Global Optimization of Polynomial Functions and Applications

by

Jiawang Nie

B.S. (Xi'an Jiaotong University) 1997 M.S. (Chinese Academy of Sciences) 2000

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy

 in

Applied Mathematics

in the

GRADUATE DIVISION of the UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge: Professor James Demmel, Co-Chair Professor Bernd Sturmfels, Co-Chair Professor Laurent El Ghaoui Professor Ming Gu

Fall 2006

The dissertation of Jiawang Nie is approved:

Co-Chair	Date
Co-Chair	Date
	Date
	Date

University of California, Berkeley

Fall 2006

Global Optimization of Polynomial Functions and Applications

Copyright 2006 by

Jiawang Nie

Abstract

Global Optimization of Polynomial Functions and Applications

by

Jiawang Nie Doctor of Philosophy in Applied Mathematics University of California, Berkeley Professor James Demmel, Co-Chair Professor Bernd Sturmfels, Co-Chair

This thesis discusses the global optimization problem whose objective function and constraints are all described by (multivariate) polynomials. The motivation is to find the global solution. For this problem, sum of squares (SOS) relaxations are able to get guaranteed lower bounds.

For unconstrained polynomial optimization problem, SOS relaxation generally only provides a lower bound. Sometimes this lower bound may be strictly smaller than the global minimum. In such situations, how can we do better? Much better lower bounds can be obtained if we apply SOS relaxation over the gradient ideal of the polynomial function. In fact, we can always get the exact lower bound, and have finite convergence, under some conditions that hold generically.

For constrained polynomial optimization, when the feasible set is compact, Lasserre's procedure is usually applied to to get a sequence of lower bounds. Under a certain condition, these lower bounds will converge to the global minimum. However, no estimates of the speed of the convergence were available. For this purpose, we obtain the first upper bound on the convergence rate. When the feasible set is *not* compact, *Lasserre's procedure* may not converge. In such situations, better lower bounds can be obtained if we apply SOS relaxation over the Kuhn-Karush-Tucker (KKT) ideal. This new sequence of lower bounds has finite convergence under some generic conditions.

SOS relaxations can also be applied to minimize rational functions. The new features of SOS relaxations for this problem are studied in this thesis.

Polynomial optimization has wide applications. We studied the applications in shape optimization of transfer functions, finding minimum ellipsoid bounds for polynomial systems, solving the nearest GCD problem, maximum likelihood optimization, and sensor network localization.

> James Demmel Dissertation Committee Co-Chair

> Bernd Sturmfels Dissertation Committee Co-Chair

Contents

Li	List of Figures ii			
\mathbf{Li}	st of	Tables	iv	
1	\mathbf{Intr}	oduction	1	
	1.1	Prior work	4	
		1.1.1 Sum of squares (SOS) relaxations	4	
		1.1.2 SOS methods versus symbolic methods	9	
	1.2	Contributions of this thesis	10	
2	Sor	ne Basic Tools for Polynomial Optimization	12	
	2.1	SOS and nonnegative polynomials	12	
	2.2	Elementary algebraic geometry	15	
	2.3	Positive polynomials on semialgebraic sets	16	
3	On	the Convergence Rate of Lasserre's Procedure	21	
	3.1	Convergence rate of Lasserre's procedure	22	
	3.2	On the complexity of Putinar's positivstellensatz	26	
	3.3	The proof of Theorem 3.2.4	31	
4	so	S Methods Based on the Gradient Ideal	42	
	4.1	Introduction	43	
	4.2	Polynomials over their gradient varieties	45	
	4.3	Convergence analysis and the algorithm	50	
		4.3.1 What if the gradient ideal $\mathcal{I}_{grad}(f)$ is not radical?	55	
	4.4	Numerical experiments	56	
		4.4.1 Comparison of lower bounds	57	
		4.4.2 Comparison of computational efficiency	58	

5	SOS Methods based on the Kuhn-Karush-Tucker (KKT) Ideal		62	
	5.1	Introduction	63	
	5.2	Representations in \mathcal{P}_{KKT} and \mathcal{M}_{KKT}	67	
	5.3	Convergence of the lower bounds	72	
	5.4	Structures over some special constraints	75	
		5.4.1 Nonnegative orthant \mathbb{R}^n_+	75	
		5.4.2 Box $[a, b]_n$	79	
6	Mir	nimizing Rational Functions	83	
	6.1	SOS relaxation for unconstrained minimization	84	
	6.2	What if $r_{sos}^* < r^*$?	91	
	6.3	Constrained minimization	97	
7	Apj	plications of Polynomial Optimization	106	
	7.1	Shape optimization of transfer functions	106	
	7.2	Minimum ellipsoid bounds for polynomial systems	113	
	7.3	Nearest greatest common divisor	121	
	7.4	Maximum likelihood optimization	126	
	7.5	Sensor network localization	128	
Bi	bliog	graphy	133	

List of Figures

1.1	Contour of $f(x)$
1.2	An example of Lasserre's procedure
7.1	A bandpass filter
7.2	A 3-step constant filter
7.3	The ellipsoid for polynomial system $(7.2.11)$ - $(7.2.13)$
7.4	Ellipsoid bound for each component $\hfill \ldots \hfill \ldots \hfill \ldots \hfill 121$
7.5	Ellipsoid bound for polynomial system $(7.2.11)$ - $(7.2.15)$
7.6	500 sensors, sparse SOS relaxation

List of Tables

- The relative difference $\frac{|f_{N,grad}^* f_{sos}^*|}{|f_{sos}^*|} \times 10^{10}$, with N = d/2. Running time in CPU seconds via traditional SOS approach (4.1.2) . 4.160
- 4.260
- 4.3Running time in CPU seconds via our approach (4.1.5), with N = d/2. 60
- The ratio of CPU seconds between (4.1.2) and (4.1.5), with N = d/2. 604.4

Acknowledgments

Firstly, I would express my greatest appreciation to my co-advisors, James Demmel and Bernd Sturmfels for their helpfulness and encouragement during my five years of Ph.D studies. Their fantastic supervision and unconditional support make my graduate study unique. Our collaboration is joyful and instructive. During these five years in Berkeley, their wonderful advising and encouragement helped me go through many difficult times in completing this thesis. I will be forever thankful for this special opportunity.

I am also grateful to other members of my thesis committee: Laurent El Ghaoui and Ming Gu. Their professional advice and expertise are encouraging and helpful. It is a joy to have discussions with them. I can never forget that, in one of Laurent's classes (*Convex Optimization*) in Fall 2001, I had the chance to learn about the new field of polynomial optimization, which become the theme of this thesis.

The academic environment of the Berkeley Mathematics Department is unique and provides much more than what I had expected. The department provided fertile ground and an unsurpassable research environment for the completion of this thesis. Its fantastic faculty and wonderful staff gave warm support to my graduate life.

I would like to take this opportunity to thank my colleagues Vicki Powers and Markus Schweighofer for their wonderful collaborations. Their professional expertise and wonderful communication have helped me learn much in real algebraic geometry, especially about positive polynomials.

Finally, I would like to give my deepest gratitude to my family for their endless support and love.

Chapter 1

Introduction

Many problems in engineering and in science can be formulated as finding the optimal value of some objective function subject to some constraints on the decision variables. Finding the optimal decision variables is one main theme of the discipline of *Mathematical Programming*. There is a huge volume of work on the theory, algorithms and applications of Mathematical Programming.

When the objective and constraints are convex, the problem is called *convex* optimization. In this case, every local minimizer is also a global minimizer, the set of minimizers is convex, and specialized, very efficient algorithms are available. We refer to [15, 68, 94, 96] for the theory and methods for convex optimization.

When the objective and constraints are general nonlinear functions (often nonconvex), the optimization problem is called *nonlinear programming*. In such situations, a local minimizer might not be a global minimizer, and the set of minimizers may be nonconvex or even disconnected. There is much work on developing numerical methods to find local minimizers. We refer to [3, 7, 5, 58, 65, 75] for the theory

and methods for nonlinear programming.

Another important and active area of mathematical programming is *global* optimization — the theory and methods for finding global minimizers. Many global methods are based on branch-and-bound algorithms. They are often very expensive to implement. For computational efficiency, *heuristic methods* are developed. But the solutions returned by heuristic methods might not be globally optimal (or only globally optimal under some assumptions). We refer to [76, 77, 78] for the theory and methods of global optimization.

Usually, it is very difficult to find the global minimizer(s) of a general nonconvex nonlinear function. However, if the objective and constraints are described by multivariate polynomials (not necessarily convex), there are some certificates on the emptiness of polynomial systems from *real algebra* [13, 6, 27]. Therefore certificates of global solutions for polynomials can be computed and tractable algorithms can be developed.

This thesis concentrates on the special optimization problems whose objective and constraints are all polynomials. The main problem to be considered is of the form

$$f^* = \min_{x \in \mathbb{R}^n} \qquad f(x) \tag{1.0.1}$$

s.t.
$$g_1(x) \ge 0, \cdots, g_m(x) \ge 0$$
 (1.0.2)

where $f(x), g_i(x) \in \mathbb{R}[X]$, the ring of real multivariate polynomials in $X = (x_1, \dots, x_n)$. Denote by S be the feasible set defined by constraint (1.0.2). It is a *basic closed semialgebraic set* [13]. In this thesis, we do not have any convexity/concavity assumptions on f(x) or $g_i(x)$. The goal is to find the global minimum f^* and global minimizers (if any). When (1.0.2) defines the whole space \mathbb{R}^n , the problem (1.0.1)-(1.0.2) becomes an unconstrained problem.

The formulation (1.0.1)-(1.0.2) contains quite a broad class of optimization problems, including some NP-hard problems.

• (Linear Programming (LP)) When f(x) and all $g_i(x)$ are all affine functions, (1.0.1)-(1.0.2) becomes a linear programming of the form:

$$f^* = \min_{x \in \mathbb{R}^n} \quad c^T x$$

s.t. $a_i^T x + b_i \ge 0, \ i = 1, \cdots, m$

where c and a_i are all vectors in \mathbb{R}^n .

• (Nonconvex Quadratic Programming (QP)) When f(x) and all $g_i(x)$ are all quadratic functions (not necessarily convex or concave), (1.0.1)-(1.0.2) becomes a QP:

$$f^* = \min_{x \in \mathbb{R}^n} \quad \frac{1}{2} x^T A_0 x + a_0^T x$$

s.t. $\frac{1}{2} x^T A_i x + a_i^T x + c_i \ge 0, \ i = 1, \cdots, m$

where $A_i \in S^n$ are symmetric matrices and $a_i \in \mathbb{R}^n$ are vectors. To solve the nonconvex QP is NP-hard [64].

Many combinatorial problems can be formulated as polynomial optimization problems. Here are some examples.

• (Matrix Copositivity) A symmetric matrix $A \in S^n$ is *copositive* if the quartic form $\sum_{i,j} A_{ij} x_i^2 x_j^2$ is always nonnegative. This can be decided by finding the global minimum of $\sum_{i,j} A_{ij} x_i^2 x_j^2$ on the unit ball, which is a polynomial optimization problem. Testing the copositivity of a matrix is NP-hard [34].

 (Partition Problem) Given a vector a ∈ Zⁿ, can it be partitioned into two parts with equal sums? This can be formulated as the polynomial optimization problem

$$\min_{x \in \mathbb{R}^n} \quad (a^T x)^2 + \sum_i (x_i^2 - 1)^2.$$

The global minimum is zero if and only if the vector a can be partitioned into two parts with equal sums.

• (Maxcut problem) Given a graph G = (V, E) with edge weights W_{ij} , how do we partition G into two parts such that the edges connecting these two parts have maximum sum of weights? This can be formulated as

$$\max_{x \in \mathbb{R}^n} \sum_{i,j} W_{ij} (x_i - x_j)^2$$

s.t. $x_i (x_i - 1) = 0, i = 1, \cdots, n$

1.1 Prior work

There has been a great deal of recent work in using Sum of Squares (SOS) relaxations to find global solutions to polynomial optimization problems. Here we give a very brief review of SOS methods.

1.1.1 Sum of squares (SOS) relaxations

The basic idea of relaxation is to approximate nonnegative polynomials by Sum of Squares (SOS) polynomials, i.e., those polynomials that can be written as a summation of squares of other polynomials (e.g., $x_1^2 - x_1x_2 + x_2^2 = \frac{1}{2}x_1^2 + \frac{1}{2}(x_1 - x_2)^2 + \frac{1}{2}x_2^2$). See Section 2.1 for an introduction to SOS polynomials.

To see the application of SOS relaxation in optimization, let us first consider problem (1.0.1) without any constraints. Assume the degree of f(x) is even (otherwise f(x) is unbounded from below). Obviously, the minimum f^* equals the maximum lower bound of f(x), i.e.,

$$f^* = \max \quad \gamma \tag{1.1.3}$$

s.t.
$$f(x) - \gamma \ge 0 \quad \forall x \in \mathbb{R}^n.$$
 (1.1.4)

When $\deg(f) \ge 4$, it is NP-hard [67] to find f^* and the minimizing values of the argument (if any). So in practice, one is interested in finding a lower bound of f^* and extracting some approximate solutions. SOS relaxation is such a method, and it provides exact lower bounds in many cases. If we relax the nonnegativity condition (1.1.4) to an SOS condition, we get the convex optimization problem:

$$f_{sos}^* = \max_{\gamma} \quad \gamma \tag{1.1.5}$$

s.t.
$$f(x) - \gamma \in \sum \mathbb{R}[X]^2$$
. (1.1.6)

Here $\sum \mathbb{R}[X]^2$ denotes the set of all polynomials that can be represented as sums of squares of polynomials. Notice that the decision variable above is γ instead of $x \in \mathbb{R}^n$. The attractive property of (1.1.5)-(1.1.6) is that it can be transformed to Semidefinite Programming (SDP) problem, for which efficient algorithms exist (e.g., interior-point methods). Notice that $f(x) - \gamma$ being SOS implies that $f(x) \geq \gamma$ for any $x \in \mathbb{R}^n$. Thus f_{sos}^* is a lower bound for f(x), that is, $f_{sos}^* \leq f^*$. And $f_{sos}^* = f^*$ if and only if the polynomial $f(x) - f^*$ is SOS [52]. From Theorem 2.1.3 below, we know in many occasions $f(x) - f^*$ may be nonnegative but *not* SOS. Thus the lower bound may not be exact.

Let $\ell = \lceil deg(f)/2 \rceil$ and write $f(x) = \sum_{\alpha} f_{\alpha} x^{\alpha}$, where the indices $\alpha = (\alpha_1, \dots, \alpha_n)$ are in \mathbb{N}^n and $x^{\alpha} := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. Then the dual of problem (1.1.5)-(1.1.6) is

$$\min_{y} \quad \sum_{\alpha} f_{\alpha} y_{\alpha} \tag{1.1.7}$$

s.t.
$$M_{\ell}(y) \succeq 0, \ y_{(0,\dots,0)} = 1.$$
 (1.1.8)

Here $y = (y_{\alpha})$ is a monomial-indexed vector, i.e., indexed by integer vectors in \mathbb{N}^n , and $M_{\ell}(y)$ is the moment matrix generated by vector y (see Definition 2.3.6 in Section 2.3). Here $A \succeq 0$ means the symmetric matrix A is positive semidefinite.

For example, consider minimizing the polynomial

$$f(x) = 4x_1^2 - \frac{21}{10}x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4.$$

The contour of f(x) is shown in Figure 1.1. We can see that f(x) is highly nonconvex and has several local minimizers. Now we apply SOS relaxation to find its global minimum and minimizers. SOS relaxation gives exact lower bound

$$f_{sos}^* \approx -1.03$$

and extracts two points

$$x^* \approx \pm (0.09, -0.71).$$



Figure 1.1: Contour of f(x)

We may plug x^* into f(x), evaluate it and find that $f(x^*) = f_{sos}^*$. In other words, we find a point where the value of the polynomial equals its lower bound f_{sos}^* . Therefore, this lower bound f_{sos}^* equals the minimum of f(x) and x^* is a global minimizer. (More precisely, this is only true up to roundoff. One may construct examples with several points at which the global minimum is nearly attained; in such cases roundoff may prevent us from identifying the correct minimum, or the minimum value very precisely. We will not consider these possibilities in this thesis.)

For the constrained problem (1.0.1)-(1.0.2), SOS relaxations can also be applied in a similar way. This is the frequently used *Lasserre's procedure* in polynomial optimization. For a fixed integer N, one lower bound of f^* can be obtained by the SOS relaxation:

$$f_N^* = \max \quad \gamma \tag{1.1.9}$$

s.t.
$$f(x) - \gamma \equiv \sigma_0(x) + \sigma_1(x)g_1(x) + \dots + \sigma_m(x)g_m(x)$$
(1.1.10)

where $\deg(\sigma_i g_i) \leq 2N$ and σ_i are all SOS polynomials. The integer N is called the degree of the SOS relaxation. The dual problem of (1.1.9)-(1.1.10) is

$$\min_{y} \quad \sum_{\alpha} f_{\alpha} y_{\alpha} \tag{1.1.11}$$

s.t.
$$M_N(y) \succeq 0, \ y_0 = 1$$
 (1.1.12)

$$M_{N-w_i}(g_i * y) \succeq 0 \tag{1.1.13}$$

where $w_i = \lceil \deg(g_i)/2 \rceil$. Here $g_i * y$ denotes another monomial-indexed vector defined as

$$(g_i * y)_{\alpha} = \sum_{\beta} g_{i,\beta} y_{\alpha+\beta} \quad \text{where } g_i(x) = \sum_{\beta} g_{i,\beta} x^{\beta}.$$

Obviously, if γ is feasible in (1.1.10), then $f(x) - \gamma$ must be nonnegative for all feasible points x. Thus every feasible γ and f_N^* are lower bounds of minimum f^* . Furthermore, the lower bound f_N^* is increasing as N increases, since the feasible region of γ defined by (1.1.10) is increasing. Under the *constraint qualification condition*, i.e., supposing there exists M > 0 and SOS polynomials $s_i(x)$ such that the following identity holds

$$M - \|x\|^2 \equiv s_0(x) + s_1(x)g_1(x) + \dots + s_m(x)g_m(x)$$
(1.1.14)

where $||x||^2 = \sum_{i=1}^n x_i^2$, Lasserre [52] showed convergence $\lim_{N \to \infty} f_N^* = f^*$. If this condition holds, we can see that the set of feasible points must be bounded. But the converse might not be true (see Section 2.3). When this constraint qualification condition fails, it might happen that $\lim_{N \to \infty} f_N^* < f^*$.

For an example, consider the following optimization problem

$$\min_{x_1, x_2} \quad -x_1 - x_2$$
s.t. $x_2 \le 2x_1^4 - 8x_1^3 + 8x_1^2 + 2$
 $x_2 \le 4x_1^4 - 32x_1^3 + 88x_1^2 - 96x_1 + 36$
 $0 \le x_1 \le 3, \ 0 \le x_2 \le 4.$

Now we apply Lasserre's procedure to find the global solution. Since the highest degree of the polyno-

mials is 3 and $2N \ge \max_i \deg(g_i) = 3$, the relaxation order N should be at least 2. For N = 2, 3, 4, we solve the relaxation (1.1.9)-(1.1.10) and get the Results

N	f_N^*	minimizer
2	-7	no sol. extracted
3	-6.667	no sol. extracted
4	-5.5080	(2.3295, 3.1785)

Figure 1.2: An example of Lasserre's procedure

in Figure 1.2. When N = 2 or 3, only a

lower bound is returned and no minimizer can be extracted. When N = 4, one lower bound $f_3^* \approx -5.5080$ is returned and a feasible point (2.3295, 3.1785) is extracted from the dual solutions (see [41]). We plug this point into the objective polynomial, evaluate it and find that the value equals the lower bound f_3^* . This implies that the global minimum is $f^* \approx -5.5080$ and one global minimizer is (2.3295, 3.1785).

SOS relaxations are very attractive for solving some hard global optimization problems. There has been a lot of work in this exciting area. We refer to Kojima [47, 50, 114], Laurent [55, 45, 26, 56], Henrion [40, 41], Lasserre [52, 53, 54], Parrilo [79, 80, 81, 26], Schweighofer [101, 103], Sturmfels [80] and many others.

1.1.2 SOS methods versus symbolic methods

The optimization problem (1.0.1)-(1.0.2) can be formulated as a solvability problem of a particular system of polynomial equalities and inequalities, and therefore can be solved using a special case of symbolic methods like Quantifier Elimination (QE). Geometric algorithms for QE exist. We refer the reader to [6, 93]. (In fact, describing an algorithm for the global optimization problem using QE is given as Exercise 14.23 in [6].) The complexity of these methods (e.g., the QE algorithms in [6, 93]) is

$$((1+m)\max\{\deg(f), \deg(g_1), \cdots, \deg(g_m)\})^{O(n)}$$

This exponential complexity is consistent with the NP-hardness of general polynomial optimization problems. On the other hand, very few of these QE algorithms have been implemented and, to our best knowledge, they are rarely applied to polynomial optimization. Therefore it is of interest to find approximation methods for polynomial optimization. SOS relaxation is a special approximation method. It has the advantage that it is easy to implement and can be solved numerically, since the relaxations (1.1.5)-(1.1.6) or (1.1.9)-(1.1.10) are SDPs.

1.2 Contributions of this thesis

The main contributions of this thesis are as follows:

- 1. When the feasible set S is compact and the constraint qualification condition (1.1.14) holds, Lasserre's procedure converges, that is, the lower bounds $\{f_N^*\}$ from (1.1.9)-(1.1.10) converge to the minimum f^* . However, no estimates of the speed of the convergence were available. The author obtained the first upper bound on the convergence rate, in cooperation with M. Schweighofer. The convergence rate analysis is based on the degree bounds in Putinar's Theorem. This will be presented in Chapter 3.
- 2. To solve the unconstrained optimization (1.0.1), SOS relaxation (1.1.5)-(1.1.6) generally only provides a lower bound f_{sos}^* . But sometimes it happens that $f_{sos}^* < f^*$. In such situations, how can we get better lower bounds? A very good lower bound can be obtained if we apply SOS relaxation over the gradient ideal of the polynomial f(x). In fact, we can always get the exact lower bound, and have finite convergence, under some conditions that hold generically. This will be presented in Chapter 4.
- 3. In the constrained optimization problem (1.0.1)-(1.0.2), we may not have convergence $\lim_{N\to\infty} f_N^* = f^*$ if the semialgebraic set S is not compact. How can we get better lower bounds in such situations? Similarly to the unconstrained

case, a very good lower bound can be obtained if we apply SOS relaxation over the Kuhn-Karush-Tucker (KKT) ideal. This lower bound is exact, and we have finite convergence, under some conditions that hold generically. This will be presented in Chapter 5.

- 4. There is a broader class of optimization problems which are described by rational functions. To our best knowledge, there is little work on the global optimization of rational function using SOS methods, even though there is a direct way to do so. We studied SOS methods for minimizing rational functions exploiting their special features. This will be presented in Chapter 6.
- 5. Polynomial optimization problems have wide applications. We studied applications in shape optimization of transfer functions, finding minimum ellipsoid bounds for polynomial systems, solving the nearest GCD problem, maximum likelihood optimization, and sensor network localization. These will be presented in Chapter 7.

Chapter 2

Some Basic Tools for Polynomial Optimization

The basic idea in polynomial optimization is to use sum of squares (SOS) representations of polynomials which are positive on some feasible sets defined by polynomial equalities and/or inequalities. The attractive property of the SOS representation is that it can be reduced to some particular semidefinite program (SDP). To study the SOS representation, we need some basic tools from algebraic geometry, real algebra and the theory of positive polynomials. They are the fundamentals of polynomial optimization.

2.1 SOS and nonnegative polynomials

A polynomial $p(x) \in \mathbb{R}[X]$ is nonnegative if $p(x) \ge 0$ for all $x \in \mathbb{R}^n$. A polynomial $p(x) \in \mathbb{R}[X]$ is a Sum Of Squares (SOS) if $p(x) = \sum_i q_i^2(x)$ for some finite number of $q_i(x) \in \mathbb{R}[X]$. Denote by $\sum \mathbb{R}[X]^2$ the set of all SOS polynomials. Obviously, if p(x) is SOS, then p(x) is always nonnegative. Example 2.1.1. The following identity

$$3 \cdot (x_1^4 + x_2^4 + x_3^4 + x_4^4 - 4x_1x_2x_3x_4)$$

= $(x_1^2 - x_2^2 - x_4^2 + x_3^2)^2 + (x_1^2 + x_2^2 - x_4^2 - x_3^2)^2 + (x_1^2 - x_2^2 - x_3^2 + x_4^2)^2 + 2(x_1x_4 - x_2x_3)^2 + 2(x_1x_2 - x_3x_4)^2 + 2(x_1x_3 - x_2x_4)^2$

shows that the polynomial $x_1^4 + x_2^4 + x_3^4 + x_4^4 - 4x_1x_2x_3x_4$ is SOS and hence nonnegative.

Example 2.1.2. The following polynomial

$$2x_1^4 + 2x_1^3x_2 - x_1^2x_2^2 + 5x_2^4 = \frac{1}{2} \left[(2x_1^2 - 3x_2^2 + x_1x_2)^2 + (x_2^2 + 3x_1x_2)^2 \right]$$

is SOS and hence nonnegative.

The set $\sum \mathbb{R}[X]^2$ is a cone within the polynomial ring $\mathbb{R}[X]$, since the following three properties hold: (i) if $f, g \in \sum \mathbb{R}[X]^2$, then $f + g \in \sum \mathbb{R}[X]^2$; (ii) if $f, g \in \sum \mathbb{R}[X]^2$, then $f \cdot g \in \sum \mathbb{R}[X]^2$; (iii) for any $f \in \mathbb{R}[X]$, $f^2 \in \sum \mathbb{R}[X]^2$.

As we have seen, p(x) being SOS implies that p(x) is nonnegative. However, the converse may not be true. For instance, the *Motzkin* polynomial

$$M(x) := x_1^4 x_2^2 + x_1^2 x_2^4 + x_3^6 - 3x_1^2 x_2^2 x_3^2$$

is nonnegative, but not SOS [95]. The following theorem characterizes the relationship between nonnegative and SOS polynomials:

Theorem 2.1.3 (Hilbert, 1888). Let $\mathcal{P}_{n,d}$ be the set of all nonnegative polynomials in n variables with degree at most d, and $\Sigma_{n,d}$ be the set of all SOS polynomials in nvariables with degree at most d. Then $\mathcal{P}_{n,d}=\Sigma_{n,d}$ if and only if n = 1, or d = 2, or (n,d) = (2,4). Denote by m(x) the column vector of monomials up to degree d

$$m(x)^T = [1, x_1, \cdots, x_n, x_1^2, x_1x_2, \cdots, x_n^2, x_1^3, \cdots, x_n^d]$$

Notice that the length of vector m(x) is $\binom{n+d}{d}$. Let p(x) be a polynomial with degree 2d. Then p(x) is SOS if and only if [81, 105] there exists a real symmetric matrix $W \succeq 0$ of dimension $\binom{n+d}{d}$ such that the identity holds:

$$p(x) \equiv m(x)^T W m(x).$$

Now we write p(x) as $\sum_{\alpha \in P} p_{\alpha} x^{\alpha}$, where P is a finite subset of \mathbb{N}^n , i.e., P is the support of polynomial p(x). Let B_{α} be the $\binom{n+d}{d}$ -dimensional coefficient matrix of x^{α} in $m(x) \cdot m(x)^T$, i.e,.

$$m(x) \cdot m(x)^T = \sum_{|\alpha| \le 2d} B_{\alpha} x^{\alpha}.$$

Then we can see that p(x) is SOS if and only if there exists a symmetric matrix W such that

$$W \succeq 0, \quad \langle W, B_{\alpha} \rangle = p_{\alpha}, \quad \forall \alpha \in P.$$

The inner product $\langle \cdot, \cdot \rangle$ above is defined as $\langle A, B \rangle = Trace(AB)$ for any two symmetric matrices A, B. Testing whether a polynomial is SOS or not can be done by solving a SDP feasibility problem. The condition that a polynomial is SOS poses an Linear Matrix Inequality (LMI) constraint on the coefficients of the polynomial. See [81] for more detailed descriptions of connections between SOS polynomials and SDP.

