasing the degree of \( p \) allows to capture point clusters in distinct connected components.

It should be pointed out that all the semialgebraic sets in the previous examples were computed in a few seconds of CPU time on a standard PC.
8 Conclusion

In this paper, we proposed a simple technique for approximating a given set of complicated shape by means of the polynomial sub level set of minimal size that contains it. The proposed approximation has been shown to be convex in the case the polynomial is assumed to be homogeneous. Then, a tractable relaxation based on Chebychev’s inequality has been introduced. Interestingly, this relaxation reduces to the classical trace minimization heuristic in the case of quadratic polynomials, thus indirectly providing an intuitive explanation to this widely used criterion.

Ongoing research is devoted to utilize the proposed approximation to construct new classes of set-theoretic filters, in the spirit of the works [1, 9].

Acknowledgments

This work was partially supported by a Bilateral Project CNRS-CNR.
Figure 2: Minimum $L^1$-norm polynomials $p$ (red) and low surface semialgebraic sets (yellow) such that $p \geq 0$ on $[-1, 1]^2$ and $p \geq 1$ at 100 points (black), for degree 2 (upper left), 5 (upper right), 9 (lower left) and 14 (lower right).

Figure 3: Including the same 10 space points (black) in low volume semialgebraic sets (yellow) of degree 4 (left), 10 (center), 14 (right).

References