2.2 Elementary algebraic geometry

This section will introduce some basic tools in algebraic geometry. Readers may consult [13, 21, 22, 29] for more details.

A subset I of $\mathbb{R}[X]$ is an *ideal* if $p \cdot q \in I$ for any $p \in I$ and $q \in \mathbb{R}[X]$. For $p_1, \ldots, p_r \in \mathbb{R}[X], \langle p_1, \cdots, p_r \rangle$ denotes the smallest ideal containing the p_i . Equivalently, $\langle p_1, \cdots, p_r \rangle$ is the set of all polynomials that are polynomial linear combinations of the p_i . Every ideal arises in this way:

Theorem 2.2.1 (Hilbert Basis Theorem). Every ideal $I \subset \mathbb{R}[x]$ has a finite generating set, i.e., $I = \langle p_1, \dots, p_\ell \rangle$ for some $p_1, \dots, p_\ell \in I$.

The variety of an ideal I is the set of all common **complex** zeros of the polynomials in I:

$$V(I) = \{x \in \mathbb{C}^n : p(x) = 0 \text{ for all } p \in I\}.$$

The subset of all real points in V(I) is the real variety of I. It is denoted

$$V^{\mathbb{R}}(I) = \{x \in \mathbb{R}^n : p(x) = 0 \text{ for all } p \in I\}.$$

If $I = \langle p_1, \ldots, p_r \rangle$ then $V(I) = V(p_1, \ldots, p_r) = \{x \in \mathbb{C}^n : p_1(x) = \cdots = p_r(x) = 0\}$. An ideal $I \subseteq \mathbb{R}[X]$ is zero-dimensional if its variety V(I) is a finite set. This condition is much stronger than requiring that the real variety $V^{\mathbb{R}}(I)$ be a finite set. For example, $I = \langle x_1^2 + x_2^2 \rangle$ is not zero-dimensional, however the real variety $V^{\mathbb{R}}(I) = \{(0,0)\}$ consists of one point of the curve V(I).

A variety $V \subseteq \mathbb{C}^n$ is *irreducible* if there do not exist two proper subvarieties $V_1, V_2 \subseteq V$ such that $V = V_1 \cup V_2$. Here "irreducible" means that the set of **complex** zeros cannot be written as a proper union of subvarieties defined by **real** polynomials. Given a variety $V \subseteq \mathbb{C}^m$, the set of all polynomials that vanish on V is an ideal

$$I(V) = \{ p \in \mathbb{R}[X] : p(u) = 0 \text{ for all } u \in V \}.$$

Given any ideal I of $\mathbb{R}[X]$, its *radical* is the ideal

$$\sqrt{I} = \{ q \in \mathbb{R}[X] : q^{\ell} \in I \text{ for some } \ell \in \mathbb{N} \}.$$

Note that $I \subseteq \sqrt{I}$. We say that I is a *radical ideal* if $\sqrt{I} = I$. Clearly, the ideal I(V) defined by a variety V is a radical ideal. The following theorems offer a converse to this observation:

Theorem 2.2.2 (Hilbert's Weak Nullstellensatz).

If I is an ideal in $\mathbb{R}[X]$ such that $V(I) = \emptyset$ then $1 \in I$.

Theorem 2.2.3 (Hilbert's Strong Nullstellensatz).

If I is an ideal in $\mathbb{R}[X]$ then $I(V(I)) = \sqrt{I}$.

2.3 Positive polynomials on semialgebraic sets

In polynomial optimization problems, we are often interested in a feasible set S of the form

$$S = \{ x \in \mathbb{R}^n : g_1(x) \ge 0, \cdots, g_m(x) \ge 0 \}$$

where $g_i(x) \in \mathbb{R}[X]$. Such an S is called a *basic closed semialgebraic set*, and plays an important role in real algebraic geometry [13]. Notice that different polynomial tuples $(g_1(x), \dots, g_m(x))$ may define the same semialgebraic set S in \mathbb{R}^n , but these tuples might have different algebraic properties (e.g., archimedeanness as defined below). So when we refer to a semialgebraic set S, we assume that a polynomial tuple $(g_1(x), \dots, g_m(x))$ (which is often clear from the text) is associated with it. Given S with polynomial tuple $(g_1(x), \dots, g_m(x))$, the *preorder* and *linear* cones associated with S are defined as

$$\mathcal{P}(S) = \left\{ \sum_{\theta \in \{0,1\}^m} s_{\theta}(x) g_1^{\theta_1}(x) \cdots g_{\ell}^{\theta_{\ell}}(x) \middle| s_{\theta} \in \Sigma \mathbb{R}[X]^2 \right\}$$
$$\mathcal{M}(S) = \left\{ \sigma_0(x) + \sum_{j=1}^{\ell} g_j(x) \sigma_j(x) \middle| \sigma_i \in \Sigma \mathbb{R}[X]^2 \right\}.$$

 $\mathcal{M}(S)$ is also called the *quadratic module* generated by S. We also denote by $\mathcal{P}(S)_N$ (and $\mathcal{M}(S)_N$ respectively) the subset of $\mathcal{P}(S)$ (and $\mathcal{M}(S)$ respectively) such that the degree in each summand is no greater than N.

A subset $M \subseteq \mathbb{R}[X]$ is called a quadratic module if it contains 1 and it is closed under addition and under multiplication with squares, i.e.,

$$1 \in M$$
, $M + M \subseteq M$ and $\mathbb{R}[X]^2 M \subseteq M$.

A subset $T \subseteq \mathbb{R}[x]$ is called a *preordering* if it contains all squares in $\mathbb{R}[X]$ and it is closed under addition and multiplication, i.e.,

$$\mathbb{R}[X]^2 \subseteq T, \quad T+T \subseteq T \quad \text{and} \quad TT \subseteq T.$$

In other words, the preorderings are exactly the multiplicatively closed quadratic modules. In 1991, Schmüdgen [98] proved the following "Positivstellensatz" (a commonly used German term explained by the analogy with Hilbert's Nullstellensatz).

Theorem 2.3.1 (Schmüdgen's Positivstellensatz, [98]). Suppose the set S is compact. Then every polynomial p(x) which is positive on S belongs to $\mathcal{P}(S)$.

The quadratic module $\mathcal{M}(S)$ is archimedean if there exists $\rho(x) \in \mathcal{M}(S)$ such that the set $\{x \in \mathbb{R}^m : \rho(x) \ge 0\}$ is compact, equivalently, if there exists $N \in \mathbb{N}$ such that $N - \sum_{i=1}^n x_i^2 \in \mathcal{M}(S)$. The condition that $\mathcal{M}(S)$ is archimedean is also called *Putinar's constraint qualification* [91], or constraint qualification condition [52].

In particular, we see that S is compact if and only if $\mathcal{P}(S)$ is archimedean. Unfortunately, S might be compact without $\mathcal{M}(S)$ being archimedean (see [27, Example 6.3.1]). What has to be added to compactness of S in order to ensure that $\mathcal{M}(S)$ is archimedean has been extensively investigated by Jacobi and Prestel [44, 27]. Now we can state the Positivstellensatz proved by Putinar [91] in 1993.

Theorem 2.3.2 (Putinar's Positivstellensatz, [91]). Suppose $\mathcal{M}(S)$ is archimedean. Then every polynomial p(x) which is positive on S belongs to $\mathcal{M}(S)$.

Remark 2.3.3. There are examples of compact S for which $\mathcal{M}(S)$ is not archimedean and the conclusion of Putinar's Theorem does not hold. For instance, for S = $\{(x_1, x_2) : 2x_1 - 1 \ge 0, 2x_2 - 1 \ge 0, 1 - x_1x_2 \ge 0\}, M - x_1^2 - x_2^2 \notin \mathcal{M}(S)$ for any M > 0. Otherwise, suppose $M - x_1^2 - x_2^2 = s_0 + s_1(2x_1 - 1) + s_2(2x_2 - 1) + s_3(1 - x_1x_2)$ for some SOS polynomials s_0, s_1, s_2, s_3 . Since the highest degree on the left hand side is 2, the highest degree in the right hand side must come from $s_0 + s_3(1 - x_1x_2)$. If $s_3 \equiv 0$, the leading coefficient on the right hand side is nonnegative, which is a contradiction. If $s_3 \neq 0$, since the leading coefficient of s_0 is nonnegative, the leading term must come from $s_3(1 - x_1x_2)$. But the leading term of $s_3(1 - x_1x_2)$ is of the form $x_1^{2k+1}x_2^{2\ell+1}$, which is a contradiction by comparison with the left hand side. In practice, if we know some integer N such that S is contained in the sphere $\{x \in \mathbb{R}^n : N - \sum_{i=1}^n x_i^2 \ge 0\}$, we can add the redundant constraint $N - \sum_{i=1}^n x_i^2 \ge 0$ to force $\mathcal{M}(S)$ to be archimedean.

Theorem 2.3.4 (Parrilo, [89]). Suppose $\mathcal{M}(S)$ contains an ideal $J = \langle p_1, \cdots, p_r \rangle$. If J is a zero-dimensional radical ideal in $\mathbb{R}[X]$, then a polynomial $w(X) \in \mathbb{R}[X]$ is nonnegative on S if and only if $w(X) \in \mathcal{M}(S)$.

Theorem 2.3.5 (Stengle's Positivstellensatz,[108]). Suppose S and $\mathcal{P}(S)$ are defined as above. Then $S = \emptyset$ if and only if $-1 \in \mathcal{P}(S)$.

Definition 2.3.6 (Moment Matrix). Let $y = (y_{\alpha})_{\alpha \in \mathbb{N}^n}$ be a sequence indexed by $\alpha = (\alpha_1, \dots, \alpha_n)$, i.e., it is multi-indexed or monomial-indexed. Then the moment matrix induced by the vector y is $M(y) = (y_{\alpha+\beta})$, i.e., the (α, β) -th entry of M(y) is $y_{\alpha+\beta}$.

The N-th truncation matrix $M_N(y)$ of M(y) is the leading submatrix such that $M_N(y) = (y_{\alpha+\beta})_{|\alpha| \le N, |\beta| \le N}$. For instance, when n = 1,

$$M_3(y) = \begin{bmatrix} y_0 & y_1 & y_2 & y_3 \\ y_1 & y_2 & y_3 & y_4 \\ y_2 & y_3 & y_4 & y_5 \\ y_3 & y_4 & y_5 & y_6 \end{bmatrix}.$$

For a polynomial $h = \sum_{\beta} h_{\beta} x^{\beta}$, define the convolution of h and vector y as the new multi-indexed vector $(h * y)_{\alpha} = \sum_{\beta} h_{\beta} y_{\alpha+\beta}$. The multi-indexed vector y is a moment sequence if there exists a measure μ on \mathbb{R}^n such that $y_{\alpha} = \int_{\mathbb{R}^n} x^{\alpha} \mu(dx)$. See [24] for more details about moment theories.

When n = 1, i.e., in case of univariate polynomials, there are some characterizations of polynomials which are nonnegative on some interval. We refer to [85]. We will use these results in Section 7.1. **Theorem 2.3.7 (Markov, Lukacs [59, 61, 84]).** Let $q(t) \in \mathbb{R}[t]$ be a real polynomial of degree n. Let $n_1 = \lfloor \frac{n}{2} \rfloor$ and $n_2 = \lfloor \frac{n-1}{2} \rfloor$. If $q(t) \ge 0$ for all $t \ge 0$, then $q(t) = q_1(t)^2 + tq_2(t)^2$ where $deg(q_1) \le n_1$ and $deg(q_2) \le n_2$.

Theorem 2.3.8 (Markov, Lukacs [59, 61, 84]). Let $q(t) \in \mathbb{R}[t]$ be a real polynomial. Suppose $q(t) \ge 0$ for all $t \in [a, b]$, then one of the following holds.

- 1. If deg(q) = n = 2m is even, then $q(t) = q_1(t)^2 + (t-a)(b-t)q_2(t)^2$ where $deg(q_1) \le m$ and $deg(q_2) \le m 1$.
- 2. If deg(q) = n = 2m + 1 is odd, then $q(t) = (t a)q_1(t)^2 + (b t)q_2(t)^2$ where $deg(q_1) \le m$ and $deg(q_2) \le m$.

Chapter 3

On the Convergence Rate of Lasserre's Procedure

Consider the constrained polynomial optimization problem

$$f^* = \min_{x \in S} \quad f(x)$$

where $S = \{x \in \mathbb{R}^n : g_1(x) \ge 0, \cdots, g_m(x) \ge 0\}$. Lasserre [52] proposed the SOS relaxation (k is the order)

$$f_k^* = \sup \quad \gamma \quad \text{ s.t. } \quad f(x) - \gamma \in \mathcal{M}_{2k}(S).$$

Obviously each f_k^* is a lower bound of f^* . Under condition (1.1.14), Lasserre [52] showed convergence $\lim_{N\to\infty} f_k^* = f^*$. A naturally arising question is how fast does f_k^* converge to f^* ? This chapter will give the *first* estimate on the convergence rate of $f_k^* \to f^*$ as k goes to infinity. This is joint work with Markus Schweighofer [72].

3.1 Convergence rate of Lasserre's procedure

Let $S = S(\bar{g}) := \{x \in \mathbb{R}^n : g_1(x) \ge 0, \cdots, g_m(x) \ge 0\}$ be the feasible set, where $\bar{g} = (g_1, \cdots, g_m)$ is the tuple of polynomials defining the constraints. For convenience, set $g_0(x) \equiv 1$. Denote by $\mathcal{M}(S)$ ($\mathcal{M}(S)_N$) the (truncated) quadratic module generated by the tuple \bar{g} , i.e.,

$$\mathcal{M}(S) = \left\{ \sum_{j=0}^{m} \sigma_j(x) g_j(x) \middle| \sigma_j \in \Sigma \mathbb{R}[X]^2 \right\}$$
$$\mathcal{M}(S)_N = \left\{ \sum_{j=0}^{m} \sigma_j(x) g_j(x) \middle| \sigma_j \in \Sigma \mathbb{R}[X]^2, \deg(\sigma_j g_j) \le N \right\}.$$

For each integer k, we can see that

$$f_k^* := \sup\{\gamma \in \mathbb{R} \mid f - \gamma \in \mathcal{M}(S)_{2k}\} \in \mathbb{R} \cup \{-\infty\}.$$
(3.1.1)

The problem of finding f_k^* is essentially a semidefinite program (SDP) whose size gets bigger as k grows (see [52, 80, 81]). One can now solve a sequence of larger and larger semidefinite programs in order to get tighter and tighter lower bounds for f^* . Lasserre [52] showed convergence by applying Putinar's Positivstellensatz.

Indeed, it is easy to see that Putinar's theorem just says that the ascending sequence $(f_k^*)_{k\in\mathbb{N}}$ converges to f^* under the condition that $\mathcal{M}(S)$ be archimedean (see Section 2.3). In this section, we will interpret our bound for Putinar's Positivstellensatz as a result about the speed of convergence of this sequence.

To get the bound for Putinar's Positivstellensatz, we will need a convenient measure of the size of the coefficients of a polynomial. For $\alpha \in \mathbb{N}^n$, we introduce the notation

$$|\alpha| := \alpha_1 + \dots + \alpha_n$$
 and $x^{\alpha} := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$

as well as the multinomial coefficient

$$\binom{|\alpha|}{\alpha} := \frac{|\alpha|!}{\alpha_1! \dots \alpha_n!}.$$

For a polynomial $f(x) = \sum_{\alpha} a_{\alpha} x^{\alpha} \in \mathbb{R}[X]$ with coefficients $a_{\alpha} \in \mathbb{R}$, we set

$$||f|| := \max_{\alpha} \frac{|a_{\alpha}|}{\binom{|\alpha|}{\alpha}}.$$

This defines a norm on the real vector space $\mathbb{R}[X]$ with convenient properties illustrated by Proposition 3.3.1 in Section 3.3. The following technical lemma estimates the value of a polynomial in term of its norm on the unit box, which will be needed in Section 3.3.

Lemma 3.1.1. For any polynomial $f \in \mathbb{R}[X]$ of degree $d \ge 1$ and all $x \in [-1, 1]^n$,

$$|f(x)| \le 2dn^d ||f||.$$

Proof. Writing $f = \sum_{\alpha} a_{\alpha} {|\alpha| \choose \alpha} \bar{X}^{\alpha}$ $(a_{\alpha} \in \mathbb{R})$, we have $||f|| = \max_{\alpha} |a_{\alpha}|$ and

$$|f(x)| = \left|\sum_{\alpha} a_{\alpha} \binom{|\alpha|}{\alpha} x_{1}^{\alpha_{1}} \cdots x_{n}^{\alpha_{n}}\right| \leq \sum_{\alpha} |a_{\alpha}| \binom{|\alpha|}{\alpha} |x_{1}|^{\alpha_{1}} \cdots |x_{n}|^{\alpha_{n}}.$$

for all $x \in [-1,1]^n$. Using that $|a_{\alpha}| \leq ||f||$ and $|x_i| \leq 1$, the multinomial identity now shows that $|f(x)| \leq ||f|| \sum_{k=0}^d n^k \leq (d+1)n^d ||f|| \leq 2dn^d ||f||$.

Now we are ready to prove the main theorem of this section, which is based on the complexity result in Theorem 3.2.4.

Theorem 3.1.1. For every polynomial tuple \bar{g} defining an archimedean quadratic module $\mathcal{M}(S)$ and a set $\emptyset \neq S = S(\bar{g}) \subseteq (-1,1)^n$, there is some c > 0 (depending on \bar{g}) such that for all $f \in \mathbb{R}[X]$ of degree d with minimum f^* on S and for all integers $k > \frac{1}{2}c \exp((2d^2n^d)^c)$, we have

$$(f - f^*) + \frac{6d^3n^{2d}||f||}{\sqrt[c]{\log \frac{2k}{c}}} \in \mathcal{M}(S)_{2k}$$

and hence

$$0 \le f^* - f_k^* \le \frac{6d^3n^{2d} \|f\|}{\sqrt[c]{\log \frac{2k}{c}}}$$

where f_k^* is defined as in (3.1.1).

Proof. Given \bar{g} , we choose c > 0 as in Theorem 3.2.4. Now let $f \in \mathbb{R}[X]$ be of degree d with minimum f^* on S and let

$$k > \frac{1}{2}c\exp((2d^2n^d)^c)$$
(3.1.2)

be an integer. The case d = 0 is trivial. We assume therefore $d \ge 1$. Note that $k > \frac{c}{2}$ and hence $\log(2k/c) > 0$. Setting

$$a := \frac{6d^3n^{2d}||f||}{\sqrt[c]{\log\frac{2k}{c}}},\tag{3.1.3}$$

all we have to prove is $h := f - f^* + a \in \mathcal{M}(S)_{2k}$ because the second claim follows from this. By our choice of c and the observation deg $h = \deg f = d$, it is enough to show that

$$c \exp\left(\left(d^2 n^d \frac{\|h\|}{a}\right)^c\right) \le 2k,$$

or equivalently

$$d^{2}n^{d} \|h\| \le a \sqrt[c]{\log \frac{2k}{c}} = 6d^{3}n^{2d} \|f\|.$$

Observing that $||h|| \le ||f|| + |f^*| + a$, it suffices to show that

$$||f|| + |f^*| + a \le 6dn^d ||f||.$$

Lemma 3.1.1 tells us that $|f^*| \leq 2dn^d ||f||$ and we are thus reduced to verifying that

$$a \le (4dn^d - 1) \|f\|$$

which is by (3.1.3) equivalent to

$$6d^3n^{2d} \le (4dn^d - 1)\sqrt[c]{\log\frac{2k}{c}}.$$

By (3.1.2), it is finally enough to check that $6d^3n^{2d} \leq (4dn^d - 1)(2d^2n^d)$.

The hypothesis that $S(\bar{g})$ is contained in the open unit hypercube is just a technical assumption, which helps avoid a more complicated bound (see Remark 3.2.1). If one does not insist on all the information given in Theorem 3.1.1, one can get a corollary which is easy to remember and still gives the most important part of the information.

Corollary 3.1.2. Suppose $\mathcal{M}(S)$ is archimedean, $S(\bar{g}) \neq \emptyset$ and $f \in \mathbb{R}[X]$. There is

- a constant c > 0 depending only on \bar{g} and
- a constant c' > 0 depending on \overline{g} and f

such that for f^* and f^*_k as defined in (3.1.1)),

$$0 \le f^* - f_k^* \le \frac{c'}{\sqrt[c]{\log \frac{2k}{c}}} \quad \text{for all large } k \in \mathbb{N}.$$

Proof. Without loss of generality, assume $f \neq 0$. Set $d := \deg f$. Since $\mathcal{M}(S)$ is archimedean, S is compact. We can hence choose a rescaling factor r > 0 depending only on \bar{g} such that $S(\bar{g}(rx)) \subseteq (-1,1)^n$. Here $\bar{g}(rx)$ denotes the tuple of rescaled polynomials $g_i(r\bar{X})$. Now Theorem 3.1.1 applied to g(rx) instead of \bar{g} yields c > 0that will together with $c' := 6d^3n^{2d}||f(rx)||$ have the desired properties by simple scaling arguments. **Remark 3.1.3.** The bound on the difference $f^* - f_k^*$ presented in this section is much worse than the corresponding one presented in [100, Section 2] which is based on preordering representations (i.e., where f_k^* would be defined using $\mathcal{P}(S)$ instead of $\mathcal{M}(S)$. This raises the question whether it is after all not such a bad thing to use preordering (instead of quadratic module) representations for optimization though they involve the 2^m products \bar{g}^{δ} thereby letting the semidefinite programs get huge when m grows. However, it is not known if Theorem 3.1.1 holds perhaps even with the bound from [100, Theorem 4]. Compare also [100, Remark 5].

3.2 On the complexity of Putinar's positivstellensatz

Denote by $\bar{g} := (g_1, \ldots, g_m)$ the tuple of polynomials defining the constraints, and set $g_0 := 1 \in \mathbb{R}[X]$ for convenience. The quadratic module $\mathcal{M}(S)$ generated by \bar{g} is

$$\mathcal{M}(S) := \left\{ \sum_{i=0}^{m} \sigma_i g_i \mid \sigma_i \in \sum \mathbb{R}[X]^2 \right\}.$$
(3.2.4)

Using the notation

$$\bar{g}^{\delta} := g_1^{\delta_1} \dots g_m^{\delta_m},$$

the preordering $\mathcal{P}(S)$ generated by \bar{g} can be written as

$$\mathcal{P}(S) := \left\{ \sum_{\delta \in \{0,1\}^m} \sigma_{\delta} \bar{g}^{\delta} \mid \sigma_{\delta} \in \sum \mathbb{R}[X]^2 \right\},$$
(3.2.5)

i.e., $\mathcal{P}(S)$ is the quadratic module generated by the 2^m products of g_i . It is obvious that all polynomials lying in $\mathcal{P}(S) \supseteq \mathcal{M}(S)$ are nonnegative on the feasible set

$$S := \{ x \in \mathbb{R}^n \mid g_1(x) \ge 0, \dots, g_m(x) \ge 0 \}.$$
Sets of this form are called *basic closed semialgebraic sets* (see [13]). In 1991, Schmüdgen [98] proved the following "Positivstellensatz" (a commonly used German term explained by the analogy with Hilbert's Nullstellensatz).

Theorem 3.2.1 (Schmüdgen). Suppose the feasible set S is compact. Then for every $f \in \mathbb{R}[X]$,

$$f > 0 \text{ on } S \implies f \in \mathcal{P}(S).$$

Under a certain extra property that $\mathcal{M}(S)$ is archimedean, i.e., $N - ||x||^2 \in \mathcal{M}(S)$ for some N > 0 (see Section 2.3), the Theorem 3.2.1 remains true if $\mathcal{P}(S)$ is replaced by $\mathcal{M}(S)$.

In particular, we see that $S = S(\bar{g})$ is compact if and only if $\mathcal{P}(S)$ is archimedean. Unfortunately, S might be compact without $\mathcal{M}(S)$ being archimedean (see [27, Example 6.3.1]). What has to be added to compactness of S in order to ensure that $\mathcal{M}(S)$ is archimedean has been extensively investigated by Jacobi and Prestel [44, 27]. Now we can state the Positivstellensatz proved by Putinar [91] in 1993.

Theorem 3.2.2 (Putinar). Suppose $\mathcal{M}(S)$ is archimedean. Then for every $f \in \mathbb{R}[X]$,

$$f > 0 \text{ on } S(\bar{g}) \implies f \in \mathcal{M}(S).$$

Both the proofs of Schmüdgen and Putinar use functional analysis and real algebraic geometry. They do not give information how to construct a representation of f showing that f lies in the preordering (an expression like (3.2.5) involving 2^m sums of squares) or the quadratic module (a representation like (3.2.4) with m + 1 sums of squares).

Based on an old theorem of Pólya [83], new proofs of both Schmüdgen's and Putinar's Positivstellensatz have been given in [99, 101] which are to some extent constructive. By carefully analyzing a tame version of [101] and using an effective version of Pólya's theorem [87], upper bounds on the degrees of the sums of squares appearing in Schmüdgen's preordering representation have been obtained in [100]. The aim of this section is to prove bounds on Putinar's Theorem. They will depend on the same data but will be worse than the ones known for Schmüdgen's theorem.

For any $k \in \mathbb{N}$, define the truncated convex cones $\mathcal{P}(S)_k$ and $\mathcal{M}(S)_k$ in the finite-dimensional vector space $\mathbb{R}[X]_{\leq k}$ of polynomials of degree at most k by setting

$$\mathcal{P}(S)_{k} = \left\{ \sum_{\delta \in \{0,1\}^{m}} \sigma_{\delta} \bar{g}^{\delta} \mid \sigma_{\delta} \in \sum \mathbb{R}[X]^{2}, \deg(\sigma_{\delta} \bar{g}^{\delta}) \leq k \right\} \quad \subseteq \mathcal{P}(S) \cap \mathbb{R}[X]_{\leq k},$$
$$\mathcal{M}(S)_{k} = \left\{ \sum_{i=0}^{m} \sigma_{\delta} \bar{g}^{\delta} \mid \sigma_{\delta} \in \sum \mathbb{R}[X]^{2}, \deg(\sigma_{\delta} \bar{g}^{\delta}) \leq k \right\} \quad \subseteq M(s) \cap \mathbb{R}[X]_{\leq k}$$

We now recall the previously proved bound for Schmüdgen's theorem.

Theorem 3.2.3 ([100]). For all \bar{g} defining a basic closed semialgebraic set $S(\bar{g})$ which is non-empty and contained in the open hypercube $(-1,1)^n$, there is some $c \geq 1$ (depending on \bar{g}) such that for all $f \in \mathbb{R}[X]$ of degree d with

$$f^* := \min\{f(x) \mid x \in S(\bar{g})\} > 0,$$

we have $f \in \mathcal{P}(S)_N$ with

$$N = cd^2 \left(1 + \left(d^2 n^d \frac{\|f\|}{f^*} \right)^c \right).$$

In this chapter, we will prove the following bound for Putinar's theorem.

Theorem 3.2.4. For all \bar{g} defining an archimedean quadratic module $\mathcal{M}(S)$ and a set $\emptyset \neq S(\bar{g}) \subseteq (-1,1)^n$, there is a constant c > 0 (depending on \bar{g}) such that for all $f \in \mathbb{R}[X]$ of degree d with

$$f^* := \min\{f(x) \mid x \in S(\bar{g})\} > 0,$$

we have $f \in \mathcal{M}(S)_N$ with

$$N = c \exp\left(\left(d^2 n^d \frac{\|f\|}{f^*}\right)^c\right).$$

Remark 3.2.1. In both theorems above, there have been additional assumptions made compared to the original theorems. But these are not very serious and have only been made to simplify the statements. For example, if $S = \emptyset$, then $-1 \in \mathcal{P}(S)_k$ for some $k \in \mathbb{N}$ by Schmüdgen's theorem. Therefore $4f = (f+1)^2 + (f-1)^2(-1) \in$ $\mathcal{P}(S)_{2d+k}$ for each $f \in \mathbb{R}[X]$ of degree $d \ge 0$. The other hypothesis that $S(\bar{g})$ be contained in the open hypercube $(-1,1)^n$ is only a matter of rescaling by a linear (or affine linear) transformation on \mathbb{R}^n . For example, if r > 0 is such that $S \subseteq (-r, r)^n$, then Theorem 3.2.3 remains true with ||f|| replaced by ||f(rx)||. Here it is important to note that the property that $\mathcal{M}(S)$ be archimedean is preserved under affine linear coordinate changes.

In both Theorem 3.2.3 and 3.2.4, the bound depends on three parameters:

- The description \bar{g} of the basic closed semialgebraic set,
- the degree d of f and
- a measure of how close f comes to have a zero on $S(\bar{g})$, namely $||f||/f^*$.

The main difference between the two bounds is the exponential function appearing in the degree bound for the quadratic module representation. It is an open problem whether this exponential function can be avoided. It could even be possible that the same bound for Schmüdgen's theorem holds also for Putinar's theorem. In view of the impact on the convergence rate of Lasserre's optimization procedure (see Section 3.1), this question seems very interesting for applications. Whereas the bound for the preordering representation cannot be improved significantly (see [109]), we are not sure whether this is possible for the quadratic module representation.

The dependence on the third parameter $||f||/f^*$ is consistent with the fact that the condition $f^* > 0$ cannot be weakened to $f^* \ge 0$ in either Schmüdgen's nor Putinar's theorem. Under certain conditions (e.g., on the derivatives of f), both theorems can however be extended to nonnegative polynomials (see [97, 63]). With the partially constructive approach from [102] applied to representations of nonnegative polynomials with zeros, one might perhaps in the future gain bounds even for the case of nonnegative polynomials which depend on further data (for example the norm of the Hessian at the zeros).

In contrast to this, our more constructive approach yields information in what way the above bound depends on the two parameters d and $||f||/f^*$. The constant c depends on the description \bar{g} of the semialgebraic set, but no explicit formula is given. For a concretely given \bar{g} , one could possibly determine a constant cin Theorems 3.2.3 and 3.2.4 by a very tedious analysis of the proofs (cf. [100, Remark 10]).

3.3 The proof of Theorem 3.2.4

In this section, we give the proof of Theorem 3.2.4. The three main ingredients in the proof are

- the bound for Schmüdgen's theorem presented in Theorem 3.2.3 above,
- ideas from the (to some extent constructive) proof of Putinar's theorem in [101, Section 2] and
- the Łojasiewicz inequality from semialgebraic geometry.

We start with some simple facts from calculus.

Lemma 3.3.1. If $0 \neq f \in \mathbb{R}[X]$ has degree d, then

$$|f(x) - f(y)| \le ||x - y||_2 d^2 n^{d-1} \sqrt{n} ||f||$$

for all $x, y \in [-1, 1]^n$.

Proof. Denoting by Df the derivative of f, by the mean value theorem, it is enough to show that

$$|Df(x)(e)| \le d^2 n^{d-1} \sqrt{n} ||f||$$
(3.3.6)

for all $x \in [-1,1]^n$ and $e \in \mathbb{R}^n$ with $||e||_2 = 1$. A small computation (compare the proof of Lemma 3.1.1) shows that

$$\left|\frac{\partial f(x)}{\partial x_i}\right| \le \|f\| \sum_{k=1}^d k(|x_1| + \dots + |x_n|)^{k-1} \le \|f\| \sum_{k=1}^d kn^{k-1} \le \|f\| d^2 n^{d-1},$$

from which we conclude for all $x \in [-1,1]^n$ and $e \in \mathbb{R}^n$ with ||e|| = 1,

$$|Df(x)(e)| = \left|\sum_{i=1}^{n} \frac{\partial f(x)}{\partial x_i} e_i\right| \le \sum_{i=1}^{n} \left|\frac{\partial f(x)}{\partial x_i}\right| \cdot |e_i| \le ||f|| d^2 n^{d-1} \sum_{i=1}^{n} |e_i|.$$

Because for a vector e on the unit sphere in \mathbb{R}^n , $\sum_{i=1}^n |e_i|$ can reach at most \sqrt{n} , this implies (3.3.6).

Remark 3.3.2. For all $k \in \mathbb{N}$ and $y \in [0, 1]$, $(y - 1)^{2k} y \le \frac{1}{2k+1}$.

In [101, Lemma 2.3], it is shown that, if $C \subseteq \mathbb{R}^n$ is any compact set, $g_i \leq 1$ on C for all i and $f \in \mathbb{R}[X]$ is a polynomial with f > 0 on $S(\bar{g})$, then there exists $\lambda \geq 0$ such that for all sufficiently large $k \in \mathbb{N}$,

$$f - \lambda \sum_{i=1}^{m} (g_i - 1)^{2k} g_i > 0$$
 on C . (3.3.7)

The idea is that, to show $f \in \mathcal{M}(S)$, you first subtract another polynomial from fwhich lies obviously in $\mathcal{M}(S)$ such that the difference can be proved to lie in $\mathcal{M}(S)$ as well. This other polynomial must necessarily be nonnegative on $S(\bar{g})$ but it should take on only very small values on $S(\bar{g})$ so that the difference is still positive on $S(\bar{g})$. On the region where it is outside but not too far away from $S(\bar{g})$, the polynomial you subtract should take large negative values so that the difference gets positive on this region outside of $S(\bar{g})$ (where f itself might be negative). The hope is that the difference satisfies an improved positivity condition which will help us to show that it lies in $\mathcal{M}(S)$. To understand the lemma, it is helpful to observe that the pointwise limit for $k \to \infty$ of this difference, which is the left hand side of (3.3.10), is f on $S(\bar{g})$ and ∞ outside of $S(\bar{g})$. This is the motivation of the following lemma:

Lemma 3.3.3. For all \bar{g} such that $S := S(\bar{g}) \cap [-1, 1]^n \neq \emptyset$ and $g_i \leq 1$ on $[-1, 1]^n$, there are $c_0, c_1, c_2 > 0$ with the following property:

set

For all polynomials $f \in \mathbb{R}[X]$ of degree d with minimum $f^* > 0$ on S, if we

$$L := d^2 n^{d-1} \frac{\|f\|}{f^*}, \qquad \lambda := c_1 d^2 n^{d-1} \|f\| L^{c_2}$$
(3.3.8)

and if $k \in \mathbb{N}$ satisfies

$$2k+1 \ge c_0(1+L^{c_0}), \tag{3.3.9}$$

then the inequality

$$f - \lambda \sum_{i=1}^{m} (g_i - 1)^{2k} g_i \ge \frac{f^*}{2}$$
(3.3.10)

holds on $[-1,1]^n$.

Proof. By the Lojasiewicz inequality for semialgebraic functions (Corollary 2.6.7 in [13]), we can choose $c_2, c_3 > 0$ such that

$$dist(x,S)^{c_2} \le -c_3 \min\{g_1(x), \dots, g_m(x), 0\}$$
(3.3.11)

for all $x \in [-1,1]^n$ where dist(x,S) denotes the Euclidean distance of x to S. Set

$$c_4 := c_3(4n)^{c_2}, \tag{3.3.12}$$

$$c_1 := 4nc_4 \tag{3.3.13}$$

and choose $c_0 \in \mathbb{N}$ big enough to guarantee that

$$c_0(1+r^{c_0}) \ge 2(m-1)c_4r^{c_2}$$
 and (3.3.14)

$$c_0(1+r^{c_0}) \ge 4mc_1 r^{c_2+1} \tag{3.3.15}$$

for all $r \ge 0$. Now suppose $f \in \mathbb{R}[X]$ is of degree d with minimum $f^* > 0$ on S and consider the set

$$A := \left\{ x \in [-1,1]^n \mid f(x) \le \frac{3}{4} f^* \right\}.$$

By Lemma 3.3.1, we get for all $x \in A$ and $y \in S$

$$\frac{f^*}{4} \le f(y) - f(x) \le \|x - y\| d^2 n^{d-1} \sqrt{n} \|f\| \le \|x - y\| d^2 n^d \|f\|.$$

Since this is valid for arbitrary $y \in S$, it holds that

$$\frac{f^*}{4d^2n^d\|f\|} \le \operatorname{dist}(x,S)$$

for all $x \in A$. We combine this now with (3.3.11) and get

$$\min\{g_1(x), \dots, g_m(x)\} \le -\frac{1}{c_3} \left(\frac{f^*}{4d^2n^d ||f||}\right)^{c_2}$$

for $x \in A$. We have omitted the argument 0 in the minimum which is here redundant because of $A \cap S = \emptyset$. By setting

$$\delta := \frac{1}{c_4 L^{c_2}} > 0, \tag{3.3.16}$$

where we define L like in (3.3.8), and having a look at (3.3.12), we can rewrite this as

$$\min\{g_1(x), \dots, g_m(x)\} \le -\delta.$$
(3.3.17)

Define λ and k like in (3.3.8) and (3.3.9). For later use, we note

$$\lambda = c_1 L^{c_2 + 1} f^*. \tag{3.3.18}$$

We claim now that

$$f + \frac{\lambda \delta}{2} \ge \frac{f^*}{2} \text{ on } [-1, 1]^n,$$
 (3.3.19)

$$\frac{\delta}{2} \ge \frac{m-1}{2k+1} \qquad \text{and} \tag{3.3.20}$$

$$\frac{J^*}{4} \ge \frac{\lambda m}{2k+1}.\tag{3.3.21}$$

Let us prove these claims. If we choose in Lemma 3.3.1 for y a minimizer of f on S, we obtain

$$|f(x) - f^*| \le \operatorname{diam}([-1, 1]^n) d^2 n^{d-1} \sqrt{n} ||f|| = 2\sqrt{n} d^2 n^{d-1} \sqrt{n} ||f|| = 2d^2 n^d ||f||$$

for all $x \in [-1,1]^n$, noting that the diameter of $[-1,1]^n$ is $2\sqrt{n}$. In particular, we observe

$$f \ge f^* - 2d^2n^d ||f|| \ge \frac{f^*}{2} - 2d^2n^d ||f||$$
 on $[-1,1]^n$.

Together with the equation

$$\frac{\lambda\delta}{2} = 2d^2n^d \|f\|,$$

which is clear from (3.3.8), (3.3.13) and (3.3.16), this yields (3.3.19). Using (3.3.9), (3.3.14) and (3.3.16), we see that

$$(2k+1)\delta \ge c_0(1+L^{c_0})\delta \ge 2(m-1)c_4L^{c_2}\delta = 2(m-1)$$

which is nothing else than (3.3.20). Finally, we exploit (3.3.9), (3.3.15) and (3.3.18), to see that

$$(2k+1)f^* \ge c_0(1+L^{c_0})f^* \ge 4mc_1L^{c_2+1}f^* = 4m\lambda,$$

i.e., (3.3.21) holds. Now (3.3.19), (3.3.20) and (3.3.21) will enable us to show our claim (3.3.10). If $x \in A$, then in the sum

$$\sum_{i=1}^{m} (g_i(x) - 1)^{2k} g_i(x)$$
(3.3.22)

at most m-1 summands are nonnegative. By Remark 3.3.2, these nonnegative summands add up to at most (m-1)/(2k+1). At least one summand is negative, in fact $\leq -\delta$ by (3.3.17). All in all, if we evaluate the left hand side of inequality (3.3.10) at a point $x \in A$, then we get

$$f - \lambda \sum_{i=1}^{m} (g_i - 1)^{2k} g_i \ge f(x) - \lambda \frac{m-1}{2k+1} + \lambda \delta \ge \underbrace{f(x) + \frac{\lambda \delta}{2}}_{\ge \frac{f^*}{2} \text{ by } (3.3.19)} + \lambda \underbrace{\left(\frac{\delta}{2} - \frac{m-1}{2k+1}\right)}_{\ge 0 \text{ by } (3.3.20)} \ge \frac{f^*}{2}.$$

When we evaluate it in a point $x \in [-1, 1]^n \setminus A$, all summands of the sum (3.3.22) might happen to be nonnegative. Again by Remark 3.3.2, they add up to at most

m/(2k+1). But at the same time, the definition of A gives us a good lower bound on f(x) so that the result is

$$\geq \frac{3}{4}f^* - \lambda \frac{m}{2k+1} \geq \frac{f^*}{2} + \underbrace{\frac{f^*}{4} - \frac{\lambda m}{2k+1}}_{\geq 0 \text{ by } (3.3.21)} \geq \frac{f^*}{2}.$$

Proposition 3.3.1. If $p, q \in \mathbb{R}[X]$ are both homogeneous (i.e., all of their respective monomials have the same degree), then $||pq|| \leq ||p|| ||q||$. For arbitrary $s \in \mathbb{N}$ and polynomials $p_1, \ldots, p_s \in \mathbb{R}[X]$, we have

$$||p_1 \cdots p_s|| \le (1 + \deg p_1) \cdots (1 + \deg p_s) ||p_1|| \cdots ||p_s||.$$

Proof. The statement for homogeneous p and q can be found in [100, Lemma 8]. The second claim follows from this by writing each p_i as a sum $p_i = \sum_k p_{ik}$ of homogeneous degree k polynomials p_{ik} . Multiply the p_i by distributing out all such sums and apply the triangle inequality to the sum which arises in this way. Then use

$$||p_{1k_1}\cdots p_{sk_s}|| \le ||p_{1k_1}||\cdots ||p_{sk_s}|| \le ||p_1||\cdots ||p_s||.$$

Now factor out $||p_1|| \cdots ||p_s||$ and recombine the terms of the sum which now are all equal to 1.

Lemma 3.3.4. For all $c_1, c_2, c_3 > 0$, there is c > 0 such that

$$c_1 \exp(c_2 r^{c_3}) \le c \exp(r^c)$$
 for all $r \ge 0$.

Proof. Choose any $c \ge c_1 \exp(c_2 2^{c_3})$ such that $c_3 \le c/2$ and $c_2 \le 2^{c/2}$. Then for $r \in [0, 2]$,

$$c_1 \exp(c_2 r^{c_3}) \le c_1 \exp(c_2 2^{c_3}) \le c \le c \exp(r^c)$$

and for $r \ge 2$ (observing that $c_1 \le c$), $c_1 \exp(c_2 r^{c_3}) \le c \exp(2^{c/2} r^{c/2}) \le c \exp(r^c)$. \Box

We resume the discussion before Lemma 3.3.3. With regard to (3.3.10), we can for the moment concentrate on polynomials positive on the hypercube $[-1, 1]^n$. If this hypercube could be described by a single polynomial inequality, i.e., if we had $[-1, 1]^n = S(p)$ for some $p \in \mathbb{R}[X]$, then the idea would be to apply the bound for Schmüdgen's Positivstellensatz now. The clue is here that p is a single polynomial and hence preordering and quadratic module representations are the same, i.e., $\mathcal{P}(p) = \mathcal{M}(p)$. The following lemma works around the fact that $[-1, 1]^n = S(p)$ can only happen when n = 1. We round the edges of the hypercube.

Lemma 3.3.5. Let $S \subseteq (-1,1)^n$ be compact. Then $1 - \frac{1}{d} - (X_1^{2d} + \dots + X_n^{2d}) > 0$ on S for all sufficiently large $d \in \mathbb{N}$.

Proof. Consider for each $1 \leq d \in \mathbb{N}$ the set

$$A_d := \left\{ x \in S \mid x_1^{2d} + \dots + x_n^{2d} \ge 1 - \frac{1}{d} \right\}.$$

This gives a decreasing sequence $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots$ of compact sets whose intersection $\cap_{d=1}^{\infty} A_d$ is empty. By compactness, a finite subintersection is empty, i.e., $A_d = \emptyset$ for all large $d \in \mathbb{N}$.

Finally, we are ready to give the proof of Theorem 3.2.4.

Proof of Theorem 3.2.4. By a simple scaling argument, we may assume that $||g_i|| \leq 1$ and $g_i \leq 1$ on [-1, 1] for all *i*. According to Lemma 3.3.5, we can choose $d_0 \in \mathbb{N}$ such that

$$p := 1 - \frac{1}{d_0} - (X_1^{2d} + \dots + X_n^{2d}) > 0 \text{ on } S(\bar{g}).$$

By Putinar's Theorem 3.2.2, we have $p \in \mathcal{M}(S)$ and therefore

$$p \in \mathcal{M}(S)_{d_1} \tag{3.3.23}$$

for some $d_1 \in \mathbb{N}$. Choose $d_2 \in \mathbb{N}$ such that

$$1 + \deg g_i \le d_2$$
 for all $i \in \{1, \dots, m\}$. (3.3.24)

Now we choose c_0, c_1, c_2 as in Lemma 3.3.3, define L and λ as in (3.3.8) and choose the smallest $k \in \mathbb{N}$ satisfying (3.3.9). Then

$$2k+1 \le c_0(1+L^{c_0})+2. \tag{3.3.25}$$

Let $c_3 \ge 1$ denote the constant existing by Theorem 3.2.3 (which is called *c* there and gives the bound for preordering representations of polynomials positive on $S(\bar{g})$). Using Lemma 3.3.4, it is easy to see that we can choose $c_4, c_5, c_6, c_7, c \ge 0$ satisfying

$$c_3 2^{c_3} r^{2+2c_3} n^{c_3 r} \le c_4(\exp(c_4 r)) \tag{3.3.26}$$

$$2r + 2c_1 r^{c_2+1} d_2^{r(1+r^{c_0})+1} \le c_5 \exp(r^{c_5})$$
(3.3.27)

$$c_4 \exp(2c_4 d_2 r(1 + r^{c_0} + 3)) \le c_6 \exp(r^{c_6})$$
(3.3.28)

$$c_5^{c_3}c_6\exp(c_3r^{c_5}+r^{c_6}) \le c_7\exp(r^{c_7}) \tag{3.3.29}$$

$$c_7 \exp(r^{c_7}) + d_1 \le c \exp(r^c) \tag{3.3.30}$$

for all $r \ge 0$. Now let $f \in \mathbb{R}[X]$ be a polynomial of degree $d \ge 1$ with

$$f^* := \min\{f(x) \mid x \in S(\bar{g})\} > 0.$$

We are going to apply Theorem 3.2.3 to

$$h := f - \lambda \sum_{i=1}^{m} (g_i - 1)^{2k} g_i.$$

By Lemma 3.3.3, (3.3.10) holds for this polynomial, in particular

$$h^* := \min\{h(x) \mid x \in S(p)\} \ge \frac{f^*}{2}.$$
 (3.3.31)

By Proposition 3.3.1 and the definition of d_2 in (3.3.24),

$$\|h\| \le \|f\| + \lambda d_2^{2k+1} \tag{3.3.32}$$

$$\deg h \le \max\{d, (2k+1)d_2, 1\} =: d_h.$$
(3.3.33)

By Theorem 3.2.3 (respectively the above choice of $c_3 \ge 1$), we get

$$h \in \mathcal{P}(\{p\})_{k_h}$$
 where $k_h := c_3 d_h^2 \left(1 + d_h^2 n^{d_h} \frac{\|h\|}{h^*} \right)^{c_3}$. (3.3.34)

Note that $||h||/h^* \ge 1$ since $0 < h^* \le h(0) \le ||h||$. We use this to simplify the degree bound in (3.3.34). Obviously

$$k_{h} \leq c_{3}d_{h}^{2} \left(2d_{h}^{2}n^{d_{h}}\frac{\|h\|}{h^{*}}\right)^{c_{3}} \leq c_{3}2^{c_{3}}d_{h}^{2+2c_{3}}n^{c_{3}d_{h}} \left(\frac{\|h\|}{h^{*}}\right)^{c_{3}} \leq c_{4}\exp(c_{4}d_{h}) \left(\frac{\|h\|}{h^{*}}\right)^{c_{3}}$$

$$(3.3.35)$$

by choice of c_4 in (3.3.26). Moreover, we have

$$\frac{\|h\|}{h^*} \le \frac{2}{f^*} (\|f\| + \lambda d_2^{2k+1}) = 2 \frac{\|f\|}{f^*} + 2c_1 d_2^{2k+1} L^{c_2+1}$$
$$\le 2L + 2c_1 d_2^{2k+1} L^{c_2+1} = 2L + 2c_1 L^{c_2+1} d_2^{c_0(1+L^{c_0})+1} \le c_5 \exp(L^{c_5}) \quad (3.3.36)$$

by (3.3.32), (3.3.31), (3.3.25), (3.3.18) and by the choice of c_5 in (3.3.27). It follows

that

$$\begin{aligned} d_h &\leq d(2k+2)d_2 & \text{(by (3.3.33))} \\ &\leq d(c_0(1+L^{c_0})+3)d_2 & \text{(by (3.3.25))} \\ &\leq 2d_2d^2n^d\frac{\|f\|}{2dn^d\|f\|}(c_0(1+L^{c_0})+3) & \\ &\leq 2d_2d^2n^d\frac{\|f\|}{f^*}(c_0(1+L^{c_0})+3) & \text{(by Lemma 3.1.1)} \\ &\leq 2d_2nL(c_0(1+(nL)^{c_0}+3)) & \text{(by (3.3.8))} \end{aligned}$$

and therefore

$$c_4 \exp(c_4 d_h) \le c_6 \exp((nL)^{c_6})$$
 (3.3.37)

for the constant c_6 chosen in (3.3.28). We now get

$$k_{h} \leq c_{4} \exp(c_{4}d_{h}) \left(\frac{\|h\|}{h^{*}}\right)^{c_{3}}$$
 (by (3.3.35))
$$\leq c_{6} \exp((nL)^{c_{6}})(c_{5} \exp(L^{c_{5}}))^{c_{3}}$$
 (by (3.3.37) and (3.3.36))
$$= c_{5}^{c_{3}}c_{6} \exp(c_{3}(nL)^{c_{5}} + (nL)^{c_{6}})$$

$$\leq c_{7} \exp((nL)^{c_{7}})$$
 (by choice of c_{7} in (3.3.29)).

Combining this with (3.3.34) and (3.3.23), i.e.,

$$h \in \mathcal{P}(\{p\})_{c_7 \exp((nL)^{c_7})}$$
 and $p \in \mathcal{M}(S)_{d_1}$,

yields (by composing corresponding representations)

$$h \in \mathcal{M}(S)_{c \exp((nL)^c)}$$

according to the choice of c in (3.3.30). Finally, we have that

$$f = h + \lambda \sum_{i=1}^{m} (g_i - 1)^{2k} g_i \in \mathcal{M}(S)_{c \exp((nL)^c)}$$

since

$$\deg((g_i - 1)^{2k} g_i) \le d_h \le k_h \le c_7 \exp((nL)^{c_7}) \le c \exp((nL)^c)$$

by choice of d_2 in (3.3.24), d_h in (3.3.33), k_h in (3.3.34) and c in (3.3.30).

Chapter 4

SOS Methods Based on the Gradient Ideal

As we have seen in Chapter 1, a very good lower bound f_{sos}^* of the polynomial f(x) can be found by applying SOS relaxations. But sometimes the SOS relaxation may not be exact, i.e., $f_{sos}^* < f^* = \min_{x \in \mathbb{R}^n} f(x)$. In such situations, how can we improve the quality of the lower bound by applying some appropriately modified SOS relaxations?

This chapter will introduce a new method to get a sequence of better lower bounds $\{f_{N,grad}^*\}_{N=1}^{\infty}$. Every lower bound $f_{N,grad}^*$ is better than f_{sos}^* . The method combines the SOS relaxation and gradient of f(x). It has the nice property that $\lim_{N\to\infty} f_{N,grad}^* = f^*$ whenever the minimum f^* is attainable. Furthermore, the method also has finite convergence under some generic conditions, i.e., with probability one. A full version of this chapter is in [71].

4.1 Introduction

In this chapter, we consider the unconstrained polynomial optimization problem

$$f^* = \min_{x \in \mathbb{R}^n} f(x) \tag{4.1.1}$$

where $x \in \mathbb{R}^n$ and f(x) is a real multivariate polynomial of degree d. As is wellknown, the optimization problem (4.1.1) is NP-hard even when d is fixed to be four [67]. A lower bound can be computed efficiently using the Sum Of Squares (SOS) relaxation

$$f_{sos}^* = \text{maximize } \gamma \quad \text{subject to} \quad f(x) - \gamma \succeq_{sos} 0,$$
 (4.1.2)

where the inequality $g \succeq_{sos} 0$ means that the polynomial g is SOS, i.e. a sum of squares of other polynomials. See Section 2.1 for an elementary introduction to SOS polynomials. The relationship between (4.1.1) and (4.1.2) is as follows: $f_{sos}^* \leq f^*$ and the equality holds if and only if $f(x) - f^*$ is SOS.

Blekherman [12] recently showed that, for fixed even degree $d \ge 4$, the ratio between the volume of all nonnegative polynomials and the volume of all SOS polynomials tends to infinity when n goes to infinity. In other words, for large n, there are many more nonnegative polynomials than SOS polynomials. For dealing with the challenging case when $f_{sos}^* < f^*$, Lasserre [52] proposed finding a sequence of lower bounds for f(x) in some large ball $\{x \in \mathbb{R}^n : ||x||_2 \le R\}$. His approach is based on the result [4] that SOS polynomials of all possible degrees are dense among polynomials which are nonnegative on some compact set. This sequence converges to f^* when the degrees of the polynomials introduced in the algorithm go to infinity. But it may not converge in finitely many steps, and the degrees of the required auxiliary polynomials can be very large. In this chapter, we introduce a method which can find the global minimum and terminate in finitely many steps, under some weak assumptions. Our point of departure is the observation that all local minima and global minima of (4.1.1) occur at points in the *real gradient variety*

$$V_{grad}^{\mathbb{R}}(f) = \{ u \in \mathbb{R}^n : \ (\nabla f)(u) = 0 \}.$$
(4.1.3)

The gradient ideal of f is the ideal in $\mathbb{R}[x_1, \ldots, x_n]$ generated by all partial derivatives of f:

$$\mathcal{I}_{grad}(f) = \langle \nabla f(x) \rangle = \langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n} \rangle.$$
(4.1.4)

There are several recent references on minimizing polynomials by way of the gradients. Hanzon and Jibetean [39] suggest applying perturbations to f to produce a sequence of polynomials f_{λ} (for small λ) with the property that the gradient variety of f_{λ} is finite and the minima f_{λ}^* converge to f^* as λ goes to 0. Laurent [55] and Parrilo [89] discuss the more general problem of minimizing a polynomial subject to polynomial equality constraints (not necessarily partial derivatives). Under the assumption that the variety defined by the equations is finite, the matrix method proposed in [55] has finite convergence even if the ideal generated by the constraints is not radical. Building on [39, 55], Jibetean and Laurent [45] propose to compute f^* by solving a single SDP, provided the gradient variety is finite (radicalness is not necessary).

The approach of this chapter is to find a lower bound f_{grad}^* for (4.1.1) by requiring $f - f_{grad}^*$ to be SOS in the quotient ring $\mathbb{R}[X]/\mathcal{I}_{grad}(f)$ instead of in $\mathbb{R}[X]$. Let $\mathbb{R}[X]_m$ denote the vector space of polynomials with degree up to m. We consider the SOS relaxation

$$f_{N,grad}^* = \sup \quad \gamma \tag{4.1.5}$$

s.t.
$$f(x) - \gamma - \sum_{j=1}^{n} \phi_j(x) \frac{\partial f}{\partial x_j} \in \sum \mathbb{R}[X]^2$$
 (4.1.6)

$$\phi_j(x) \in \mathbb{R}[X]_{2N-d+1}.\tag{4.1.7}$$

Here $d = \deg(f)$, N is an fixed integer, and $\sum \mathbb{R}[X]^2$ denotes the cone of SOS polynomials. Obviously $f_{N,grad}^* \leq f^*$ for all $2N \geq d$, provided f^* is attained at one point. The lower bound $f_{N,grad}^*$ is monotonically increasing as N increases, since the feasible domain of (4.1.5) is also increasing, i.e.,

$$\cdots f_{N,grad}^* \le f_{N+1,grad}^* \le \cdots \le f^*.$$
 (4.1.8)

The convergence of these lower bounds is summarized in the following theorem.

Theorem 4.1.1. Let f(x) be a polynomial in n real variables which attains its infimum f^* over \mathbb{R}^n . Then $\lim_{N\to\infty} f^*_{N,grad} = f^*$. Furthermore, if the gradient ideal $\mathcal{I}_{grad}(f)$ is radical, then f^* is attainable, i.e., there exists an integer N such that $f^*_{N,grad} = f^*_{grad} = f^*$.

The proof of this theorem will be given in Section 4.3.

4.2 Polynomials over their gradient varieties

Consider a polynomial $f \in \mathbb{R}[X]$ and its gradient ideal $\mathcal{I}_{grad}(f)$ as in (4.1.4). A natural idea in solving (4.1.1) is to apply Theorem 2.3.4 to the ideal $I = \mathcal{I}_{grad}(f)$, since the minimum of f over \mathbb{R}^n will be attained at a subset of $V^{\mathbb{R}}(I)$ if it is attained at all. However, the hypothesis of Theorem 2.3.4 requires that I be zerodimensional, which means that the complex variety $V_{grad}(f) = V(I)$ of all critical points must be finite. Our results in this section remove this restrictive hypothesis. We shall prove that every nonnegative f is SOS in $\mathbb{R}[X]/I$ as long as the gradient ideal $I = \mathcal{I}_{grad}(f)$ is radical.

Theorem 4.2.1. Assume that the gradient ideal $\mathcal{I}_{grad}(f)$ is radical. If the real polynomial f(x) is nonnegative over $V_{grad}^{\mathbb{R}}(f)$, then there exist real polynomials $q_i(x)$ and $\phi_j(x)$ so that

$$f(x) = \sum_{i=1}^{s} q_i(x)^2 + \sum_{j=1}^{n} \phi_j(x) \frac{\partial f}{\partial x_j}.$$
 (4.2.9)

The proof of this theorem will be based on the following two lemmas. The first is a generalization of the Lagrange Interpolation Theorem from sets of points to disjoint varieties.

Lemma 4.2.2. Let V_1, \ldots, V_r be pairwise disjoint varieties in \mathbb{C}^n . Then there exist polynomials $p_1, \ldots, p_r \in \mathbb{R}[X]$ such that $p_i(V_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta function.

Proof. Our definition of variety requires that each V_j is actually defined by polynomials with real coefficients. If $I_j = I(V_j)$ is the radical ideal of V_j then we have $V_j = V(I_j)$. Fix an index j and let W_j denote the union of the varieties $V_1, \ldots, V_{i-1}, V_{i+1}, \ldots, V_r$. Then

$$I(W_j) = I_1 \cap \cdots \cap I_{j-1} \cap I_{j+1} \cap \cdots \cap I_r.$$

Our hypothesis implies that $V_j \cap W_j = \emptyset$. By Hilbert's Weak Nullstellensatz (Theorem 2.2.2), there exist polynomials $p_j \in I(W_j)$ and $q_j \in I_j$ such that $p_j + q_j = 1$. This identity shows that $p_j(V_j) = 1$ and $p_j(V_k) = 0$ for $k \neq j$. Hence the *r* polynomials p_1, \ldots, p_r have the desired properties. \Box

Now consider the behavior of the polynomial f(x) over its gradient variety $V_{grad}(f)$. We make use of the fact that $V_{grad}(f)$ is a finite union of irreducible subvarieties ([13, §2]).

Lemma 4.2.3. Let W be an irreducible subvariety of $V_{grad}(f)$ and suppose that W contains at least one real point. Then f(x) is constant on W.

Proof. If we replace our polynomial ring $\mathbb{R}[X]$ by $\mathbb{C}[X]$ then W either remains irreducible or it becomes a union of two irreducible components $W = W_1 \cup W_2$ which are exchanged under complex conjugation. Let us first consider the case when W is irreducible in the Zariski topology induced by $\mathbb{C}[X]$. W is connected in \mathbb{C}^n (see [104]). Any two points in a connected algebraic variety in \mathbb{C}^n can be connected by an algebraic curve. This curve may be singular, but it is a projection of some non-singular curve. Let x, y be two arbitrary points in W. Hence there exists a smooth path $\varphi(t)$ ($0 \le t \le 1$) lying inside W such that $x = \varphi(0)$ and $y = \varphi(1)$. By the Mean Value Theorem of Calculus, it holds that for some $t^* \in (0, 1)$

$$f(y) - f(x) = \nabla f(\varphi(t^*))^T \varphi'(t^*) = 0,$$

since ∇f vanishes on W. We conclude that f(x) = f(y), and hence f is constant on W.

Now consider the case when $W = W_1 \cup W_2$ where W_1 and W_2 are exchanged by complex conjugation. We had assumed that W contains a real point p. Since pis fixed under complex conjugation, $p \in W_1 \cap W_2$. By the same argument as above, f(x) = f(p) for all $x \in W$. \Box Proof of Theorem 4.2.1. Consider the irreducible decomposition of $V_{grad}(f)$. We group together all components which have no real point and all components on which f takes the same real value. Hence the gradient variety has a decomposition

$$V_{grad}(f) = W_0 \cup W_1 \cup W_2 \cup \dots \cup W_r, \qquad (4.2.10)$$

such that W_0 has no real point and f is a real constant on each other variety W_i , say,

$$f(W_1) > f(W_2) > \cdots > f(W_r) \ge 0.$$

The varieties W_i are pairwise disjoint, so by Lemma 4.2.2 there exist polynomials $p_i \in \mathbb{R}[X]$ such that $p_i(W_j) = \delta_{ij}$. By Theorem 2.3.5, there exists a sum of squares $sos(x) \in \mathbb{R}[X]$ such that f(x) = sos(x) for all $x \in W_0$. Using the non-negative real numbers $\alpha_j := \sqrt{f(W_j)}$, we define

$$q(x) = sos(x) \cdot p_0^2(x) + \sum_{i=1}^r (\alpha_i \cdot p_i(x))^2.$$
 (4.2.11)

By construction, f(x) - q(x) vanishes on the gradient variety $V_{grad}(f)$. The gradient ideal $\mathcal{I}_{grad}(f)$ was assumed to be radical. Using Hilbert's Strong Nullstellensatz (Theorem 2.2.3), we conclude that f(x) - q(x) lies in $\mathcal{I}_{grad}(f)$. Hence the desired representation (4.2.9) exists. \Box

In Theorem 4.2.1, the assumption that $\mathcal{I}_{grad}(f)$ is radical cannot be removed. This is shown by the following counterexample.

Example 4.2.4. Let n = 3 and consider the polynomial

$$f(x, y, z) = x^8 + y^8 + z^8 + M(x, y, z)$$

where $M(x, y, z) = x^4 y^2 + x^2 y^4 + z^6 - 3x^2 y^2 z^2$ is the Motzkin polynomial, which is is non-negative but not a sum of squares in $\mathbb{R}[X]/\mathcal{I}_{grad}(f)$ (see [71]). In cases (like Example 4.2.4) when the gradient ideal is not radical, the following still holds.

Theorem 4.2.5. Let $f(x) \in \mathbb{R}[X]$ be a polynomial which is strictly positive on its real gradient variety $V_{qrad}^{\mathbb{R}}(f)$, Then f(x) is SOS modulo its gradient ideal $\mathcal{I}_{grad}(f)$.

Proof. We retain the notation from the proof of Theorem 4.2.1. Consider the decomposition of the gradient variety in (4.2.10). Each W_i is the union of several irreducible components. Consider a primary decomposition of the ideal $\mathcal{I}_{grad}(f)$, and define J_i to be the intersection of all primary ideals in that decomposition whose variety is contained in W_i . Then we have $\mathcal{I}_{grad}(f) = J_0 \cap J_1 \cap \cdots \cap J_r$, where $W_i = V(J_i)$ and, since the W_i are pairwise disjoint, we have $J_i + J_k = \mathbb{R}[X]$ for $i \neq k$. The Chinese Remainder Theorem [29] implies

$$\mathbb{R}[X]/\mathcal{I}_{grad}(f) \simeq \mathbb{R}[X]/J_0 \times \mathbb{R}[X]/J_1 \times \cdots \times \mathbb{R}[X]/J_r.$$
(4.2.12)

Here $V^{\mathbb{R}}(J_0) = \emptyset$. Hence, by Theorem 2.3.5, there exists a sum of squares $sos(x) \in \mathbb{R}[X]$ such that $f(x) - sos(x) \in J_0$. By assumption, $\alpha_i^2 = f(W_i)$ is strictly positive for all $i \geq 1$. The polynomial $f(x)/\alpha_i^2 - 1$ vanishes on W_i . By Hilbert's Strong Nullstellensatz, there exists an integer m > 0 such that $(f(x)/\alpha_i^2 - 1)^m$ is in the ideal J_i . We construct a square root of $f(x)/\alpha_i^2$ in the residue ring $\mathbb{R}[X]/J_i$ using the familiar Taylor series expansion for the square root function:

$$(1 + (f(x)/\alpha_i^2 - 1))^{1/2} = \sum_{k=0}^{m-1} {\binom{1/2}{k}} (f(x)/\alpha_i^2 - 1)^k \mod J_i$$

Multiplying this polynomial by α_i , we get a polynomial $q_i(x)$ such that $f(x) - q_i^2(x)$ is in the ideal J_i . We have shown that f(x) maps to the vector $(sos(x), q_1(x)^2, q_2(x)^2, \dots, q_r(x)^2)$ under the isomorphism (4.2.12). That vector is clearly a sum of squares in the ring on the right hand side of (4.2.12). We conclude that f(x) is a sum of squares in $\mathbb{R}[X]/\mathcal{I}_{grad}(f)$. \Box

Example 4.2.6. Let f be the polynomial in Example 4.2.4 and let ϵ be any positive constant. Theorem 4.2.5 says that $f + \epsilon$ is SOS modulo $\mathcal{I}_{grad}(f)$. Such a representation can be found by symbolic computation as follows. Primary decomposition over $\mathbb{Q}[x, y, z]$ yields

$$\mathcal{I}_{grad}(f) = J_0 \cap J_1,$$

where $V^{\mathbb{R}}(J_0) = \emptyset$ and and $\sqrt{J_1} = \langle x, y, z \rangle$. The ideal J_1 has multiplicity 153, and it contains the square f^2 of our given polynomial. The ideal J_0 has multiplicity 190. Its variety $V(J_0)$ consists of 158 distinct points in \mathbb{C}^3 . By elimination, we can reduce to the univariate case. Using the algorithm of [8, 9] for real radicals in $\mathbb{Q}[z]$, we find a sum of squares $sos(z) \in \mathbb{Q}[z]$ such that $f - sos(z) \in J_0$. Running Buchberger's algorithm for $J_0 + J_1 = \langle 1 \rangle$, we get polynomials $p_0 \in J_0$ and $p_1 \in J_1$ such that $p_0 + p_1 = 1$. The following polynomial is a sum of squares,

$$p_1^2 \cdot (sos(z) + \epsilon) + p_0^2 \cdot \epsilon \cdot (1 + \frac{1}{2\epsilon}f)^2,$$
 (4.2.13)

and it is congruent to $f(x, y, z) + \epsilon$ modulo $\mathcal{I}_{grad}(f) = J_0 \cap J_1 = J_0 \cdot J_1$. Note that the coefficients of the right hand polynomial in the SOS representation (4.2.13) tend to infinity as ϵ approaches zero. This is consistent with the conclusion of Example 4.2.4.

4.3 Convergence analysis and the algorithm

We are now ready to give the proof of Theorem 4.1.1.

Proof of Theorem 4.1.1. Since f(x) attains its infimum, the global minima of f(x)must occur on the real gradient variety $V_{grad}^{\mathbb{R}}(f)$. It is obvious that any real number γ which satisfies the SOS constraint in (4.1.5) is a lower bound of f(x), and we have the sequence of inequalities in (4.1.8). Consider an arbitrary small real number $\varepsilon > 0$. The polynomial $f(x) - f^* + \varepsilon$ is strictly positive on its real gradient variety $V_{grad}^{\mathbb{R}}(f)$. By Theorem 4.2.5, $f(x) - f^* + \varepsilon$ is SOS modulo $\mathcal{I}_{grad}(f)$. Hence there exists an integer $N(\epsilon)$ such that

$$f_{N,qrad}^* \ge f^* - \varepsilon$$
 for all $N \ge N(\epsilon)$.

Since the sequence $\{f_{N,grad}^*\}$ is monotonically increasing, it follows that $\lim_{N\to\infty} f_{N,grad}^* = f^*$.

Now suppose $\mathcal{I}_{grad}(f) = \mathcal{I}_{grad}(f - f^*)$ is a radical ideal. The nonnegative polynomial $f(x) - f^*$ is SOS modulo $\mathcal{I}_{grad}(f)$ by Theorem 4.2.1. Hence $f^*_{N,grad} = f^*$ for some $N \in \mathbb{Z}_{>0}$.

Remark 4.3.1. (i) The condition that f(x) attains its infimum cannot be removed. Otherwise the infimum f_{grad}^* of f(x) on $V_{grad}^{\mathbb{R}}(f)$ need not be a lower bound for f(x) on \mathbb{R}^n . A counterexample is $f(x) = x^3$. Obviously f(x) has infimum $f^* = -\infty$ on \mathbb{R}^1 . However, $f_{grad}^* = f_{grad,N}^* = 0$ for all $N \ge 1$ because $f(x) = (\frac{x}{3})f'(x)$ is in the gradient ideal $\mathcal{I}_{grad}(f) = \langle f'(x) \rangle$.

(ii) It is also not always the case that $f_{grad}^* = f^*$ when f^* is finite. Consider the bivariate polynomial $f(x, y) = x^2 + (1 - xy)^2$. We can see that $f^* = 0$ is not attained, but $f_{arad}^* = 1 > f^*$.

(iii) If f(x) attains its infimum but $\mathcal{I}_{grad}(f)$ is not radical, we have only that $\lim_{N \to \infty} f^*_{N,grad} = f^*$. But there is typically no integer N with $f^*_{N,grad} = f^*$, as shown in Example 4.2.4. In the rest of this section, we discuss the duality of problem (4.1.5), and show how to extract the global minimizer(s) (if any). Given any *multi-indexed* vector $y = (y_{\alpha})$, i.e., indexed by integer vectors $\alpha \in \mathbb{N}^n$, its *moment matrix* $M_N(y)$ to defined to be (see Section 2.3)

$$M_N(y) = (y_{\alpha+\beta})_{0 \le |\alpha|, |\beta| \le N}.$$

For polynomial $p(x) = \sum_{\beta} p_{\beta} x^{\beta}$, define the linear map $y \mapsto p * y$ such that the monomial-indexed vector p * y has coordinates $(p * y)_{\alpha} = \sum_{\beta} p_{\beta} y_{\alpha+\beta}$. Denote by f the vector of coefficients of f(x). Let f_i denote the vector of coefficients of the *i*-th partial derivative $\frac{\partial f}{\partial x_i}$. We rewrite (4.1.5) as follows:

$$f_{N,grad}^{*} = \max_{\substack{\gamma \in \mathbb{R}, \sigma \in \mathbb{R}[X]_{2N} \\ \phi_{j}(x) \in \mathbb{R}[X]_{2N-d+1}}} \gamma \quad \text{subject to} \quad \sigma(x) \succeq_{sos} 0 \quad (4.3.14)$$

and
$$f(x) - \gamma = \sigma(x) + \sum_{j=1}^{n} \phi_{j}(x) \frac{\partial f}{\partial x_{j}}.$$

$$(4.3.15)$$

The dual of above problem is the following (see also[52, 55])

$$f_{N,mom}^* = \min_{y} f^T y \tag{4.3.16}$$

s.t.
$$M_{N-d/2}(f_i * y) = 0, \ i = 1, \cdots, n$$
 (4.3.17)

$$M_N(y) \succeq 0, \ y_0 = 1.$$
 (4.3.18)

The following theorem relates the primal and dual objective function values $f_{N,mom}^*$ and $f_{N,grad}^*$, and it shows how to extract a point x^* in \mathbb{R}^n at which the minimum of f(x) is attained.

Theorem 4.3.2. Assume f(x) attains its infimum f^* over \mathbb{R}^n (hence d is even). Then we have:

- (i) $f_{N,mom}^* \ge f_{N,grad}^*$ and hence $\lim_{N \to \infty} f_{N,mom}^* = f^*$.
- (ii) Suppose $f_{N,grad}^* = f^*$ for some N. If $x^* \in \mathbb{R}^n$ minimizes f(x), then $y^* = mon_{2N}(x^*) \in \mathbb{R}^{\binom{n+2N}{2N}}$ solves the primal SDP.
- (iii) If y is a solution to the primal problem with $rank(M_N(y)) = 1$, then factoring $M_N(y)$ as column vector times row vector yields one global minimizer x^* of the polynomial f(x).
- (iv) Suppose that $f_{N,grad}^* = f^*$ and $\sigma(x) = \sum_{j=1}^{\ell} (q_j(x))^2$ solves the dual SDP. Then the set of all global minima of f(x) equals the set of solutions $x \in \mathbb{R}^n$ to the following equations:

$$q_j(x) = 0, \quad j = 1, \dots, \ell$$
$$\frac{\partial f(x)}{\partial x_i} = 0, \quad i = 1, \dots, n.$$

Proof. Parts (i) and (ii) are basically a direct application of Theorem 4.2 in [52]. The hypotheses of that theorem are verified by an "epsilon argument" and applying our Theorem 4.2.5. Let us prove part (iii). Since the moment matrix $M_N(y)$ has rank one, there exists a vector $x^* \in \mathbb{R}^n$ such that $y = mon_N(x^*)$. The strong duality result in (i) implies that

$$f(x^*) = f^T y = f^*_{N,mom} = f^*_{N,grad}.$$

Since $f_{N,grad}^*$ is a lower bound for f(x), we conclude that this lower bound is attained at the point x^* . Therefore, $f_{N,grad}^* = f^*$ and x^* is a global minimizer. Part (iv) is straightforward. \Box

From Theorem 4.3.2 (ii), we can see that there exists one optimal solution y^* to the primal SDP such that $rank(M_N(y^*)) = 1$ if $f^*_{N,qrad} = f^*$ for some integer

N. However, interior-point solvers for SDP will find a solution with moment matrix of maximum rank. So, if there are several global minimizers, the moment matrix $M_N(y^*)$ at relaxation N for which the global minimum is reached, will have rank > 1. However, if some *flat extension* condition holds at order N, i.e.,

$$\operatorname{rank} M_N(y^*) = \operatorname{rank} M_{N-d/2}(y^*) = r \tag{4.3.19}$$

where y^* is one optimal solution to the dual problem, we still can extract minimizers. The rank condition (4.3.19) can be verified very accurately by Singular Value Decomposition (SVD). Then as a consequence of Theorem 1.6 in [24], there exist rvectors $x^*(1), \dots, x^*(r) \in \mathbb{R}^n$ such that

$$M_N(y^*) = \sum_{j=1}^r \nu_j \ mon_N(x^*(j)) \cdot mon_N(x^*(j))^T$$

where $\sum_{j=1}^{r} \nu_j = 1$ and $\nu_j > 0$ for all $j = 1, \dots, r$. Henrion and Lasserre [41] proposed a detailed algorithm to find all such vectors $x^*(j)$. The condition (4.3.19) can be satisfied for some N when $V_{grad}(f)$ is finite; see [55] for a proof. We refer to [41] and [71] for more details about extracting minimizers.

Summarizing the discussion above, we get the following algorithm for minimizing polynomials globally.

Algorithm 4.3.3. Computing the global minimizer(s) (if any) of a polynomial.

Input: A polynomial f(x) of even degree d in n variables $x = (x_1, \ldots, x_n)$.

Output: Global minimizers $x^*(1), \dots, x^*(r) \in \mathbb{R}^n$ of f(x) for some $r \ge 1$.

Algorithm: Initialize N = d/2.

- Step 1 Solve the pair of primal SDP (4.3.14)-(4.3.15) and dual SDP (4.3.16)-(4.3.18).
- Step 2 Check rank condition (4.3.19). If it is satisfied, extract r solutions $x^*(1), \dots, x^*(r)$ by using the method in [41], where r is the rank of $M_N(y^*)$, and then stop.

Step 3 If (4.3.19) is not satisfied, N = N + 1 and then go to Step 1.

As we pointed out after (4.3.19) ([55]), this algorithm will terminate if $V_{grad}(f)$ is finite. If $V_{grad}(f)$ is infinite, it is possible to have infinitely many global minimizers and the extraction method in [41] can not be applied generally (it may work sometimes). In such situations we need to solve the equations in (iv) of Theorem 4.3.2 to obtain the minimizers.

4.3.1 What if the gradient ideal $\mathcal{I}_{grad}(f)$ is not radical ?

The lack of radicalness of the gradient ideal $\mathcal{I}_{grad}(f)$ would be an obstacle for our algorithm. Fortunately, this does not happen often in practice because the ideal $\mathcal{I}_{grad}(f)$ is generically radical, as shown by Proposition 4.3.4. It can be proved by standard arguments of algebraic geometry. We omit the proof.

Proposition 4.3.4. For almost all polynomials f in the finite-dimensional vector space $\mathbb{R}[X]_d$, the gradient ideal $\mathcal{I}_{grad}(f)$ is radical and the gradient variety $V_{grad}(f)$ is a finite subset of \mathbb{C}^n .

Proposition 4.3.4 means that, for almost all polynomials f which attain their minimum f^* , Algorithm 4.3.3 will compute the minimum in finitely many steps. An a priori bound for a degree N with $f^*_{N,grad} = f^*$ is given in [55].

Let us now consider the unlucky case when $\mathcal{I}_{grad}(f)$ is not radical. This happened for instance, in Example 4.2.4. In theory, one can replace the gradient ideal $\mathcal{I}_{grad}(f)$ by its radical $\sqrt{\mathcal{I}_{grad}(f)}$ in our SOS optimization problem. This is justified by the following result.

Corollary 4.3.5. If a polynomial f(x) attains its infimum f^* over \mathbb{R}^n then $f(x) - f^*$ is SOS modulo the radical $\sqrt{\mathcal{I}_{grad}(f)}$ of the gradient ideal.

Proof. Consider the decomposition (4.2.10) and form the SOS polynomial q(x) in (4.2.11). Since f(x)-q(x) vanishes on the gradient variety $V(\mathcal{I}_{grad}(f)) = V(\sqrt{\mathcal{I}_{grad}(f)})$, Hilbert's Strong Nullstellensatz implies that $f(x) - q(x) \in \sqrt{\mathcal{I}_{grad}(f)}$. \Box

There are some known algorithms for computing radicals (see e.g. [33, 51]), and they are implemented in various computer algebra systems. But running these algorithms is usually very time-consuming. In practice, replacing $\mathcal{I}_{grad}(f)$ by its radical $\sqrt{\mathcal{I}_{grad}(f)}$ is not a viable option for efficient optimization algorithms. However, if some polynomials in $\sqrt{\mathcal{I}_{grad}(f)} \setminus \mathcal{I}_{grad}(f)$ are known to the user (for instance, from the geometry of the problem at hand), including these polynomials in (4.1.5) will probably speedup convergence of Algorithm 4.3.3.

4.4 Numerical experiments

In this section, we show some numerical examples by implementing Algorithm 4.3.3. Firstly we show examples where Algorithm 4.3.3 provides much better lower bounds than the standard SOS relaxations, which is consistent with Theorem 4.1.1. Secondly, we show that Algorithm 4.3.3 is more computationally efficient than the standard SOS relaxation.

4.4.1 Comparison of lower bounds

The following examples demonstrate the effectiveness of our Algorithm 4.3.3 for a sample of polynomials that have been discussed in the polynomial optimization literature.

Example 4.4.1 (Homogeneous Polynomials). Let f(x) be a homogeneous polynomial. Regardless of whether f(x) is non-negative, we always have $f_{N,grad}^* = 0$ for any $N \ge d/2$. This comes from the identity $f(x) = \frac{1}{d} \cdot \sum_i x_i \frac{\partial f}{\partial x_i}$, which implies that f(x) lies in its gradient ideal $\mathcal{I}_{grad}(f)$. In order to test global non-negativity of a homogeneous polynomial f(x), we can apply Algorithm 4.3.3 to a dehomogenization of f(x), as shown in examples below.

Example 4.4.2. $f(x,y) = x^2y^2(x^2 + y^2 - 1)$. This polynomial is taken from [52]. It has global minimum value $f^* = -1/27 = -0.03703703703703...$ However, $f^*_{sos} = -\infty$ is considerably smaller than f^* . If we minimize f(x) over its gradient ideal with N = 4, then we get $f^*_{4,grad} = -0.03703703703706212$. The difference equals $f^* - f^*_{4,grad} \approx 2.50 \cdot 10^{-11}$. The solutions extracted by GloptiPoly ([41]) are $(\pm 0.5774, \pm 0.5774)$.

Example 4.4.3. The polynomial $f(x,y) = x^4y^2 + x^2y^4 + 1 - 3x^2y^2$ is obtained from the Motzkin polynomial by substituting z = 1 as in [81]. We have $f^* = 0 >$ $f^*_{sos} = -\infty$. However, if we minimize f(x,y) over its gradient ideal with N =4, we get $f^*_{4,grad} = -6.1463 \cdot 10^{-10}$. The solutions extracted by GloptiPoly are $(\pm 1.0000, \pm 1.0000)$.

Example 4.4.4. The polynomial $f(x, y) = x^4 + x^2 + z^6 - 3x^2z^2$ is obtained from the Motzkin polynomial by substituting y = 1. Now, $f^* = 0 > f^*_{sos} = -729/4096$. However, if we minimize f(x, z) over its gradient ideal with N = 4, we get $f^*_{4,grad} =$ $-9.5415 \cdot 10^{-12}$. The solutions extracted by GloptiPoly are (0.0000, 0.0000) and (±1.0000, ±1.0000).

4.4.2 Comparison of computational efficiency

We test the efficiency of Algorithm 4.3.3 on the Parrilo-Sturmfels family of polynomials of the form

$$f(x_1, \cdots, x_n) = x_1^d + \cdots + x_n^d + g(x_1, \cdots, x_n),$$

where $g \in \mathbb{R}[X]$ is a random polynomial of degree $\leq d - 1$ whose coefficients are uniformly distributed between -K and K, for a fixed positive integer K. This family of polynomials was considered in [80] where it was shown experimentally that the SOS formulation (4.1.2) almost always yields the global minimum. Without loss of generality, we can set K = 1, because any f(x) in the above form can be scaled to have coefficients between -1 and 1 by taking

$$f_s(x_1, \cdots, x_n) = \alpha^{-d} \cdot f(\alpha x_1, \cdots, \alpha x_n)$$

for some properly chosen α . As observed in [80], this scaling will greatly increase the stability and speed of the numerical computations involved in solving the primal-dual SDP.

We ran a large number of randomly generated examples for various values of d and n. The comparison results are in listed in Table 4.1 and Table 4.4. The computations were performed on a Dell Laptop with a 2.0 GHz Pentium IV and 512MB of memory. Table 4.1 is the comparison of the lower bounds by formulation (4.1.2) and (4.1.5). Taking N = d/2 in Algorithm 4.3.3 appears to be good enough in practice for minimizing the Parrilo-Sturmfels polynomials. Our experiments show that increasing N above d/2 will not increase the lower bound significantly.

From Table 4.1, we can see that the lower bounds f_{sos}^* and $f_{N,grad}^*$ are close, agreeing to their leading 8 to 10 decimal digits, which confirms the observation made in [80] that almost *all* the polynomials gotten by subtracting their infima are SOS. Tables 4.2-4.4 are comparisons of running time in CPU seconds for formulations (4.1.2) and (4.1.5). The symbol "-" in the tables means that the computation takes more than one hour and we then terminate it. And "*" means we use a different scaling as described below.

Our formulation (4.1.5) uses about three quarters of the running time used by formulation (4.1.2). This may be unexpected since the use of gradients introduces many new variables. While we are not sure of the reason, one possible explanation is that adding gradients improves the conditioning and makes the interior-point algorithm for solving the SDP converge faster.

The numerical performance is subtle in this family of test polynomials. In the cases (n,d) = (4,10) or (n,d) = (5,10), our formulation (4.1.5) has numerical trouble, while (4.1.2) does not, and yet (4.1.5) is still faster than (4.1.2). However, for these two cases, if we scale $f(x_1, \ldots, x_n)$ so that the coefficients of $g(x_1, \ldots, x_n)$ belong to [-0.1, 0.1], both (4.1.2) and (4.1.5) do not have numerical trouble, and formulation (4.1.5) is still faster than (4.1.2). In Table 4.4 we see that the time ratio between (4.1.5) and (4.1.2) under this scaling is smaller than the time ratio for other values of (n, d). So numerical comparisons in Tables 4.1-4.4 for (n, d) = (4, 10)or (n, d) = (5, 10) are implemented under this new scaling, while for other values of (n, d) we still use the old scaling where the coefficients of $g(x_1, \ldots, x_n)$ belong

"-" means the computation is terminated if it takes more than one hour; "*" means the coefficients of $g(x_1, \dots, x_n)$ are scaled to belong to [-0.1, 0.1].

$d \setminus n$	3	4	5	6	7	8	9	10
4	5	7	9	10	11	13	14	15
6	10	19	38	41	232	-	-	-
8	17	78	186	233	-	-	-	-
10	40	39^{*}	102^{*}	-	-	-	-	-

Table 4.1: The relative difference $\frac{|f_{N,grad}^* - f_{sos}^*|}{|f_{sos}^*|} \times 10^{10}$, with N = d/2.

$d \setminus n$	3	4	5	6	7	8	9	10
4	0.16	0.24	0.42	0.86	1.86	7.56	25.85	73.69
6	0.32	1.17	8.40	49.04	309.66	-	-	-
8	1.10	12.23	173.98	1618.86	-	-	-	-
10	3.15	64.48^{*}	2144.04*	-	-	-	-	-

Table 4.2: Running time in CPU seconds via traditional SOS approach (4.1.2)

$d \setminus n$	3	4	5	6	7	8	9	10
4	0.12	0.18	0.32	0.68	1.46	5.65	18.85	54.97
6	0.23	0.91	6.39	35.16	241.71	-	-	-
8	0.84	9.54	129.53	1240.23	-	-	-	-
10	2.59	45.14^{*}	1539.80^{*}	-	-	-	-	-

Table 4.3: Running time in CPU seconds via our approach (4.1.5), with N = d/2.

$d \setminus n$	3	4	5	6	7	8	9	10
4	0.75	0.75	0.76	0.79	0.78	0.74	0.73	0.75
6	0.72	0.77	0.76	0.72	0.78	-	-	-
8	0.76	0.78	0.74	0.76	-	-	-	-
10	0.82	0.70^{*}	0.71^{*}	-	-	-	-	-

Table 4.4: The ratio of CPU seconds between (4.1.2) and (4.1.5), with N = d/2.

to [-1, 1]. A stability analysis for the scaling and the speed-up caused by adding gradients may be a future research topic.

Chapter 5

SOS Methods based on the Kuhn-Karush-Tucker (KKT) Ideal

As shown in Chapter 1, a sequence of lower bounds of $f^* = \min_{x \in S} f(x)$ can be obtained by solving the SOS program

$$f_N^*$$
 $(p_N^* resp.) = \max \ \gamma \ s.t. \ f(x) - \gamma \in \mathcal{M}(S)_{2N} \ (\mathcal{P}(S)_{2N} resp.).$

Lasserre [52] showed convergence $\lim_{N\to\infty} f_N^* = f^*$ under condition (1.1.14). If (1.1.14) fails but S is compact, we still have $\lim_{N\to\infty} p_N^* = f^*$. When S is *not* compact, we may not have convergence. In such situations, the gradient SOS methods introduced in Chapter 4 can be generalized to get a new sequence of lower bound of better properties. This chapter is based on joint work with Demmel and Powers [73].
5.1 Introduction

Consider the constrained polynomial optimization problem

$$f^* = \min \quad f(x) \tag{5.1.1}$$

s.t.
$$g_i(x) = 0, \quad i = 1, \cdots, s,$$
 (5.1.2)

$$h_j(x) \ge 0, \quad j = 1, \cdots, t$$
 (5.1.3)

where $x = [x_1 \cdots x_n] \in \mathbb{R}^n$ and $f(x), g_i(x), h_j(x) \in \mathbb{R}[X]$, the ring of real multivariate polynomials in $X = (x_1, \cdots, x_n)$. Let S be the feasible set defined by the constraints (5.1.2) - (5.1.3). Many optimization problems in practice can be formulated as (5.1.1)-(5.1.3). Finding the global optimal solutions to (5.1.1) - (5.1.3) is an NPhard problem, even if f(x) is quadratic and g_i, h_j are linear. For instance, the Maximum-Cut problem and nonconvex quadratic programming (QP) are NP-hard ([34, 64]).

Recently, the techniques of sum of squares (SOS) relaxations and moment matrix methods have made it possible to find globally optimal solutions to (5.1.1)-(5.1.3) by SOS relaxations (also called SDP relaxations in some references). For more details about these methods and their applications, see [45, 52, 53, 54, 55, 70, 71, 80, 81]. To have convergence for these methods, it is often necessary to assume that the feasible region S is compact or even finite. In [89], it is shown that SOS relaxations can solve (5.1.1)-(5.1.3) globally in finitely many steps in the case where $\{x \in \mathbb{C}^n : g_1(x) = \cdots = g_s(x) = 0\}$ is finite and the ideal $\langle g_1(x), \cdots, g_s(x) \rangle$ is radical. If we only assume that $\{x \in \mathbb{C}^n : g_1(x) = \cdots = g_s(x) = 0\}$ is finite, it is shown in [55] that the moment matrix method can solve (5.1.1)-(5.1.3) globally in finitely many steps. Finally, if S is compact and its quadratic module $\mathcal{M}(\mathcal{F})$ is archimedean (see Theorem 2.3.2), then arbitrarily close lower bounds for f^* can be obtained by SOS relaxations or moment matrix methods [52]. In this case, a convergence rate is given in Chapter 2.

The above global optimization methods are based on representation theorems from real algebraic geometry for polynomials positive and nonnegative on semialgebraic sets. On the other hand, the traditional local methods in optimization often follow the first order optimality conditions. The underlying idea in [71] and the present paper is to combine these two types of methods in order to more efficiently solve (5.1.1)-(5.1.3) globally. In [71], SOS relaxations are applied on the gradient ideal \mathcal{I}_{grad} (the ideal generated by all the partial derivatives of f(x)) in the unconstrained case, and on the KKT (Kuhn-Karush-Tucker) ideal I_{KKT} (defined below) in the constrained case, where only equality constraints are allowed. When \mathcal{I}_{grad} or I_{KKT} is radical, which is generically true in practice, the method in [71] can solve the optimization (5.1.1)-(5.1.2) globally; otherwise, arbitrarily close lower bounds of f^* can be obtained. No assumptions about S are made, i.e., it need not be finite or even compact. Jibetean and Laurent [45] also proposed a method to minimize polynomials by using the gradient ideal.

The KKT system of problem (5.1.1)-(5.1.3) is

$$F \stackrel{\Delta}{:=} \nabla f(x) + \sum_{i=1}^{s} \lambda_i \nabla g_i(x) - \sum_{j=1}^{t} \nu_j \nabla h_j(x) = 0, \qquad (5.1.4)$$

$$h_j(x) \ge 0, \nu_j h_j(x) = 0, \ j = 1, \cdots, t,$$
 (5.1.5)

$$g_i(x) = 0, \ i = 1, \cdots, s,$$
 (5.1.6)

where vectors $\lambda = [\lambda_1 \cdots \lambda_s]^T$ and $\nu = [\nu_1 \cdots \nu_t]^T$ are called Lagrange multipliers. See [75] for some regularity conditions that make the KKT system hold at local or global minimizers. For an example where the KKT system fails to define the global minimum, see Example 5.3.2 in Section 5.3.

Notice that we do not require $\nu \geq 0$ above; this makes the SOS relaxations simpler and does not affect the convergence of the method, since omitting the constraint $\nu \geq 0$ means simply that there are more feasible points for (5.1.4)-(5.1.6), including maxima as well as minima.

Define the KKT ideal I_{KKT} and its varieties as follows:

$$I_{KKT} = \langle F_1, \cdots, F_n, g_1, \cdots, g_s, \nu_1 h_1, \cdots, \nu_t h_t \rangle,$$

$$V_{KKT} = \{ (x, \lambda, \nu) \in \mathbb{C}^n \times \mathbb{C}^s \times \mathbb{C}^t : p(x, \lambda, \nu) = 0, \ \forall p \in I_{KKT} \},$$

$$V_{KKT}^{\mathbb{R}} = \{ (x, \lambda, \nu) \in \mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^t : p(x, \lambda, \nu) = 0, \ \forall p \in I_{KKT} \}.$$

Here $F = [F_1, \cdots, F_n]^T$ is defined in (5.1.4). Let

$$\mathcal{H} = \{ (x, \lambda, \nu) \in \mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^t : h_j(x) \ge 0, \ j = 1, \cdots, t \}.$$

The preorder cone \mathcal{P}_{KKT} associated with the KKT system is defined as

$$\mathcal{P}_{KKT} = \left\{ \left. \sum_{\theta \in \{0,1\}^t} \sigma_{\theta} h_1^{\theta_1} h_2^{\theta_2} \cdots h_t^{\theta_t} \right| \sigma_{\theta} \text{ are SOS } \right\} + I_{KKT}.$$

The *quadratic module* (sometimes called *linear cone*) associated with the KKT system is defined to be

$$\mathcal{M}_{KKT} = \left\{ \left. \sigma_0 + \sum_{j=1}^t \sigma_j h_j \right| \sigma_0, \cdots, \sigma_t \text{ are SOS} \right\} + I_{KKT}.$$

Notice that $I_{KKT} \subseteq \mathcal{M}_{KKT} \subseteq P_{KKT} \subseteq \mathbb{R}[x, \lambda, \nu].$

In solving SOS programs, we often set an upper bound on the degrees of the involved polynomials. Define the truncated KKT ideal

$$I_{N,KKT} = \Big\{ \sum_{k=1}^{n} \phi_k F_k + \sum_{i=1}^{s} \varphi_i g_i + \sum_{j=1}^{t} \psi_j \nu_j h_j \Big| deg(\phi_k F_k), deg(\varphi_i g_i), deg(\psi_j \nu_j h_j) \le N \Big\}.$$

and truncated preorder and linear cones

$$P_{N,KKT} = \left\{ \sum_{\theta \in \{0,1\}^t} \sigma_{\theta} h_1^{\theta_1} h_2^{\theta_2} \cdots h_t^{\theta_t} \middle| deg(\sigma_{\theta} h_1^{\theta_1} \cdots h_t^{\theta_t}) \leq N \right\} + I_{N,KKT}$$
$$\mathcal{M}_{N,KKT} = \left\{ \sigma_0 + \sum_{j=1}^t \sigma_j h_j \middle| \begin{array}{c} \sigma_0, \cdots, \sigma_t \text{ are SOS} \\ deg(\sigma_0), deg(\sigma_j h_j) \leq N \end{array} \right\} + I_{N,KKT}.$$

A sequence $\{p_N^*\}$ of lower bounds of (5.1.1)-(5.1.3) can be obtained by SOS relaxations:

$$p_N^* = \sup \ \gamma \tag{5.1.7}$$

s.t.
$$f(x) - \gamma \in P_{N,KKT}$$
. (5.1.8)

Since $P_{N,KKT}$ has a summation over 2^t terms like $\sigma_{\theta}h_1^{\theta_1}h_2^{\theta_2}\cdots h_t^{\theta_t}$, it is usually very expensive to solve the SOS program (5.1.7)-(5.1.8) in practice. So it is natural to replace the truncated preorder cone $P_{N,KKT}$ by the truncated linear cone $\mathcal{M}_{N,KKT}$, which leads to the SOS relaxations:

$$f_N^* = \max_{\gamma \in \mathbb{R}} \quad \gamma \tag{5.1.9}$$

s.t.
$$f(x) - \gamma \in \mathcal{M}_{N,KKT}$$
. (5.1.10)

Thus we get monotonically increasing lower bounds $\{f_N^*\}_{N=2}^{\infty}$ and $\{p_N^*\}_{N=2}^{\infty}$ such that $f_N^* \leq p_N^* \leq f^*$. The following is the convergence theorem, which will be proved in Section 5.3.

Theorem 5.1.1. Assume f(x) has a minimum $f^* := f(x^*)$ at one KKT point x^* of (5.1.1)-(5.1.3). Then $\lim_{N\to\infty} p_N^* = f^*$. Furthermore, if I_{KKT} is radical, then there exists some $N \in \mathbb{N}$ such that $p_N^* = f^*$, i.e., the SOS relaxations (5.1.7)-(5.1.8) converge in finitely many steps. The lower bounds $\{f_N^*\}$ are not guaranteed converge to f^* . However, if \mathcal{M}_{KKT} is archimedean (see Section 2.3), then we have convergence $\lim_{N\to\infty} f_N^* = f^*$ by Theorem 2.3.2 (Putinar's Positivestellensatz). We will return to this claim again in Section 5.3.

5.2 Representations in \mathcal{P}_{KKT} and \mathcal{M}_{KKT}

This section discusses the representations of objective polynomial f(x) in cones \mathcal{P}_{KKT} and \mathcal{M}_{KKT} .

Theorem 5.2.1. Assume I_{KKT} is radical. If f(x) is nonnegative on $V_{KKT}^{\mathbb{R}} \cap \mathcal{H}$, then f(x) belongs to \mathcal{P}_{KKT} .

To prove Theorem 5.2.1, we need the following lemma, which is a generalization of Lemma 4.2.3.

Lemma 5.2.2. Let W be an irreducible component of V_{KKT} . Then f(x) is constant on W.

Proof. Since W is irreducible and contains a real point, it remains irreducible if we replace $\mathbb{R}[X, \lambda]$ by $\mathbb{C}[X, \lambda]$. Thus W is connected in the strong topology on \mathbb{C}^{n+s} and hence is path-connected (see e.g. [107, 4.1.3]).

We notice that the Lagrangian function

$$\mathcal{L}(x,\lambda,\nu) = f(x) + \sum_{i=1}^{s} \lambda_i g_i(x) + \sum_{j=1}^{t} \nu_j h_j(x)$$

is equal to f(x) on V_{KKT} , which contains W. Choose two arbitrary points $(x^{(1)}, \lambda^{(1)}, \nu^{(1)})$, $(x^{(2)}, \lambda^{(2)}, \nu^{(2)})$ in W. We claim that $f(x^{(1)}) = f(x^{(2)})$. Firstly assume both $(x^{(1)}, \lambda^{(1)}, \nu^{(1)})$ and $(x^{(2)}, \lambda^{(2)}, \nu^{(2)})$ are nonsingular points. The set of nonsingular points consists a manifold. Since W is path-connected, there exists a piecewise-smooth path $\varphi(\tau) = (x(\tau), \lambda(\tau), \nu(\tau))$ $(0 \le \tau \le 1)$ lying inside W such that $\varphi(0) = (x^{(1)}, \lambda^{(1)}, \nu^{(1)})$ and $\varphi(1) = (x^{(2)}, \lambda^{(2)}, \nu^{(2)})$. Let $\mu_j(\tau)$ be the principle square root of $\nu_j(\tau)$, $1 \le j \le t$ (for a complex number $z = |z| \exp(\sqrt{-1\theta})$ with $0 \le \theta < 2\pi$, its principle square root is defined to be $\sqrt{|z|} \exp\{\frac{1}{2}\sqrt{-1\theta}\}$). From the KKT system (5.1.4)-(5.1.6), we can see that the function

$$f(x) + \sum_{i=1}^{s} \lambda_i g_i(x) + \sum_{j=1}^{t} \mu_j^2 h_j(x)$$

has zero gradient on the path $\varphi(\tau)$ $(0 \le \tau \le 1)$. By the Mean Value Theorem, we have $f(x^{(1)}) = f(x^{(2)})$.

Secondly consider the case that at least one of $(x^{(1)}, \lambda^{(1)}, \mu^{(1)})$ and $(x^{(2)}, \lambda^{(2)}, \mu^{(2)})$ is singular. Since the set of nonsingular points of W is dense and open in W([107, Chap. 4]), we can choose arbitrarily close nonsingular points to approximate $(x^{(1)}, \lambda^{(1)}, \mu^{(1)})$ and $(x^{(2)}, \lambda^{(2)}, \mu^{(2)})$. By continuity of f(x), we immediately have $f(x^{(1)}) = f(x^{(2)})$ and hence that f is constant on W. \Box

Proof of Theorem 5.2.1. Decompose V_{KKT} into its irreducible components, then by Lemma 5.2.2, f(x) is constant on each of them. Let W_0 be the union of all the components whose intersection with \mathcal{H} is empty, and group together the components on which f(x) attains the same value, say W_1, \ldots, W_r . Suppose $f(x) = \alpha_i \ge 0$ on W_i .

We have $V_{KKT} = W_0 \cup W_1 \cup \cdots \cup W_r$, and W_i are pairwise disjoint. Note that by our definition of irreducible, each W_i is conjugate symmetric. By Lemma 4.2.2, there exist polynomials $p_0, p_1, \cdots, p_r \in \mathbb{R}[x, \lambda, \nu]$ such that $p_i(W_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta function.

By assumption, $W_0 \cap \mathcal{H} = \emptyset$ and so, by Theorem 2.3.5, there are SOS polynomials v_{θ} ($\theta \in \{0, 1\}^t$) such that

$$-1 \equiv \sum_{\theta \in \{0,1\}^t} v_{\theta} h_1^{\theta_1} \cdots h_t^{\theta_t} \stackrel{def}{=} v_0 \quad mod \quad I(W_0).$$

We have $f = (f + \frac{1}{2})^2 - (f^2 + (\frac{1}{2})^2) = f_1 + v_0 \cdot f_2$ for the SOS polynomials $f_1 = (f + \frac{1}{2})^2, f_2 = f^2 + (\frac{1}{2})^2$. Then

$$f \equiv f_1 + v_0 f_2 \equiv \sum_{\theta \in \{0,1\}^t} u_\theta h_1^{\theta_1} \cdots h_t^{\theta_t} \stackrel{def}{=} q_0 \mod I(W_0)$$

for some SOS polynomials u_{θ} ($\theta \in \{0,1\}^t$). Recall that $f(x) = \alpha_i$, a constant, on each $W_i(1 \le i \le r)$. Set $q_i(x) = \sqrt{\alpha_i}$, then $f(x) = q_i(x)^2$ on $I(W_i)$.

Now let $q = q_0 \cdot (p_0)^2 + (\sum_{i=1}^r q_i p_i)^2$. Then f - q vanishes on V_{KKT} and hence $f - q \in I_{KKT}$ since I_{KKT} is radical. It follows that $f \in \mathcal{P}_{KKT}$. \Box

Remark 5.2.3. The assumption that I_{KKT} is radical is needed in Theorem 5.2.1, as shown by Example 3.4 in [71]. However, when I_{KKT} is not radical, the conclusion also holds if f(x) is strictly positive on $V_{KKT}^{\mathbb{R}}$.

Theorem 5.2.4. If f(x) is strictly positive on $V_{KKT}^{\mathbb{R}} \cap \mathcal{H}$ then f(x) belongs to \mathcal{P}_{KKT} .

Proof. As in the proof of Theorem 5.2.1, we decompose V_{KKT} into subvarieties W_0, W_1, \dots, W_r such that $W_0 \cap \mathcal{H} = \emptyset$, and for $i = 1, \dots, r, W_i \cap \mathcal{H} \neq \emptyset$ and f is constant on W_i . Since each $W_i, i > 0$ contains at least one real point and f(x) > 0 on $V_{KKT}^{\mathbb{R}}$, each $\alpha_i > 0$. The W_i were chosen so that each α_i is distinct, hence the W_i 's are pairwise disjoint.

Consider the primary decomposition $I_{KKT} = \bigcap_{i=0}^{r} J_i$ corresponding to our decomposition of V_{KKT} , i.e., $V(J_i) = W_i$ for $i = 0, 1, \dots, r$. Since $W_i \cap W_j = \emptyset$, we

have $J_i + J_j = \mathbb{R}[x, \lambda, \nu]$ by Theorem 2.2.2. The Chinese Remainder Theorem, see e.g. [29, 2.13], implies that there is an isomorphism

$$\rho: \mathbb{R}[x,\lambda,\nu]/I_{KKT} \to \mathbb{R}[x,\lambda,\nu]/J_0 \times \mathbb{R}[x,\lambda,\nu]/J_1 \times \cdots \times \mathbb{R}[x,\lambda,\nu]/J_r.$$

For any $p \in \mathbb{R}[x, \lambda, \nu]$, let [p] and $\rho([p])_i$ denote the equivalence classes of pin $\mathbb{R}[x, \lambda, \nu]/I_{KKT}$ and $\mathbb{R}[x, \lambda, \nu]/J_i$ respectively.

Recall that that $V(J_0) \cap \mathcal{H} = \emptyset$, hence by Theorem 2.3.5 there exist SOS polynomials u_{θ} ($\theta \in \{0, 1\}^t$) such that

$$-1 \equiv \sum_{\theta \in \{0,1\}^t} u_{\theta} \rho([h_1^{\theta_1}])_0 \cdots \rho([h_t^{\theta_t}])_0 \stackrel{def}{=} u_0 \mod J_0$$

As in the proof of Theorem 5.2.1, we write $f = f_1 - f_2$ for SOS polynomials f_1, f_2 and then we have

$$f \equiv f_1 + u_0 f_2 \equiv \sum_{\theta \in \{0,1\}^t} v_{\theta}(\rho([h_1^{\theta_1}]))_0 \cdots (\rho([h_t^{\theta_t}]))_0 \stackrel{def}{=} q_0 \mod J_0$$

for some SOS polynomials v_{θ} ($\theta \in \{0, 1\}^t$). Thus the preimage $\rho^{-1}((q_0, 0, \dots, 0)) \in \mathcal{P}_{KKT}$.

Now on each W_i , $1 \leq i \leq r$, $f(x) = \alpha_i > 0$, and hence $(f(x)/\alpha_i) - 1$ vanishes on W_i . Then by Theorem 2.2.3 there is $\ell \in \mathbb{N}$ such that $(f(x)/\alpha_i - 1)^\ell \in J_i$. From the binomial theorem, it follows that

$$(1 + (f(x)/\alpha_i - 1))^{1/2} \equiv \sum_{k=1}^{\ell-1} {\binom{1/2}{k}} (f(x)/\alpha_i - 1)^k \stackrel{def}{=} q_i/\sqrt{\alpha_i} \mod J_i \ .$$

Thus $(\rho([f]))_i = q_i^2$ is SOS in $\mathbb{R}[x, \lambda, \nu]/J_i$, and hence $\rho^{-1}(q_i^2 e_{i+1})$ is SOS in $\mathbb{R}[x, \lambda, \nu]/I_{KKT}$, where e_{i+1} is the (i+1)-st standard unit vector in \mathbb{R}^{r+1} .

Finally, we see that $\rho([f]) = (q_0, q_1^2, \cdots, q_r^2)$. The preimage of the latter is

$$\rho^{-1}((q_0, q_1^2, \cdots, q_r^2)) = \rho^{-1}(q_0 e_1)) + \sum_{i=1}^r \rho^{-1}(q_i^2 e_{i+1}),$$

which implies that $f \in \mathcal{P}_{KKT}$. \Box

Remark 5.2.5. The conclusions in Theorem 5.2.1 and Theorem 5.2.4 can not be strengthened to show that $f(x) \in \mathcal{M}_{KKT}$. The following is a counterexample.

Example 5.2.6. Consider the optimization

min
$$f(x) = (x_3 - x_1^2 x_2)^2 - 1 + \epsilon$$

s.t. $h_1(x) = 1 - x_1^2 \ge 0$
 $h_2(x) = x_2 \ge 0$
 $h_3(x) = x_3 - x_2 - 1 \ge 0$

where $0 < \epsilon < 1$. From the constraints, we can easily observe that the global minimum $f^* = \epsilon > 0$ which is attained at $x^* = (0, 0, 1)$. Its KKT ideal

$$I_{KKT} = \left\langle 2x_1 x_2 (x_3 - x_1^2 x_2) - \nu_1 x_1, 2x_1^2 (x_3 - x_1^2 x_2) + \nu_2 - \nu_3, \\ 2(x_3 - x_1^2 x_2) - \nu_3, \nu_1 (1 - x_1^2), \nu_2 x_2, \nu_3 (x_3 - x_2 - 1) \right\rangle$$

is radical (verified in Macaulay 2 [30]). However, we can not find SOS polynomials $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ and general polynomials ϕ_1, ϕ_2, ϕ_3 such that

$$f(x) = \sigma_0 + \sigma_1 h_1 + \sigma_2 h_2 + \sigma_3 h_3 + \phi_1 (\frac{\partial f}{\partial x_1} - \nu_1 x_2) + \phi_2 (\frac{\partial f}{\partial x_2} - \nu_2 + \nu_3) + \phi_3 (\frac{\partial f}{\partial x_3} - \nu_3).$$

Suppose to the contrary that they exist. Plugging $\nu = (0,0)$ in the above identity yields

$$0 = 1 - \epsilon + \sigma_0 + \sigma_1(1 - x_1^2) + \sigma_2 x_2 + \sigma_3(x_3 - x_2 - 1) + \phi(x_3 - x_1^2 x_2)$$

where $\phi = -4x_1\phi_1 - x_1^2\phi_2 + 2\phi_3 - (x_3 - x_1^2x_2)$. Now substitute $x_3 = x_1^2x_2$ in the above, yielding

$$\sigma_3((1-x_1^2)x_2+1) = 1 - \epsilon + \sigma_0 + \sigma_1(1-x_1^2) + \sigma_2 x_2.$$

Here $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ are now considered as SOS polynomials in (x_1, x_2) . Since $1 - \epsilon > 0$, σ_3 can not be the zero polynomial. If $\sigma_3 = \sigma_3(x_1)$ is independent of x_2 , we can derive a contradiction using an argument identical to the argument in the proof of of [86, Thm. 2]. Thus $2m = \deg_{x_2}\sigma_3(x_1, x_2) \ge 2$ and $2d = \deg_{x_1}\sigma_3(x_1, x_2) \ge 0$. On the left hand side, the leading term is of the form $A \cdot x_1^{2d+2} x_2^{2m+1}$ with coefficient A < 0. Since the degree in x_2 on the left hand side is odd, the leading term on the right hand side must come from $\sigma_2(x_1, x_2)x_2$, and is of the form like $B \cdot x_1^{2d} x_2^{2m+1}$ with B > 0. This is a contradiction. Therefore we can conclude that $f(x) \notin \mathcal{M}_{KKT}$.

5.3 Convergence of the lower bounds

In this section, we give the proof of Theorem 5.1.1. To get the convergence of $\{f_N^*\}$, we need some extra assumptions.

Proof of Theorem 5.1.1. The sequence $\{p_N^*\}$ is monotonically increasing, and $p_N^* \leq f^*$ for all $N \in \mathbb{N}$, since f^* is attained by f(x) in the KKT system (5.1.4)-(5.1.6) by assumption and the constraint (5.1.10) implies that $\gamma \leq f^*$. Now for arbitrary $\epsilon > 0$, let $\gamma_{\epsilon} = f^* - \epsilon$ and replace f(x) by $f(x) - \gamma_{\epsilon}$ in (5.1.1)-(5.1.3). The KKT system remains unchanged, and $f(x) - \gamma_{\epsilon}$ is strictly positive on $V_{KKT}^{\mathbb{R}}$. By Theorem 5.2.4, $f(x) - \gamma_{\epsilon} \in \mathcal{P}_{KKT}$. Since $f(x) - \gamma_{\epsilon}$ is fixed, there must exist some integer N_1 such that $f(x) - \gamma_{\epsilon} \in P_{N_1,KKT}$. Hence $f^* - \epsilon \leq p_{N_1}^* \leq f^*$. Therefore we have that $\lim_{N \to \infty} p_N^* = f^*$.

Now assume that I_{KKT} is radical. Replace f(x) by $f(x) - f^*$ in (5.1.1)-(5.1.3). The KKT system still remains the same, and $f(x) - f^*$ is now nonnegative on $V_{KKT}^{\mathbb{R}}$. By Theorem 5.2.1, $f(x) - f^* \in \mathcal{P}_{KKT}$. So there exists some integer N_2 such that $f(x) - f^* \in P_{N_2, KKT}$, and hence $P_{N_2}^* \ge f^*$. Then $p_N^* \le f^*$ for all N implies that $p_{N_2}^* = f^*$.

Remark 5.3.1. (i) In Lasserre's method [52], a sequence of lower bounds that converge to f^* asymptotically can be obtained when the feasible region S is compact; but those lower bounds usually do not converge in finitely many steps. However, from Theorem 5.1.1, we see that when I_{KKT} is radical then the lower bounds $\{p_N^*\}$ converge in finitely many steps, even if S is not compact. This implies that the lower bounds $\{p_N^*\}$ may have better convergence even in the case where S is compact. (ii) The assumption in Theorem 5.1.1 is non-trivial and can not be removed, which is illustrated by the following example.

Example 5.3.2. Consider the optimization: min x s.t. $x^3 \ge 0$. Obviously $f^* = 0$ and the global minimizer $x^* = 0$. However, the KKT system

$$1 - \nu \cdot 3x^2 = 0, \quad \nu \cdot x^3 = 0, \quad x^3 \ge 0, \quad \nu \ge 0$$

is not satisfied, since $V_{KKT} = \emptyset$. Actually we can see that the lower bounds $\{f_N^*\}$ given by (5.1.9)-(5.1.10) tend to infinity. By Theorem 2.2.2, $V_{KKT} = \emptyset$ implies that $1 \in \mathcal{P}_{KKT}$, i.e.,

$$(1+3\nu x^2)(1-3\nu x^2)+9\nu^2 x\cdot\nu x^3=1.$$

In the SOS relaxation (5.1.9)-(5.1.10), for arbitrarily large γ , $x - \gamma \in \mathcal{P}_{KKT}$, since

$$x - \gamma = (x - \gamma)(1 + 3\nu x^2)(1 - 3\nu x^2) + 9\nu^2 x(x - \gamma) \cdot \nu x^3 \in \mathcal{P}_{KKT}.$$

Thus $p_8^* = \infty$. In this example, the conclusion in Theorem 5.1.1 does not hold.

The convergence of lowers bounds $\{f_N^*\}$ cannot be guaranteed, as we see in Example 5.2.6. In that example, replace the objective by the perfect square $(x_3 - x_1^2 x_2)^2$. Then $f^* = 1$, but we do not have $\lim_{N \to \infty} f_N^* = 1$. From the arguments there, we can see that $f(x) - (1 - \epsilon) \notin \mathcal{M}_{KKT}$ for all $0 < \epsilon < 1$, which implies that $f_N^* \leq 0$. But $f_N^* \geq 0$ is obvious since $(x_3 - x_1^2 x_2)^2$ is a perfect square. Therefore $\lim_{N \to \infty} f_N^* = 0 < 1 = f^*$, i.e., the lower bounds $\{f_N^*\}$ obtained from (5.1.9)-(5.1.10) may not converge.

On the other hand, the situation is often not that bad in practice. In the examples in the rest of this paper, it always happens that $\lim_{N\to\infty} p_N^* = \lim_{N\to\infty} f_N^* = f^*$. If we further assume that \mathcal{M}_{KKT} is archimedean then it must hold that $\lim_{N\to\infty} p_N^* = \lim_{N\to\infty} f_N^* = f^*$ from Theorem 2.3.2 (Putinar). This is the generalization of assumption 4.1 in [52]. See also the remark after Theorem 2.3.2.

The SOS relaxation (5.1.9)-(5.1.10) can be solved using software SOS-TOOLS [88], or GloptiPoly [40]. The SOS relaxations (5.1.9)-(5.1.10) not only give the lower bounds f_N^* , but also the information about global minimizers x^* and their Lagrange multipliers (λ^*, ν^*). SOSTOOLS can extract the minimizer when the moment matrix has rank one. Gloptipoly can also find the lower bounds, and extract the global minimizers when the moment matrix satisfies some rank condition ([41]). Gloptipoly does not need the moment matrix to be rank one.

Example 5.3.3 (Exercise 2.18, [43]). Consider the global optimization:

min
$$(-4x_1^2 + x_2^2)(3x_1 + 4x_2 - 12)$$

s.t. $3x_1 - 4x_2 \le 12$, $2x_1 - x_2 \le 0$, $-2x_1 - x_2 \ge 0$.

The global minimum is $f^* \approx -18.6182$ and the minimizer is $x^* = (-24/55, 128/55) \approx$ (-0.4364, 2.3273). The lower bound obtained from (5.1.9)-(5.1.10) is $f_4^* \approx -18.6182$. The extracted minimizer is $\hat{x} \approx (-0.4364, 2.3273)$, which coincides with x^* . **Example 5.3.4.** Consider the Quadratically Constrained Quadratic Program (QCQP):

min
$$-\frac{4}{3}x_1^2 + \frac{2}{3}x_2^2 - 2x_1x_2$$

s.t. $x_2^2 - x_1^2 \ge 0, -x_1x_2 \ge 0.$

The global minimum is $f^* = 0$ and minimizer is $x^* = (0, 0)$. The feasible region S defined by the constraints is non-compact. The lower bound returned by (5.1.9)-(5.1.10) is $f_4^* \approx -2.6 \times 10^{-15}$ (Note: this computation was done in double precision floating point, with round off error bounded by $2^{-53} \approx 10^{-16}$). The extracted minimizer is $\hat{x} \approx (6.1 \times 10^{-16}, -9.0 \times 10^{-17})$ and the Lagrange multiplier is $\hat{\nu} \approx (0.3884, 0.3909)$.

5.4 Structures over some special constraints

In SOS relaxation (5.1.9)-(5.1.10), the polynomials are in $(x, \lambda, \nu) \in \mathbb{R}^{n+s+t}$. It is very expensive to implement when there are many constraints. In practice, if the polynomials $g_i(x)$ or $h_j(x)$ are of special forms, the KKT system (5.1.4)- (5.1.6) can be simplified and so can (5.1.9)-(5.1.10). In this section, we consider the case where the constraints include the nonnegative orthant \mathbb{R}^n_+ or some box $[a, b]_n = \{x \in \mathbb{R}^n : a \leq x \leq b\}.$

5.4.1 Nonnegative orthant \mathbb{R}^n_+

In this subsection, suppose the inequality constraints (5.1.3) are the nonnegative orthant \mathbb{R}^{n}_{+} . Then (5.1.2)-(5.1.3) have the form

$$g_1(x) = \dots = g_s(x) = 0, \quad x \in \mathbb{R}^n_+.$$

The KKT system (5.1.4)-(5.1.6) becomes

$$\nabla f(x) + \sum_{i=1}^{s} \lambda_i \nabla g_i(x) - \nu = 0,$$

$$g_1(x) = \dots = g_s(x) = 0,$$

$$x_k \nu_k = 0, \quad k = 1, \dots, n,$$

$$x \in \mathbb{R}^n_+, \quad \nu \in \mathbb{R}^n.$$

We can see that Lagrange multiplier ν can be solved for explicitly. By eliminating ν , the above system simplifies to

$$x_k\left(\frac{\partial f}{\partial x_k} + \sum_{i=1}^s \lambda_i \frac{\partial g_i}{\partial x_k}\right) = 0, \quad k = 1, \cdots, n$$
(5.4.11)

$$g_1(x) = \dots = g_s(x) = 0.$$
 (5.4.12)

We define cones $\mathcal{M}_{KKT}^{\mathbb{R}^n_+}$ and $\mathcal{M}_{N,KKT}^{\mathbb{R}^n_+}$, associated to the above simplified system, similar to the definition of \mathcal{M}_{KKT} and $\mathcal{M}_{N,KKT}$. Note that $\mathcal{M}_{KKT}^{\mathbb{R}^n_+}, \mathcal{M}_{KKT}^{\mathbb{R}^n_+} \subset \mathbb{R}[x,\lambda]$ and the Lagrange multiplier ν does not appear. Similar to (5.1.9)-(5.1.10), a sequence $\{\hat{f}_N^*\}$ of lower bounds of (5.1.1)-(5.1.3) can be obtained by the following SOS relaxations:

$$\hat{f}_N^* = \max_{\gamma \in \mathbb{R}} \quad \gamma \tag{5.4.13}$$

s.t.
$$f(x) - \gamma \in \mathcal{M}_{N,KKT}^{\mathbb{R}^n_+}$$
 (5.4.14)

Now the indeterminates in the above SOS program are (x, λ) instead of (x, λ, ν) .

Since ν is eliminated by direct substitutions, systems (5.1.4)-(5.1.6) and (5.4.11)-(5.4.12) are equivalent. Thus we see that $f(x) - \gamma \in \mathcal{M}_{N_1,KKT}$ if and only if $f(x) - \gamma \in \mathcal{M}_{N_2,KKT}^{\mathbb{R}^n_+}$, for some integers N_1 and N_2 . Therefore the lower bounds $\{\hat{f}_N^*\}$ have the same property of convergence as $\{f_N^*\}$ obtained from (5.1.9)-(5.1.10). If, in addition, the constraints (5.1.2) are the standard simplex:

$$Ax = b, \quad x \ge 0$$

where $A \in \mathbb{R}^{s \times n}, b \in \mathbb{R}^{s}$, the KKT system (5.1.4)-(5.1.6) can be furtherly reduced to

$$x_k(\frac{\partial f}{\partial x_k} + a_k^T \lambda) = 0, \ k = 1, \cdots, n$$

 $Ax = b, \ x \ge 0$

where $a_k \in \mathbb{R}^s$ is the k-th column of matrix A.

Furthermore, if Ax = b consists of a single equation $a^T x = b \neq 0$, then $\lambda = -\frac{x^T \nabla f(x)}{b}$ and the KKT system has the simpler form

$$x_k(\frac{\partial f}{\partial x_k} - \alpha_k \frac{x^T \nabla f(x)}{b}) = 0, \ k = 1, \cdots, n$$
$$a^T x = b, \ x \ge 0$$

where $a = [\alpha_1, \cdots, \alpha_n]^T$.

Based on the reduced KKT systems, simpler SOS relaxations can be obtained.

Example 5.4.1 (Test Problem 2.9, [32]). Consider the Maximum Clique Problem for n = 5:

$$\min - \left(\sum_{i=1}^{4} x_i x_{i+1} + x_1 x_5 + x_1 x_4 + x_2 x_5 + x_3 x_5\right)$$

$$s.t. \quad x_1 + x_2 + x_3 + x_4 + x_5 = 1$$

$$x_1, x_2, x_3, x_4, x_5 \ge 0.$$

The global minimum $f^* = -1/3$ and minimizers x^* are (1/3, 1/3, 0, 0, 1/3), (1/3, 0, 0, 1/3, 1/3), (0, 1/3, 1/3, 0, 1/3), and (0, 0, 1/3, 1/3, 1/3). The lower bound obtained from (5.4.13)-(5.4.14) is $\hat{f}_4^* \approx -0.333333333378814$. The difference is $f^* - \hat{f}_4^* \approx 4.5 \times 10^{-10}$. Example 5.4.2 (Exercise 1.20, [43]). Consider optimization problem:

min
$$\sum_{i=1}^{n-1} x_i^2 x_{i+1} + x_n^2 x_1$$

s.t. $\sum_{x_i=1}^n x_i = 1, \quad x \ge 0.$

The global minimum $f^* = 0$ and the minimizers are the vertices of the simplex defined by the constraints. The lower bound obtained from (5.4.13)-(5.4.14) is $\hat{f}_4^* = -4.0 \cdot 10^{-8}$.

Example 5.4.3. $f(x) = x^T H x$ and the constraints are $0 \le x \le e$, where $x \in \mathbb{R}^5$ and $e = [1, 1, 1, 1, 1]^T$, and

$$H = \begin{bmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{bmatrix}$$

is a co-positive matrix ([79, 81]), i.e., $f(x) \ge 0 \quad \forall x \ge 0$. If each x_i is replaced by x_i^2 , then the resulting quartic polynomial is nonnegative, but not SOS. Consider the Quadratic Program (QP):

$$\begin{array}{ll} \min & x^T H x \\ s.t. & x_1, x_2, x_3, x_4, x_5 \geq \end{array}$$

The lower bound obtained from (5.4.13)-(5.4.14) is $\hat{f}_2^* = -3.35 \times 10^{-9}$. Actually, we have the following decomposition

0.

$$x^{T}Hx = 0 + \sum_{i=1}^{5} 2 \cdot (x_{i} \cdot h_{i}^{T}x)$$

in (5.4.13)-(5.4.14). Here h_i is the *i*-th column of matrix H.

5.4.2 Box $[a, b]_n$

Consider the case that (5.1.3) is given by $a \le x \le b$ where $a = [a_1, \dots, a_n]^T$ and $b = [b_1, \dots, b_n]^T$, and a < b. Now the KKT system (5.1.4)-(5.1.6) has the form

$$\nabla f(x) + \sum_{i=1}^{s} \lambda_i \nabla g_i(x) - \nu + \mu = 0,$$

$$g_1(x) = \dots = g_s(x) = 0,$$

$$(x_k - a_k)\nu_k = 0, \quad (b_k - x_k)\mu_k = 0, \quad k = 1, \dots, n,$$

$$x - a \ge 0, \quad b - x \ge 0,$$

where $\nu_i(\mu_i, \lambda_i)$ is the *i*-th component of Lagrange multipliers $\nu(\mu, \lambda)$ respectively. One good property of this KKT system is that (ν, μ) can be solved for explicitly. Eliminating ν and μ , we have that

$$\left(\frac{\partial f}{\partial x_k} + \sum_{i=1}^s \lambda_i \frac{\partial g_i}{\partial x_k}\right)(x_k - a_k)(b_k - x_k) = 0, \quad k = 1, \cdots, n,$$
$$g_1(x) = \cdots = g_s(x) = 0, \quad x - a \ge 0, \quad b - x \ge 0.$$

Like the definitions of $\mathcal{M}_{KKT}^{\mathbb{R}^{n}_{+}}$ and $\mathcal{M}_{N,KKT}^{\mathbb{R}^{n}_{+}}$ (see the preceding subsection), define the cones $\mathcal{M}_{KKT}^{[a,b]_{n}}$ and $\mathcal{M}_{N,KKT}^{[a,b]_{n}}$ associated with the above simplified KKT system, where $\mathcal{M}_{KKT}^{[a,b]_{n}}$, $\mathcal{M}_{N,KKT}^{[a,b]_{n}} \subset \mathbb{R}[x,\lambda]$. Similar to (5.4.13)-(5.4.14), a sequence of lower bounds $\{\tilde{f}_{N}^{*}\}$ of (5.1.1)-(5.1.3) can be obtained by the following SOS relaxations:

$$\tilde{f}_N^* = \max_{\gamma \in \mathbb{R}} \quad \gamma \tag{5.4.15}$$

s.t.
$$f(x) - \gamma \in \mathcal{M}_{N,KKT}^{[a,b]_n}$$
 (5.4.16)

Now a polynomial $u(x, \lambda)$ of degree d in $\mathcal{M}_{d,KKT}^{[a,b]_n}$ has at most $\binom{n+s+d}{d}$ coefficients, which is much smaller than $\binom{n+s+2n+d}{d}$, the number of coefficients of one polynomial of degree d in $\mathcal{M}_{N,KKT}$. So (5.4.15)-(5.4.16) can be solved much more efficiently. Similarly as $\{\hat{f}_N^*\}$, the lower bounds $\{\tilde{f}_N^*\}$ have the same properties of convergence as $\{f_N^*\}$.

Consider the special case that $f(x) = \frac{1}{2}x^T H x + g^T x$ is a quadratic function and there are no equality constraints. Here $g \in \mathbb{R}^n$ and $H = H^T \in \mathbb{R}^{n \times n}$ is symmetric. The the above KKT system can be further reduced to

$$(h_k^T x + g_k)(x_k - a_k)(b_k - x_k) = 0, \quad k = 1, \cdots, n,$$

 $x - a \ge 0, \quad b - x \ge 0.$

Here h_k is the k-th row of matrix H and g_k is the k-entry of g. Finding the global minimum of a general nonconvex quadratic function over a box is an NP-hard problem. The relaxations (5.4.15)-(5.4.16) provide a new approach for such nonconvex quadratic programming problem.

Example 5.4.4 (Test Problem 4.7, [32]). Consider optimization problem

min
$$-12x_1 - 7x_2 + x_2^2$$

s.t. $-2x_1^4 + 2 - x_2 = 0$
 $0 \le x_1 \le 2, \quad 0 \le x_2 \le 3.$

The best known objective value ≈ -16.73889 . The lower bound obtained from (5.4.15)-(5.4.16) is $\tilde{f}_6^* \approx -16.73889$. So $f^* \approx \tilde{f}_6^*$. The extracted minimizer $\tilde{x} \approx$ (0.7175, 1.4698) and Lagrange multiplier $\tilde{\lambda} \approx -4.0605$.

Example 5.4.5 (Test Problem 2.1, [32]). Consider optimization problem

min
$$42x_1 + 44x_2 + 45x_3 + 47x_4 + 47.5x_5 - 50\sum_{i=1}^5 x_i^2$$

s.t. $20x_1 + 12x_2 + 11x_3 + 7x_4 + 4x_5 \le 40$
 $0 \le x_1, x_2, x_3, x_4, x_5 \le 1.$

The global minimum $f^* \approx -17$ and the minimizer $x^* = (1, 1, 0, 1, 0)$. The lower bound obtained from (5.4.15)-(5.4.16) is $\tilde{f}_6^* \approx -17.00$. The extracted minimizer $\tilde{x} \approx (1.00, 1.00, 0.00, 1.00, 0.00)$ and Lagrange multiplier $\tilde{\nu} \approx 0.1799$.

Example 5.4.6 (Exercise 2.22, [43]). Consider the Maximum Independent Set Problem

$$\min \quad -\sum_{i=1}^{n} x_i + \sum_{(i,j)\in E} x_i x_j$$
$$s.t. \quad 0 \le x_i \le 1, \quad i = 1, \cdots, n.$$

The negative of the global minimum $-f^*$ equals the cardinality of the maximum independent vertex set of G = (V, E). Let G be a pentagon with two diagonals which do not intersect in the interior. Now n = 5 and $f^* = -2$. The lower bound obtained from (5.4.15)-(5.4.16) is $\tilde{f}_4^* \approx -2.00$.

Example 5.4.7 (Exercise 1.32, [43]). Consider optimization problem

min
$$\prod_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i$$

s.t. $0 \le a \le x_1, \cdots, x_n \le b.$

The global minimum is $f^* = a^n - na$ when $a \ge 1$. For n = 4, a = 2, b = 3, the

lower bound obtained from (5.4.15)-(5.4.16) is $\tilde{f}_6^* \approx 8.00$. The extracted minimizer is $\tilde{x} \approx (2.00, 2.00, 2.00, 2.00)$.

Chapter 6

Minimizing Rational Functions

This chapter discusses the global minimization of rational functions. Consider the problem of minimizing a rational function

$$r^* = \min_{x \in \mathbb{R}^n} \quad r(x) := \frac{f(x)}{g(x)}$$
 (6.0.1)

s.t.
$$h_1(x) \ge 0, \cdots, h_m(x) \ge 0$$
 (6.0.2)

where f(x), g(x), $h_i(x) \in \mathbb{R}[X]$. The motivation is to find the global minimum r^* of the rational function r(x), and if possible, one or more global minimizer(s) x^* such that $r(x^*) = r^*$, subject to constraints. This contains a broad class of nonlinear global optimization problems. Without loss of generality, assume that g(x) is nonnegative and not identically zero on the feasible set; as long as g(x) is not identically zero, we can replace $\frac{f(x)}{g(x)}$ by $\frac{f(x)g(x)}{g^2(x)}$. The sum of squares (SOS) methods can be generalized to solve this problem. Some special features arise that differ from the polynomial case. The difficulty appears when the minimum occurs on the common zeros of f(x) and g(z).

6.1 SOS relaxation for unconstrained minimization

In this section, we discuss the global minimization of (6.0.1) without any constraints.

Obviously, γ is a lower bound for r^* if and only if the polynomial $f(x) - \gamma g(x)$ is nonnegative. Now by approximating the nonnegativity of $f(x) - \gamma g(x)$ by a sum of squares, we get the following SOS relaxation

$$\begin{array}{rl} r^*_{sos} := \sup_{\gamma} & \gamma \\ s.t. & f(x) - \gamma g(x) \in \sum \mathbb{R}[X]^2 \end{array}$$

For any γ feasible in the above formulation, we immediately have $r(x) \geq \gamma$ for every $x \in \mathbb{R}^n$. Thus every feasible γ (and hence including r_{sos}^*) is a lower bound for r(x), i.e., $r_{sos}^* \leq r^*$.

Let $2d = \max(\deg(f), \deg(g))$ (it must be even for r(x) to have a finite minimum) and m(x) be the column vector of monomials up to degree d

$$m(x)^T = [1, x_1, \cdots, x_n, x_1^2, x_1x_2, \cdots, x_n^2, x_1^3, \cdots, x_n^d]$$

Notice that the length of vector m(x) is $\binom{n+d}{d}$. As discussed in Section 2.1, the polynomial $f(x) - \gamma g(x)$ is SOS if and only there exists a symmetric matrix $W \succeq 0$ of length $\binom{n+d}{d}$ such that the following identity holds:

$$f(x) - \gamma g(x) \equiv m(x)^T W m(x). \tag{6.1.3}$$

Now we write f(x) (resp. g(x)) as $\sum_{\alpha \in F} f_{\alpha} x^{\alpha}$ (resp. $\sum_{\alpha \in F} g_{\alpha} x^{\alpha}$), where F is a finite subset of \mathbb{N}^n . i.e., F is the support of polynomials f(x) and g(x).

Throughout this chapter, we index the rows and columns of matrix W by monomials up to degree d, i.e., the indices for the entries in W have the form (α, β) where $\alpha, \beta \in \mathbb{N}^n$. For any $\alpha \in \mathbb{N}^n$, denote by B_α the coefficient matrix of x^α in $m(x)m(x)^T$ (see Section 2.1) When n = 1, the B_α are Hankel matrices. Now we can see that (6.1.3) holds if and only if

$$f_{\alpha} - \gamma g_{\alpha} = \langle B_{\alpha}, W \rangle, \quad \forall \alpha \in F.$$

Therefore the SOS relaxation of problem (6.0.1) is essentially the following (SDP):

$$r_{sos}^* := \sup_{\gamma, W} \quad \gamma \tag{6.1.4}$$

s.t.
$$f_{\alpha} - \gamma g_{\alpha} = \langle B_{\alpha}, W \rangle, \quad \forall \alpha \in F$$
 (6.1.5)

$$W \succeq 0. \tag{6.1.6}$$

Notice that the decision variables are γ and W instead of x.

Now let us derive the dual problem to SDP (6.1.4)-(6.1.5). Its Lagrange function is

$$\mathcal{L}(\gamma, W, y, S) = \gamma + \sum_{\alpha \in F} (f_{\alpha} - \gamma g_{\alpha} - \langle B_{\alpha}, W \rangle) y_{\alpha} + W \bullet S$$
$$= \sum_{\alpha \in F} f_{\alpha} y_{\alpha} + (1 - \sum_{\alpha \in F} g_{\alpha} y_{\alpha}) \gamma + (S - \sum_{\alpha \in F} y_{\alpha} B_{\alpha}) \bullet W$$

where $y = (y_{\alpha})$ and W are dual decision variables (Lagrange multipliers). The vector y is monomial-indexed, and S is a symmetric matrix of the same size as W. And

 $S \succeq 0$ corresponds to the constraint $W \succeq 0$. Obviously the following holds

$$\sup_{\gamma,W} \mathcal{L}(\gamma, W, y, S) = \begin{cases} \sum_{\alpha \in F} f_{\alpha} y_{\alpha} & \text{if } \sum_{\alpha \in F} g_{\alpha} y_{\alpha} = 1, \\ & \sum_{\alpha \in F} y_{\alpha} B_{\alpha} = S \\ +\infty & \text{otherwise.} \end{cases}$$

Therefore, the dual problem of (6.1.4)-(6.1.6) is

$$r_{mom}^* := \inf_{y} \quad \sum_{\alpha \in F} f_{\alpha} y_{\alpha} \tag{6.1.7}$$

s.t.
$$\sum_{\alpha} g_{\alpha} y_{\alpha} = 1$$
 (6.1.8)

$$M_d(y) \succeq 0. \tag{6.1.9}$$

where the matrix $M_d(y) := \sum_{\alpha} y_{\alpha} B_{\alpha}$ is the *d*-th moment matrix of y. (6.1.7)-(6.1.9) can also be considered as an generalization of moment approaches in [52] except the equality (6.1.8).

From the derivation of dual problem (6.1.7)-(6.1.9) we immediately have that $r_{sos}^* \leq r_{mom}^*$, which is referred to as *weak duality* in optimization duality theory. Actually we have stronger properties for the SOS relaxation (6.1.4)-(6.1.6) and its dual (6.1.7)-(6.1.9) as summarized in the following theorem.

Theorem 6.1.1. Assume that the SOS relaxation (6.1.4)-(6.1.6) has a feasible solution (γ, W) . Then the following properties hold for the primal problem (6.1.4)-(6.1.6) and its dual (6.1.7)-(6.1.9):

- (i) Strong duality holds, i.e., $r_{sos}^* = r_{mom}^*$, and $f(x) r_{sos}^*g(x)$ is SOS.
- (ii) The lower bound r_{sos}^* obtained from the SOS relaxation (6.1.4)-(6.1.6) is exact, i.e., $r_{sos}^* = r^*$, if and only if $f(x) - r^*g(x)$ is SOS.

(iii) When $r_{sos}^* = r^*$ and $u^{(j)}$ $(j = 1, \dots, t)$ are global minimizers, then every vector y in the set

$$y \in \left\{ \sum_{j=1}^{t} \theta_j m_{2d}(u^{(j)}) : \theta_j \ge 0, \sum_{j=1}^{t} \theta_j = 1 \right\}$$

is an optimal solution of (6.1.7)-(6.1.9).

Proof. (i) The result can be obtained from the standard duality theory of convex programs [96, §30], if we can show that there exists a vector \hat{y} such that $\sum_{\alpha} g_{\alpha} \hat{y}_{\alpha} = 1$ and $M_d(\hat{y}) \succ 0$. Let μ be a Lebesgue measure on \mathbb{R}^n with strictly positive density everywhere on \mathbb{R}^n and finite moments, i.e., $|\int x^{\alpha} d\mu| < \infty$ for all $\alpha \in \mathbb{N}^n$ (e.g., one density function can be chosen as $\exp(-\sum_{i=1}^n x_i^2)$). Define the vector $y = (y_{\alpha})$ as follows:

$$y_{\alpha} = \int x^{\alpha} d\mu < \infty$$

Then we can claim that

$$0 < \tau := \sum_{\alpha} g_{\alpha} y_{\alpha} = \int g(x) d\mu < \infty.$$

The second inequality is obvious since all the moments of μ are finite. For the first inequality, for a contradiction, suppose $\tau \leq 0$, that is,

$$\int g(x)d\mu \le 0$$

Since g(x) is assumed to be nonnegative everywhere and μ has positive density everywhere, we must have that g(x) should be identically zero, which is a contradiction. Then we prove that $M_d(y)$ is positive definite. For any monomial-indexed nonzero vector q with the same length as $M_d(y)$ (it corresponds to a nonzero polynomial q(x)), it holds that

$$q^T M_d(y)q = \sum_{0 \le |\alpha|, |\beta| \le d} y_{\alpha+\beta} q_\alpha q_\beta = \int \left(\sum_{0 \le |\alpha|, |\beta| \le d} x^{\alpha+\beta} q_\alpha q_\beta \right) d\mu = \int q(x)^2 d\mu > 0.$$

Now let $\hat{y} = y/\tau$, which obviously satisfies $\sum g_{\alpha}\hat{y}_{\alpha} = 1$ and $M_d(\hat{y}) \succ 0$. In other words, the problem (6.1.7)-(6.1.9) has an interior point. Therefore, from the duality theory of convex optimization, we know that the strong duality holds, i.e., $r_{sos}^* = r^*$ and the optimal solution set of (6.1.4)-(6.1.6) is nonempty.

As already shown in (i), the optimal solution set of (6.1.4)-(6.1.6) is nonempty, which implies the conclusion in (ii) immediately.

(iii) When $r_{sos}^* = r^*$, the optimal value in (6.1.7)-(6.1.9) is also r^* , by strong duality as established in (i). Now choose an arbitrary monomial-indexed vector y of the form

$$y = \sum_{j=1}^t \theta_j m_{2d}(u^{(j)})$$

for any θ such that $\theta_j \ge 0$, $\sum_{j=1}^t \theta_j = 1$. Then we have

$$\sum_{\alpha \in F} f_{\alpha} y_{\alpha} = \sum_{j=1}^{t} \theta_j f(u^{(j)}) = \sum_{j=1}^{t} \theta_j r^* = r^*.$$

And obviously $M_d(y) = \sum_{j=1}^t \theta_j m_d(u^{(j)}) m_d(u^{(j)})^T \succeq 0$. So y is a feasible solution with optimal objective value. Thus y is a optimal solution to (6.1.7)-(6.1.9).

The information about the minimizers of (6.0.1) can be found from the optimal solutions to the dual problem (6.1.7)-(6.1.9). Suppose $y^* = (y^*_{\alpha})$ with $y^*_{(0,\dots,0)} \neq 0$ is one minimizer of (6.1.7)-(6.1.9) such that the moment matrix $M_d(y^*)$ has rank one. Then there is a vector w, with the same length as $M_d(y^*)$, such that

$$M_d(y^*)/y^*_{(0,\cdots,0)} = ww^T$$

where the left hand side is the called *normalized* moment matrix, with (1, 1) entry being 1. Set $x^* := w(2 : n + 1)$. So for any monomial-index α , it holds that $w(\alpha) = (x^*)^{\alpha}$. Now plug the point x^* into the rational function r(x), evaluate it, then we can see that

$$r(x^*) = \frac{f(x^*)}{g(x^*)} = \frac{\sum_{\alpha} f_{\alpha}(x^*)^{\alpha}}{\sum_{\alpha} g_{\alpha}(x^*)^{\alpha}} = \frac{\sum_{\alpha} f_{\alpha}y_{\alpha}^*}{\sum_{\alpha} g_{\alpha}y_{\alpha}^*} = r_{mom}^* = r_{sos}^*$$

In other words, we get a point x^* at which the evaluation of objective r(x) equals the lower bound r_{sos}^* . Therefore, x^* is a global minimizer and r_{sos}^* equals the global minimum r^* . When $M_d(y^*)$ (with $y_{(0,\dots,0)}^* \neq 0$) has rank more than one and satisfies the flat extension condition, there is more than one global minimizer, and they can be found numerically by solving a particular eigenvalue problem. We refer to [24, 41] for more details about the flat extension condition and extracting minimizers. When it happens that $y_{(0,\dots,0)}^* = 0$, we can not normalize the moment matrix $M_d(y^*)$ to represent some measure, which might be due to the case that the infimum of r(x) is attained at infinity. For instance, consider the example that $r(x) := 1/(1 + x_1^2)$. The optimal solution is $y^* = (0, 0, 1)$, which can not be normalized.

In the rest of this section, we show some numerical examples. The problem (6.1.4)-(6.1.6) and its dual (6.1.7)-(6.1.9) are solved by YALMIP [57] which is based on SeDuMi [111].

Example 6.1.2. Consider the global minimization of the rational function

$$\frac{(x_1^2+1)^2+(x_2^2+1)^2}{(x_1+x_2+1)^2}$$

Solving (6.1.4)-(6.1.6) yields the lower bound $r_{sos}^* \approx 0.7639$. The solution y^* to (6.1.7)-(6.1.9) is

 $y^* \approx (0.2000, 0.1236, 0.1236, 0.0764, 0.0764, 0.0764, 0.0472, 0.0472, 0.0472, 0.0472, 0.0292, 0.0292, 0.0292, 0.0292, 0.0292).$

The rank of moment matrix $M_2(y^*)$ is one, and we can extract one point $x^* \approx (0.6180, 0.6180)$. The evaluation of r(x) at x^* shows that $r(x^*) \approx 0.7639$. So x^* is a global minimizer and 0.7639 is the global minimum (approximately, or ignoring rounding errors).

Example 6.1.3. Consider the global minimization of the rational function

$$\frac{x_1^4 - 2x_1^2x_2x_3 + (x_2x_3 + 1)^2}{x_1^2}$$

The lower bound given by (6.1.4)-(6.1.6) is $r_{sos}^* \approx 2.0000$. The solution y^* to (6.1.7)-(6.1.9) is

$$y^* \approx (1.0859, -0.0000, -0.0000, -0.0000, 1.0000, 0.0000, -0.0000, 0.8150, -0.0859,$$

 $0.8150, -0.0000, -0.0000, -0.0000, -0.0000, 0.0000, -0.0000, -0.0000, -0.0000,$
 $-0.0000, -0.0000, 1.0859, 0.0000, -0.0000, 0.8150, 0.0859, 0.8150, 0.0000, 0.0000,$
 $-0.0000, -0.0000, 2.3208, -0.0000, 0.1719, 0.0000, 2.3208).$

The moment matrix $M_2(y^*)$ does not satisfy the flat extension condition, and no minimizers can be extracted. Actually one can see that 2 is the global minimum by observing the identity

$$f(x) - 2g(x) = (x_1^2 - x_2x_3 - 1)^2.$$

The lower bound 2 is achieved at (1,0,0) and hence is the global minimum. There are infinitely many global minimizers.

The relationship between the bounds is $r_{mom}^* = r_{sos}^* \leq r^*$ But it may happen that $r_{sos}^* < r^*$, just like SOS relaxations for minimizing polynomials. Let us see the following example.

$$\frac{x_1^4 x_2^2 + x_1^2 x_2^4 + x_3^6}{x_1^2 x_2^2 x_3^2}$$

The lower bound given by (6.1.4)-(6.1.6) is $r_{sos}^* = 0$, and the solution y^* to (6.1.7)-(6.1.9) is

$$y_{(2,2,2)}^* = 1, \quad y_{\alpha}^* = 0, \, \forall \alpha \neq (2,2,2).$$

The global minimum $r^* = 3$ because

$$x_1^4 x_2^2 + x_1^2 x_2^4 + x_3^6 - 3x_1^2 x_2^2 x_3^2 \ge 0 \quad \forall x \in \mathbb{R}^3$$

and r(1, 1, 1) = 3. So in this example, the SOS lower bound $r_{sos}^* < r^*$. Actually for any $0 < \gamma \leq 3$, the polynomial

$$x_1^4 x_2^2 + x_1^2 x_2^4 + x_3^6 - \gamma x_1^2 x_2^2 x_3^2$$

is nonnegative but not SOS. The proof is the the same as to prove that Motzkin polynomial

$$x_1^4 x_2^2 + x_1^2 x_2^4 + x_3^6 - 3 x_1^2 x_2^2 x_3^2$$

is not SOS [95].

6.2 What if $r_{sos}^* < r^*$?

From Theorem 6.1.1, we know that $r_{sos}^* = r^*$ if and only if the polynomial $f(x) - r^*g(x)$ is a sum of squares. But sometimes $f(x) - r^*g(x)$ might not be SOS, as we observed in Example 6.1.4. In this subsection, we discuss how to minimize a rational function r(x) when $r_{sos}^* < r^*$. Here we generalize the big ball technique

introduced in [52], but we should be very careful about the zeros of the denominator g(x) in r(x).

Suppose we know in advance that at least one global minimizer of r(x)belongs to the ball $B(c, \rho) := \{x \in \mathbb{R}^n : \rho^2 - ||x - c||_2^2 \ge 0\}$ with center c and radius $\rho > 0$. Let $\pi(x) := \rho^2 - ||x - c||_2^2$. Then we immediately have that $r^* = \min_{x \in \mathbb{R}^n} r(x) = \min_{x \in B(c,\rho)} r(x)$. In practice, we often choose the center c = 0 and radius ρ big enough. So the original unconstrained minimization problem (6.0.1) becomes the constrained problem

$$\min_{x \in B(c,\rho)} \quad r(x).$$

One natural SOS relaxation of this constrained problem is

$$r_N^* := \sup_{\gamma} \quad \gamma \tag{6.2.1}$$

s.t.
$$f(x) - \gamma g(x) \equiv \sigma_0(x) + \sigma_1(x)\pi(x)$$
(6.2.2)

$$\deg(\sigma_1) \le 2(N-1), \ \sigma_0(x), \sigma_1(x) \in \sum \mathbb{R}[X]^2.$$
 (6.2.3)

Similar to the dual of (6.1.4)-(6.1.6), the dual problem of (6.2.1)-(6.2.3) can be found to be

$$\hat{r}_N^* := \inf_y \quad \sum_{\alpha \in F} f_\alpha y_\alpha \tag{6.2.4}$$

s.t.
$$\sum_{\alpha} g_{\alpha} y_{\alpha} = 1$$
 (6.2.5)

$$M_N(y) \succeq 0 \tag{6.2.6}$$

$$M_{N-1}(\pi * y) \succeq 0 \tag{6.2.7}$$

where π is the vector of the coefficients of polynomial $\pi(x)$. For a general polynomial $p(x) = \sum_{\alpha} p_{\alpha} x^{\alpha}$, the generalized moment matrix $M_k(p * y)$ is defined as (see

Section 2.3)

$$M_k(p*y)(\beta,\tau) := \sum_{\alpha} p_{\alpha} y_{\beta+\tau+\alpha}, \ 0 \le |\beta|, |\tau| \le k.$$

We have the following theorem for the SOS relaxation (6.2.1)-(6.2.3) and its dual (6.2.4)-(6.2.7).

Theorem 6.2.1. Assume that $r^* > -\infty$ and at least one global minimizer of r(x)lies in the ball $B(c, \rho)$. If the f(x) and g(x) in r(x) have no common real zeros on $B(c, \rho)$, then

- (i) The lower bounds converge: $\lim_{N\to\infty}r_N^*=r^*.$
- (ii) For N large enough, there is no duality gap between (6.2.1)-(6.2.3) and its dual (6.2.4)-(6.2.7), i.e., $r_N^* = \hat{r}_N^*$.
- (iii) For N large enough, $r_N^* = r^*$ if and only if $f(x) r^*g(x) = \sigma_0(x) + \sigma_1(x)\pi(x)$ for some SOS polynomials σ_0, σ_1 with $\deg(\sigma_1) \le 2(N-1)$.
- (iv) If $r_N^* = r^*$ for some integer N and $u^{(j)}$ $(j = 1, \dots, t)$ are global minimizers on $B(c, \rho)$, then every vector y in the set

$$y \in \left\{ \sum_{j=1}^{t} \theta_j m_{2N}(u^{(j)}) : \theta_j \ge 0, \sum_{j=1}^{t} \theta_j = 1 \right\}$$

is an optimal solution to (6.2.4)-(6.2.7).

Proof. (i) For any fixed $\gamma < r^*$, we can see that $f(x) - \gamma g(x) > 0$ on $B(c, \rho)$ if $g(x) \neq 0$ (we have assumed that g(x) is nonnegative). When g(x) = 0, we must have $f(x) \geq 0$. Otherwise assume f(u) < 0 at some point u with g(u) = 0. Then r(x) is unbounded from the below, which contradicts the assumption that $r^* > -\infty$. Thus

g(x) = 0 implies $f(x) \ge 0$ on $B(c, \rho)$. So we have that

$$f(x) - \gamma g(x) \ge 0, \quad \forall x \in B(c, \rho)$$

Since $\gamma < r^*$, $f(x) - \gamma g(x) = 0$ implies that f(x) = g(x) = 0, which is not possible. Therefore, the polynomial $f(x) - \gamma g(x)$ is positive on ball $B(c, \rho)$. Now by Putinar's Positivstellensatz (Theorem 2.3.2), there exist SOS polynomials σ_0, σ_1 with degree high enough such that

$$f(x) - \gamma g(x) \equiv \sigma_0(x) + \sigma_1(x)\pi(x)$$

So in (6.2.1)-(6.2.3), γ can be chosen arbitrarily close to r^* . Therefore we proved the convergence of lower bounds r_N^* .

(ii) Similar to the proof of Theorem 6.1.1, it suffices to show that the problem (6.2.4)-(6.2.7) has a strictly feasible solution. Let μ be a probability measure with uniform distribution on $B(c, \rho)$. Define the monomial-indexed vector $y = (y_{\alpha})$ in the following way:

$$y := \int x^{\alpha} du.$$

Now we show that $M_N(y)$ and $M_{N-1}(\pi * y)$ are positive definite. $M_N(y) \succ 0$ can be shown in the same way as in the proof of (i) in Theorem 6.1.1. Now we show that $M_{N-1}(\pi * y) \succ 0$. For any nonzero monomial-indexed vector q of the same length as $M_{N-1}(\pi * y)$ (it corresponds to a nonzero polynomial q(x) up to degree N-1), it holds that

$$q^{T}M_{N-1}(\pi * y)q = \int q(x)^{2}\pi(x)d\mu = \frac{1}{\operatorname{Vol}(B(c,\rho))} \int_{B(c,\rho)} q(x)^{2}\pi(x)dx > 0,$$

which implies that $M_{N-1}(\pi * y)$ is positive definite. In the above, $\operatorname{Vol}(B(c, \rho))$ denotes the volume of the ball $B(c, \rho)$. Since g(x) is not identically zero and always

nonnegative, g(x) can not be always zero on $B(c, \rho)$ and hence

$$\sum_{\alpha} g_{\alpha} y_{\alpha} = \int g(x) d\mu = \frac{1}{\operatorname{Vol}(B(c,\rho))} \int_{B(c,\rho)} g(x) dx > 0.$$

Now set the vector \hat{y} as $y / \sum_{\alpha} g_{\alpha} y_{\alpha}$. Then can see that \hat{y} is an interior point for the dual problem (6.2.4)-(6.2.7).

(iii) For any fixed $\hat{\gamma} < r^*$, from the previous arguments we know that the polynomial $f(x) - \gamma g(x)$ is positive on K. Then by Putinar's Theorem, there exist SOS polynomials $s_0(x), s_1(x)$ with $\deg(\sigma_1)$ high enough such that

$$f(x) - \hat{\gamma}g(x) \equiv s_0(x) + s_1(x)\pi(x).$$

This means that the primal convex problem (6.2.1)-(6.2.3) has a feasible solution. From (ii) we know its dual problem (6.2.4)-(6.2.7) has a strict interior point. Now applying the duality theory of standard convex programming, we know the solution set of (6.2.1)-(6.2.3) is nonempty. And notice that r^* is obviously an upper bound for all r_N^* .

When $r_N^* = r^*$, we know r_N^* is optimal. For N sufficiently large, by (ii), the primal problem (6.2.1)-(6.2.3) is guaranteed to have a solution. So there exist SOS polynomials $\sigma_0(x), \sigma_1(x)$ with $\deg(\sigma_1) \leq 2(N-1)$ such that

$$f(x) - r^*g(x) \equiv \sigma_0(x) + \sigma_1(x)\pi(x).$$

The "if" direction is obvious.

The proof of (iv) is the same as (iii) of Theorem 6.1.1. $\hfill \Box$

Remark 6.2.2. In Theorem 6.2.1, we need the assumption that the numerator f(x) and denominator g(x) have no common real zeros on ball $B(c, \rho)$ to show the

convergence $\lim_{N\to\infty} r_N^* = r^*$. When they have common real zeros, for any $\gamma < r^*$, the polynomial $f(x) - \gamma g(x)$ is not strictly positive on $B(c, \rho)$ and hence Putinar's Theorem can not be applied. In such situations, the convergence is not guaranteed (see Remark 6.3.5). However, in case of two variables, i.e., n = 2, if f(x) and g(x) have at most finitely many real common zeros on $B(c, \rho)$, we still have $\lim_{N\to\infty} r_N^* = r^*$; furthermore, if the global minimizers of r(x) are finite, finite convergence holds, i.e., there exists $N \in \mathbb{N}$ such that $r_N^* = r^*$. Please see Theorem 6.3.7 in Section 4. Notice that the ball $B(c, \rho)$ satisfies both conditions (i) and (ii) there.

Remark 6.2.3. When f(x) and g(x) have common zeros on $B(c, \rho)$, the solution to dual problem (6.2.4)-(6.2.7) is not unique. To see this fact, suppose $w \in B(c, \rho)$ is such that f(w) = g(w) = 0, and y^* is an optimal solution to (6.2.4)-(6.2.7). Now let $\hat{y} = m_{2N}(w)$, which is not zero since $\hat{y}_{(0,\dots,0)} = 1$. Then $\sum_{\alpha} f_{\alpha} \hat{y}_{\alpha} = \sum_{\alpha} g_{\alpha} \hat{y}_{\alpha} = 0$ and $M_N(\hat{y}) \succeq 0$, $M_{N-1}(\pi * \hat{y}) \succeq 0$. So we can see that $y^* + \hat{y}$ is another feasible solution with the same optimal value. In such situations, some extracted points from the moment matrix $M_N(y^* + \hat{y})$ might not be global minimizers and they may be the common zeros of f(x) and g(x). See Example 6.2.5.

Example 6.2.4. Consider the global minimization of the rational function (obtained by plugging $x_3 = 1$ in Example 6.1.4)

$$\frac{x_1^4 x_2^2 + x_1^2 x_2^4 + 1}{x_1^2 x_2^2}$$

Choose c = 0 and $\rho = 2$. For N = 3, the lower bound given by (6.2.1)-(6.2.3) is $r_3^* = 3$, and the solution to (6.2.4)-(6.2.6) is

$$y^* = (1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1).$$

The moment matrix $M_3(y^*)$ has rank 4, and satisfies the flat extension condition. The following four points are extracted: $(\pm 1, \pm 1)$. They are all global minimizers.

Example 6.2.5. Consider the global minimization of the rational function (obtained by plugging $x_2 = 1$ in Example 6.1.4)

$$\frac{x_1^4 + x_1^2 + x_3^6}{x_1^2 x_3^2}$$

Choose c = 0 and $\rho = 2$. For N = 4, the lower bound given by (6.2.1)-(6.2.3) is $r_4^* = 3.0000$, and the solution to (6.2.4)-(6.2.6) is

$$y^* \approx (2.8377, 0, 0, 1, 0, 0, 1.0008, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1)$$

The moment matrix has rank 6 and satisfies flat extension condition. Six points are extracted:

$$(\pm 1.0000, \pm 1.0000), \quad (0.0000, \pm 0.0211)$$

The evaluation of r(x) at these points shows that the first four points are global minimizers. The last two points are not global minimizers, but they are approximately common zeros of the numerator and denominator. See Remark 6.2.3.

6.3 Constrained minimization

In this section, consider the constrained optimization problem

$$r^* := \min_{x \in \mathbb{R}^n} \quad r(x) := \frac{f(x)}{g(x)}$$
 (6.3.1)

s.t.
$$h_1(x) \ge 0, \cdots, h_m(x) \ge 0$$
 (6.3.2)

where $f(x), g(x), h_i(x)$ are all real multivariate polynomials in $x = (x_1, \dots, x_n)$. Without confusion, let r^* still be the minimum objective value as in the unconstrained case. If some h_i are rational functions, we can reformulate the constraints $h_i(x) \ge 0$ equivalently as some polynomial inequalities (one should be careful with the zeros of $h_i(x)$). Denote by S the feasible set. Here we assume that g(x) is not identically zero on S, and g(x) is nonnegative on S (otherwise, e.g., replace $\frac{f(x)}{g(x)}$ by $\frac{f(x)g(x)}{g^2(x)}$).

When $g(x) \equiv 1$ (or a nonzero constant), problem (6.3.1)-(6.3.2) becomes a standard constrained polynomial optimization problem. Lasserre [52] (also see Chapter 1) proposed a general procedure to solve this kind of optimization problem by a sequence of sum of squares relaxations. When g(x) is a nonconstant polynomial nonnegative on S, Lasserre's procedure can be generalized in a natural way. For each fixed positive integer N, consider the SOS relaxation

$$r_N^* := \sup \quad \gamma \tag{6.3.3}$$

s.t.
$$f(x) - \gamma g(x) \equiv \sigma_0(x) + \sum_{i=1}^m \sigma_i(x) h_i(x)$$
 (6.3.4)

$$\deg(g_i) \le 2N - d_i, \ \sigma_i(x) \in \sum \mathbb{R}[X]^2$$
(6.3.5)

where $d_i = \lceil \deg(h_i)/2 \rceil$. For any feasible γ above, it is obvious that $f(x) - \gamma g(x) \ge 0$ on S and so hence $r(x) \ge \gamma$. Thus every such γ (and hence including r_N^*) is a lower bound of r(x) on S.

We denote by M(S) the set of polynomials which can be represented as

$$\sigma_0(x) + \sigma_1(x)h_1(x) + \dots + \sigma_m(x)h_m(x)$$

with all $\sigma_i(x)$ being SOS. M(S) is the quadratic module generated by polynomial tuple (h_1, \dots, h_m) . Throughout this section, we make the following assumption for M(S):
Assumption 6.3.1 (Constraint Qualification Condition). There exist R > 0and SOS polynomials $s_0(x), s_1(x), \dots, s_m(x) \in \sum \mathbb{R}[X]^2$ such that

$$R - ||x||_2^2 = s_0(x) + s_1(x)h_1(x) + \dots + s_m(x)h_m(x)$$

Remark 6.3.2. When the assumption above is satisfied, the quadratic module M(S) is said to be *archimedean* (see Section 2.3). Obviously, when this assumption holds, the semialgebraic set S is contained in the ball $B(0, \sqrt{R})$ and hence compact; but the converse might not be true. See Example 6.3.1 in [27] for a counterexample. Under this assumption, Putinar [91] showed that every polynomial p(x) positive on S belongs to M(S) (see Theorem 2.3.2).

Remark 6.3.3. When Assumption 6.3.1 does not hold, we can add to S one redundant constraint like $R - ||x||_2^2 \ge 0$ for R sufficiently large (e.g., a norm bound is known in advance for one global minimizer). Then the new quadratic module is always archimedean.

Similar to the derivation of (6.1.7)-(6.1.9), the dual problem of (6.3.3)-(6.3.4) can be found to be

$$\inf_{y} \quad \sum_{\alpha \in F} f_{\alpha} y_{\alpha} \tag{6.3.6}$$

s.t.
$$\sum_{\alpha} g_{\alpha} y_{\alpha} = 1 \tag{6.3.7}$$

$$M_N(y) \succeq 0 \tag{6.3.8}$$

$$M_{N-d_i}(h_i * y) \succeq 0, \ i = 1, \ \cdots, \ m.$$
 (6.3.9)

The properties of SOS relaxation (6.3.3)-(6.3.5) and (6.3.6)-(6.3.9) are summarized as follows:

Theorem 6.3.4. Assume that the minimum r^* of r(x) on S is finite, and f(x) = g(x) = 0 has no solutions on S. Then the following holds:

- (i) Convergence of the lower bounds: $\lim_{N \to \infty} r_N^* = r^*$.
- If, furthermore, S has nonempty interior, then (ii) and (iii) below are true.
- (ii) For N large enough, there is no duality gap between (6.3.3)-(6.3.5) and its dual
 (6.3.6)-(6.3.9).
- (iii) For N large enough, $r_N^* = r^*$ if and only if $f(x) r^*g(x) \equiv \sigma_0(x) + \sum_{i=1}^m \sigma_i h_i(x)$ for SOS polynomials $\sigma_i(x)$ with $\deg(\sigma_i h_i) \leq 2N$.
- (iv) If $r_N^* = r^*$ for some integer N and $u^{(j)}$ $(j = 1, \dots, t)$ are global minimizers on S, then every vector y in the set

$$y \in \left\{ \sum_{j=1}^{t} \theta_j m_{2N}(u^{(j)}) : \theta_j \ge 0, \sum_{j=1}^{t} \theta_j = 1 \right\}$$

is an optimal solution to (6.3.6)-(6.3.9).

Proof. (i) For any $\gamma < r^*$, we have that the polynomial

$$\vartheta_{\gamma}(x) := f(x) - \gamma g(x)$$

is nonnegative on S. When $\vartheta_{\gamma}(u) = 0$ for some point $u \in S$, we must have f(u) = g(u) = 0, since otherwise g(u) > 0 (g(x) is assumed to be nonnegative on S) and $r(u) = \gamma < r^*$, which is impossible. Therefore $\vartheta_{\gamma}(x)$ is positive on S. By Theorem 2.3.2, there exist SOS polynomials $\sigma_i(x)$ of degree high enough such that

$$\vartheta_{\gamma}(x) \equiv \sigma_0(x) + \sum_{i=1}^m \sigma_i(x) h_i(x).$$

Therefore the claim in (i) is true.

(ii),(iii) & (iv): The proof here is almost the same as the one of Theorem 6.2.1. In a similar way, show that (6.3.3)-(6.3.5) has a feasible solution, and (6.3.6)-(6.3.9) has an interior point. Then apply the duality theory of convex programming. In (iv), check every y with given form is feasible and achieves the optimal objective value.

Remark 6.3.5. In Theorem 6.3.4, we made the assumption that f(x) and g(x) have no common zeros on S. But sometimes f(x) and g(x) may have common zeros, and it is also possible that the minimum r^* is attained at the common zero(s) (in this case, f(x) and g(x) are of the same magnitude order around the common zero(s)). In such situations, we can not apply Theorem 2.3.2 and might not have convergence. For a counterexample, consider the global minimization (with n = 1)

min
$$r(x) := \frac{1+x}{(1-x^2)^2}$$

s.t. $(1-x^2)^3 \ge 0.$

The global minimum is $r^* = \frac{27}{32}$ and the minimizer is $x^* = -\frac{1}{3}$. However, for any $\gamma < \frac{27}{32}$, there do not exist SOS polynomials $\sigma_0(x), \sigma_1(x)$ such that

$$1 + x - \gamma (1 - x^2)^2 \equiv \sigma_0(x) + \sigma_1(x)(1 - x^2)^3.$$

Otherwise, for a contradiction, suppose they exist. Then the left hand side vanishes at x = -1 and so does the right hand side. So x = -1 is a zero of $\sigma_0(x)$ with multiplicity greater than one, since σ_0 is SOS. Hence x = -1 is a multiple zero of the left hand side, which is impossible since the derivative of $1 + x - \gamma(1 - x^2)^2$ at x = -1 is 1. This counterexample is motivated by the one given by Stengle [109], which shows that the polynomial $1 - x^2$ does not belong to the quadratic module $M((1-x^2)^3)$ since $1-x^2$ is not strictly positive on $\{x : (1-x^2)^3 \ge 0\}$. On the other hand, if we can know in advance that the global minimum is not attained where the denominator g(x) vanishes, one way to overcome this difficulty is to add more constraints which keep the global minimizers but kick out the zeros of g(x).

Remark 6.3.6. When f(x) and g(x) have common zeros on S, the solution to dual problem (6.3.6)-(6.3.9) is not unique. In such situations, some extracted points from the moment matrix $M_N(y^*)$ may not be global minimizers and they might be the common zeros of f(x) and g(x). See Remark 6.2.3.

When n = 2, i.e., in case of two variables, the distinguished representations of nonnegative polynomials by Scheiderer [97] are very useful. Under some conditions on the geometry of the feasible set S, the convergence or even finite convergence holds if f(x) and g(x) has finitely many common zeros on S. This leads to our next theorem.

Theorem 6.3.7. Suppose n = 2. Let $Z(f,g) = \{u \in S : f(u) = g(u) = 0\}$ and Θ be the set of global minimizer(s) of r(x) on S. We have convergence $\lim_{N \to \infty} r_N^* = r^*$ if $\Omega = Z(f,g)$ is finite and satisfies at least one of the following two conditions:

- (i) Each curve C_i = {x ∈ C² : h_i(x) = 0} (i = 1, · · · , m) is reduced and no two of them share an irreducible component. No point in Ω is a singular point of the curve C₁ ∪ · · · ∪ C_m.
- (ii) Each point of Ω is an isolated real common zero of $f(x) r^*g(x)$ in \mathbb{R}^2 , but not an isolated point of the feasible set S.

Furthermore, if $\Omega = Z(f,g) \cup \Theta$ is finite and satisfies at least one of (i) and (ii), then we have finite convergence, i.e., there exists an integer N such that $r_N^* = r^*$.

Proof. Firstly, assume that $\Omega = Z(f,g)$ is finite and satisfies at least one of (i) and (ii). For any $\gamma < r^*$, we have that the polynomial

$$\vartheta_{\gamma}(x) := f(x) - \gamma g(x)$$

is nonnegative on S. When $\vartheta_{\gamma}(u) = 0$ for some point $u \in S$, we must have f(u) = g(u) = 0, since otherwise g(u) > 0 and $r(u) = \gamma < r^*$, which is impossible. By assumption in the theorem, the nonnegative polynomial $\vartheta_{\gamma}(x)$ has at most finitely many zeros on S. Now applying Corollary 3.7(if (i) holds) or Corollary 3.10 (if (ii) holds) in [97], we know that there exist SOS polynomials $\sigma_i(x)$ of degree high enough such that

$$\vartheta_{\gamma}(x) \equiv \sigma_0(x) + \sum_{i=1}^m \sigma_i(x) h_i(x).$$

Secondly, assume that $\Omega = Z(f,g) \cup \Theta$ is finite and satisfies at least one of (i) and (ii). Consider the polynomial $\vartheta_{r^*}(x) := f(x) - r^*g(x)$, which is nonnegative on S. When $\vartheta_{r^*}(u) = 0$ for some $u \in S$, we must have either f(u) = g(u) = 0 or $r(u) = r^*$. Thus polynomial $\vartheta_{r^*}(x)$ has at most finitely many zeros on S. Corollary 3.7(if (i) holds) or Corollary 3.10 (if (ii) holds) in [97] implies that there are SOS polynomials $\sigma_i(x)$ with deg $(\sigma_i h_i) \leq 2N$ (N is large enough) such that

$$\vartheta_{r^*}(x) \equiv \sigma_0(x) + \sum_{i=1}^m \sigma_i(x) h_i(x)$$

which completes the proof.

Example 6.3.8. Consider the problem

$$\min_{x} \quad \frac{x_1^4 x_2^2 + x_1^2 x_2^4 + 1}{x_1^2 x_2^2}$$
s.t. $x_1, x_2 \ge 0, \ 1 - x_1^2 - x_2^2 \ge 0$

The SOS relaxation (6.3.3)-(6.3.5) of order N = 3 yields the lower bound $r_3^* \approx 5.000$, and we can extract one point $x^* \approx (0.7071, 0.7071)$ from the dual solution to (6.3.6)-(6.3.9). $r(x^*) \approx 5.0000$ shows that the point x^* is a global minimizer.

Example 6.3.9. Consider the problem

$$\min_{x} \quad \frac{x_1^4 + x_1^2 + x_3^6}{x_1^2 x_3^2} \\ s.t. \quad x_1, x_3 \ge 0, \ 1 - x_1^2 - x_3^2 \ge 0$$

The SOS relaxation (6.3.3)-(6.3.5) of order N = 3 yields lower bound $r_3^* \approx 3.2324$, and we can extract one point $x^* \approx (0.6276, 0.7785)$ from the dual solution to (6.3.6)-(6.3.9). $r(x^*) \approx 3.2324$ shows that the point x^* is a global minimizer.

Example 6.3.10. Consider the problem

$$\min_{x} \quad \frac{x_1^3 + x_2^3 + 3x_1x_2 + 1}{x_1^2(x_2 + 1) + x_2^2(1 + x_1) + x_1 + x_2}$$
s.t. $2x_1 - x_1^2 \ge 0, \ 2x_2 - x_2^2 \ge 0$
 $4 - x_1x_2 \ge 0, \ x_1^2 + x_2^2 - \frac{1}{2} \ge 0.$

The SOS relaxation (6.3.3)-(6.3.5) of order N = 2 yields lower bound $r_2^* = 1$ and we can extract three points (0,1), (1,0), (1,1) from the dual solution to (6.3.6)-(6.3.9). The evaluations of r(x) at these three points show that they are all global minimizers. Example 6.3.11. Consider the problem

$$\min_{x} \quad \frac{x_{1}^{4} + x_{2}^{4} + x_{3}^{4} + x_{1}^{2} + x_{2}^{2} + x_{3}^{2} + 2x_{1}x_{2}x_{3}(x_{1} + x_{2} + x_{3})}{x_{1}^{3} + x_{2}^{3} + x_{3}^{3} + 2x_{1}x_{2}x_{3}}$$
s.t. $x_{1}^{4} + x_{2}^{4} + x_{3}^{4} = 1 + x_{1}^{2}x_{2}^{2} + x_{2}^{2}x_{3}^{2} + x_{3}^{2}x_{1}^{2}$
 $x_{3} \ge x_{2} \ge x_{1} \ge 0$

The SOS relaxation (6.3.3)-(6.3.5) of order N = 3 yields $r_3^* \approx 2.0000$ and we can extract two points

$$x^* \approx (0.0000, 0.0000, 1.0000), \qquad x^{**} \approx (-0.0032, 0.9977, 0.9974)$$

from the dual solution to (6.3.6)-(6.3.9). x^* is feasible and $r(x^*) \approx 2.0000$ implies that x^* is a global minimizer. And x^{**} is not feasible, but if we round x^{**} to the nearest feasible point we get (0, 1, 1), which is another global minimizer since r(0, 1, 1) = 2.

Example 6.3.12. Consider the problem

$$\min_{x} \quad \frac{x_{1}^{2} + x_{2}^{2} + x_{3}^{2} + x_{4}^{2} + 2(x_{2} + x_{3} + x_{1}x_{3} + x_{1}x_{4} + x_{2}x_{4}) + 1}{x_{1} + x_{4} + x_{1}x_{2} + x_{2}x_{3} + x_{3}x_{4}}$$
s.t. $x_{1}^{2} + x_{2}^{2} - 2x_{3}x_{4} = 0$
 $4 - x_{1}^{2} - x_{2}^{2} - x_{3}^{2} - x_{4}^{2} \ge 0$
 $x_{1}, x_{2}, x_{3}, x_{4} \ge 0.$

The SOS relaxation (6.3.3)-(6.3.5) of order N = 3 yields $r_3^* \approx 2.0000$ and we can extract one point

$$x^* \approx (0.0002, 0.0000, 0.0000, 0.9998).$$

from the dual solution to (6.3.6)-(6.3.9). $r(x^*) \approx 2.0000$ implies that x^* is a global minimizer (approximately). Actually the exact global minimizer is (0, 0, 0, 1).

Chapter 7

Applications of Polynomial Optimization

This chapter shows some specific applications of polynomial optimization. Shape design of transfer functions, minimum ellipsoid bounds for polynomial systems, finding polynomials with a nontrivial GCD, maximum likelihood optimization, and sensor network localization will be discussed.

7.1 Shape optimization of transfer functions

Consider the linear time invariant (LTI) single-input-single-output (SISO) system

$$\dot{x}(t) = Ax(t) + bu(t)$$
 (7.1.1)

$$y(t) = c^T x(t) + du(t)$$
 (7.1.2)

where $A \in \mathbb{R}^{n \times n}$, $b, c \in \mathbb{R}^n$, $d \in \mathbb{R}$. u(t) is the input, x(t) is the state variable, and y(t) is the output. The relationship between the Laplace transformations of u(t) and y(t) is that $\mathcal{L}(y)(s) = H(s)\mathcal{L}(u)(s)$ where

$$H(s) = d + c^T (sI - A)^{-1}b$$

is called the *transfer function* of system (7.1.1)-(7.1.2). H(s) can also be written as the rational function

$$\frac{\sum_{k=0}^{n} \alpha_k s^k}{\sum_{k=0}^{n} \beta_k s^k} \equiv \frac{q_1(s)}{q_2(s)}$$

Note that $\deg(q_1) \leq \deg(q_2) \leq n$. Actually any rational function H(s) of this form is the transfer function of some particular LTI system (it is not unique). Any such LTI system is called a realization of H(s). There are many such (algebraically equivalent) LTI systems [17, chap. 9].

In some engineering applications, designers want the transfer function to have certain desirable properties. For example, we may want the Bode plot (the graph of |H(s)| versus the pure imaginary axis $s = j \cdot \omega$) to have a certain shape corresponding to some kind of filtering. In this section, we discuss the shape optimization problem of choosing the coefficients of rational function H(s) so that its Bode plot has some desired shape. For *discrete* LTI systems, i.e. the governing differential equation (7.1.1)-(7.1.2) is replaced by difference equations (see [17]), there are several papers [1, 35, 116] that show how to formulate the filter design problem as the solution of the feasibility problem for certain convex sets. The main idea of this section is to apply the spectral factorization of trigonometric polynomials, a characterization of nonnegative univariate polynomials, and semi-infinite programming. This approach can be used to design the transfer function to be a bandpass filter, piecewise constant or polynomial, or even have an arbitrary shape. Our contribution is to extend these results to the continuous time LTI SISO systems (7.1.1)-(7.1.2). In this case the transfer function is not a trigonometric polynomial and hence we cannot directly apply spectral factorization. Fortunately our transfer function is a univariate rational function, which lets us apply certain characterizations of nonnegative univariate polynomials over the whole axis $(-\infty, \infty)$, semi-axis $(0, \infty)$, or some finite interval [a, b]. Using these characterizations, we can solve the shape optimization problem for the following shapes: (i) standard bandpass filter design; (ii) arbitrary piecewise constant shape; (iii) arbitrary piecewise polynomial shape; (iv)general nonnegative function. The first three shape design problems can be solved by testing the feasibility of certain Linear Matrix Inequalities [14]. The fourth shape design can be obtained by semi-infinite programming (SIP) [82, 116]. In this section, we show how to get the first two kinds of designs. The designs for the latter two are similar, see [69]. There is a lot of related work in various kinds of filter design problems and characterizing nonnegative polynomials on lines, circles, or intervals. We refer to [1, 35, 116, 31, 67, 36].

Now we apply Theorem 2.3.7 to characterize the transfer function, which is similar to the spectral factorization for trigonometric polynomials. Observe that

$$|H(j\omega)|^{2} = \frac{|q_{1}(j\omega)|^{2}}{|q_{2}(j\omega)|^{2}} = \frac{|q_{1,even}(j\omega) + q_{1,odd}(j\omega)|^{2}}{|q_{2,even}(j\omega) + q_{2,odd}(j\omega)|^{2}}$$
$$= \frac{[q_{11}(\omega^{2})]^{2} + \omega^{2} [q_{12}(\omega^{2})]^{2}}{[q_{21}(\omega^{2})]^{2} + \omega^{2} [q_{22}(\omega^{2})]^{2}}$$
$$\equiv \frac{p_{1}(w)}{p_{2}(w)} \text{ where } w = \omega^{2}$$

Here $q_{i,even}$ and $q_{i,odd}$ denote the even and odd parts of the polynomial q_i , and $q_{ij}, i, j = 1, 2$ are defined accordingly. Note that $p_1(w)$ and $p_2(w)$ are nonnegative polynomials on $w \in [0, \infty)$. Conversely, by Theorem 2.3.7, given any such nonneg-

ative $p_1(w)$ and $p_2(w)$, it is possible to reconstruct the $q_{ij}(w)$, and so $q_i(j\omega)$ and $H(j\omega)$. In other words, $p_1(w)$ and $p_2(w)$ with $deg(p_1) \leq deg(p_2)$ satisfy $|H(j\omega)|^2 = p_1(w)/p_2(w)$ where $w = \omega^2$ for some transfer function $H(j\omega)$ if and only if they are nonnegative on $[0, \infty)$.

First, let us design a bandpass filter. The goal is to design a transfer function $|H(j\omega)|^2 = \frac{p_1(w)}{p_2(w)}$ which is close to one on some squared frequency $(w = \omega^2)$ interval $[w^{\ell}, w^r]$ and tiny in a neighborhood just outside this interval. The design rules can be formulated as

$$p_{1}(w), p_{2}(w) \geq 0, \ \forall \ w \geq 0$$

$$1 - \alpha \leq \frac{p_{1}(w)}{p_{2}(w)} \leq 1 + \beta, \ \forall \ w \in [w^{\ell}, w^{r}]$$

$$\frac{p_{1}(w)}{p_{2}(w)} \leq \delta, \ \forall \ w \in [w^{\ell}_{1}, w^{\ell}_{2}] \cup [w^{r}_{1}, w^{r}_{2}]$$

where the interval $[w_1^{\ell}, w_2^{\ell}]$ is to the left of $[w^{\ell}, w^r]$, and $[w_1^r, w_2^r]$ is to the right. Here α, β, δ are tiny tolerance parameters (say around .05). Let p_1 and p_2 be the vectors of coefficients of $p_1(w)$ and $p_2(w)$ respectively. Then the constraints above can we restated as

$$p_{1}, p_{2} \in K_{0,\infty}$$

$$p_{1} - (1 - \alpha)p_{2} \in K_{w^{\ell},w^{r}}$$

$$(1 + \beta)p_{2} - p_{1} \in K_{w^{\ell},w^{r}}$$

$$\delta p_{2} - p_{1} \in K_{w_{1}^{\ell},w_{2}^{\ell}} \cap K_{w_{1}^{r},w_{2}^{r}}$$

where the cones $K_{[a,b]}$ are defined as

$$K_{[a,b]} = \{p(t) \in \mathbb{R}[t] : p(w) \ge 0 \,\forall w \in [a,b]\}.$$

The above cone constraints can be expressed as $Ap \in K$ where

$$A = \begin{bmatrix} I_{n+1} & 0 \\ 0 & I_{n+1} \\ I_{n+1} & (\alpha - 1)I_{n+1} \\ -I_{n+1} & (1 + \beta)I_{n+1} \\ -I_{n+1} & \delta I_{n+1} \\ -I_{n+1} & \delta I_{n+1} \end{bmatrix}, \ p = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

and $K = K_{0,\infty} \times K_{0,\infty} \times K_{w^{\ell},w^{r}} \times K_{w^{\ell},w^{r}} \times K_{w_{1}^{\ell},w_{2}^{\ell}} \times K_{w_{1}^{r},w_{2}^{r}}$. Given (α, β, δ) , we solve a cone feasibility problem and then recover the coefficient of p (see [69]). As introduced in [35] for the discrete case, we can also consider the following objectives:

- minimize $\alpha + \beta$ for fixed δ and n
- minimize δ for fixed α , β , and n
- minimize the degree n of p_1 and p_2 for fixed α , β , and δ .

These optimization problems with objectives are no longer convex, but quasi-convex. This means that we can use bisection to find the solution by solving a sequence of LMI feasibility problems. A design example is shown in Figure 7.1 (see also [69]). The parameters in Figure 7.1 are $[w^l, w^r] = [2, 3]$, $[w_1^l, w_2^l] = [0, 1.8]$, $[w_1^r, w_2^r] = [3.2, 5]$, $\alpha = \beta = 0.05$, $\delta = 0.05$, n = 10.

Second, let us show how to design a piecewise constant shape. In other words, we want the transfer function to be close to given constant values $c_1, ..., c_m$ in a set of m disjoint intervals $\omega^2 = w \in [a_k, b_k]$, where $a_1 < b_1 < a_2 < b_2 < \cdots < a_m < b_m$. More precisely we want the transfer function to lie in the interval



Figure 7.1: A bandpass filter.

 $[(1 - \alpha)c_k, (1 + \beta)c_k]$ for $w \in [a_k, b_k]$. By picking enough intervals (picking *m* large enough) we can approximate any continuous function as closely as we like.

These constraints may be written in the form

$$p_1(w), p_2(w) \ge 0, \ \forall \ w \ge 0$$

$$(1-\alpha)c_k \le \frac{p_1(w)}{p_2(w)} \le (1+\beta)c_k, \ \forall \ w \in [a_k, b_k], \ k = 1, \cdots, m.$$

Similarly, these constraints can also be written as cone constraints

$$p_1(w), p_2(w) \in K_{0,\infty}$$

 $p_1 - (1 - \alpha)c_k p_2, \ (1 + \beta)c_k p_2 - p_1 \in K_{a_k, b_k}, k = 1, \cdots, m$

Now the design problem becomes to find vector p such that $Ap \in K$ where

$$A = \begin{bmatrix} I_{n+1} & 0 \\ 0 & I_{n+1} \\ I_{n+1} & (\alpha - 1)c_1I_{n+1} \\ (1 + \beta)c_1I_{n+1} & -I_{n+1} \\ \vdots & \vdots \\ I_{n+1} & (\alpha - 1)c_mI_{n+1} \\ (1 + \beta)c_mI_{n+1} & -I_{n+1} \end{bmatrix}, \ p = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

and $K = K_{0,\infty}^2 \times K_{a_1,b_1}^2 \times \cdots \times K_{a_m,b_m}^2$. By solving a particular feasibility problem, we can find the coefficients p (see [69] for details). Similar to bandpass filter design, various design objectives can be achieved by applying bisection. A design example for a step function with 3 steps is shown in Figure 7.2. The parameters in Figure 7.2



Figure 7.2: A 3-step constant filter.

are $[a_1, b_1] = [0, 1.8], [a_2, b_2] = [2, 3], [a_3, b_3] = [3.2, 5], c_1 = 1, c_2 = 3, c_3 = 2, \alpha = \beta = 0.05, n = 10.$

Lastly, let us show how to use Theorems 2.3.7 and 2.3.8 to recover the transfer function from the polynomials $p_1(w)$ and $p_2(w)$ that are obtained by design

(see [69] for more details). For given polynomials $p_1(w)$ and $p_2(w)$ ($w = \omega^2$) such that $\frac{p_1}{p_2}$ has some desired shape, we need to find real polynomials q_1 and q_2 so that

$$\frac{p_1(w)}{p_2(w)} = \left|\frac{q_1(j\omega)}{q_2(j\omega)}\right|^2.$$

To this end, given a polynomial p(w) that is nonnegative on $[0, \infty)$, we can find two polynomials $q_e(w)$ and $q_o(w)$ such that $p(w) = q_e^2(w) + w \cdot q_o^2(w)$ (see [69]). Then q_e contains the even coefficients and q_o the odd coefficients (modulo signs) of the desired polynomials q_1, q_2 .

7.2 Minimum ellipsoid bounds for polynomial systems

This section shows how to find a minimum ellipsoid bound on the solution set of parameterized polynomial systems. A full version of this section can be found in [70]. Consider the polynomial system of equalities and inequalities of the form:

$$\begin{cases} \phi_{1}(x_{1}, \cdots, x_{n}; \mu_{1}, \cdots, \mu_{r}) = 0 \\ \vdots \\ \phi_{s}(x_{1}, \cdots, x_{n}; \mu_{1}, \cdots, \mu_{r}) = 0 \\ \rho_{1}(x_{1}, \cdots, x_{n}; \mu_{1}, \cdots, \mu_{r}) \leq 0 \\ \vdots \\ \rho_{t}(x_{1}, \cdots, x_{n}; \mu_{1}, \cdots, \mu_{r}) \leq 0 \end{cases}$$
(7.2.3)

where $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ and $\mu = (\mu_1, \dots, \mu_r)^T \in \mathbb{R}^r$. For each $1 \leq i \leq s$ and $1 \leq j \leq t$, ϕ_i and ρ_j are multivariate polynomials in $(x, \mu) \in \mathbb{R}^{n+r}$. μ can be thought of as parameters perturbing the solution x. We are only interested in bounding x for all μ determined by (7.2.3). x can also be thought of as the projection of the solution

 $(x,\mu) \in \mathbb{R}^{n+r}$ of (7.2.3) into the subspace \mathbb{R}^n . We consider only real solutions, since many practical problems concern only real solutions.

Our goal is to bound the projected solution set defined as

$$\mathcal{S} = \{ x \in \mathbb{R}^n : \exists \ \mu \in \mathbb{R}^r \text{ s.t. } (x, \mu) \text{ satisfies system (7.2.3)} \}.$$

For a given μ , there may be no real x satisfying (7.2.3), or one unique such x, or several such x, or infinitely many such x. So S can be quite complicated.

The traditional approach in perturbation analysis of a system of equations is to find the maximum distance of the perturbed solutions to the unperturbed solution, i.e. to find a bounding ball of smallest radius with the unperturbed solution at the center. This approach works well when the solution set is almost a ball and the unperturbed solution lies near the center. Unfortunately, this is often not the case in practice, when the solution set is very elongated. Instead, we seek a bounding ellipsoid of smallest volume (in a sense defined below), which can more effectively bound many elongated sets.

The particular idea for finding minimum ellipsoids was introduced in [18, 19], where the authors try to find the minimum ellipsoids for linear systems whose coefficients are rational functions of perturbing parameters. In this section, we show how to find the minimum ellipsoid bounds for the projected solution set S.

An open ellipsoid in \mathbb{R}^n can be described as

$$\mathcal{E}(P,z) = \left\{ x \in \mathbb{R}^n : (x-z)^T P^{-1} (x-z) < 1 \right\}$$
(7.2.4)

where $P \in S_{++}^n$ is the *shape matrix*, and $z \in \mathbb{R}^n$ is the *center*. By taking the Schur

complement, the ellipsoid can be written in LMI form

$$\mathcal{E}(P,z) = \left\{ x \in \mathbb{R}^n : \begin{bmatrix} P & x-z \\ (x-z)^T & 1 \end{bmatrix} \succ 0 \right\}.$$
 (7.2.5)

For example, the ellipsoid in the 2D plane given by

$$\frac{(x_1 - z_1)^2}{a^2} + \frac{(x_2 - z_2)^2}{b^2} < 1$$

$$\begin{bmatrix} a^2 & 0\\ 0 & b^2 \end{bmatrix}.$$

has the shape matrix $P = \begin{bmatrix} a^2 \\ 0 \end{bmatrix}$

How do we measure the "size" of an ellipsoid? The "best" measure would appear to be its volume, which is proportional to $\sqrt{\det P}$. However, we will instead choose trace(P) to measure the size, for two reasons: 1) trace(P) is an affine function, whereas $\sqrt{\det P}$ is not; 2) trace(P) is zero if and only if all the axes are zero, but $\sqrt{\det P}$ is zero if any axis is zero.

The minimum ellipsoid bound can be found by solving the optimization problem:

$$\inf_{P \in S^n_{++}, z \in \mathbb{R}^n} trace(P)$$

$$\begin{array}{c} (x-z)^T P^{-1}(x-z) < 1 \\ s.t. \quad \text{for all } (x,\mu) \text{ satisfying} \\ \phi_i(x,\mu) = 0, \rho_j(x,\mu) \le 0 \end{array} \right\}.$$

$$(7.2.6)$$

In the rest of this section, we will show how to relax the constraint (7.2.7) by the technique of Sum of Squares (SOS), which can be reduced to solving some SDP.

The constraint (7.2.7) holds if and only if

$$1 - (x - z)^T P^{-1}(x - z) > 0 \text{ for all } \left\{ x \in \mathbb{R}^n \middle| \begin{array}{l} \phi_i(x, \mu) = 0, i = 1, \cdots, s \\ \rho_j(x, \mu) \le 0, j = 1, \cdots, t \end{array} \right\}.$$

A certificate for the above can be obtained immediately by applying Putinar's Positivstellensatz (Theorem 2.3.2).

Theorem 7.2.1. Suppose Putinar's constraint qualification (see Section 2.3) holds for the polynomial system $\{\pm\phi_1, \dots, \pm\phi_s, -\rho_1, \dots, -\rho_t\}$. Then, if the constraint (7.2.7) holds, there exist polynomials $\lambda_i = \lambda_i(x, \mu), \sigma_j = \sigma_j(x, \mu)$ such that

$$1 - (x - z)^T P^{-1}(x - z) + \sum_{i=1}^s \lambda_i \phi_i + \sum_{j=1}^t \sigma_j \rho_j \succeq_{sos} 0$$
$$\sigma_1, \cdots, \sigma_t \succeq_{sos} 0$$

where the inequality $q(x,\mu) \succeq_{sos} 0$ means that the polynomial $q(x,\mu)$ is SOS.

Proof. Let $p = 1 - (x - z)^T P^{-1}(x - z)$ and $\{\pm \phi_1, \dots, \pm \phi_s, -\rho_1, \dots, -\rho_t\}$ be the polynomials defining the semi-algebraic set in Theorem 2.3.2. Notice that p(x) is strictly positive on the semialgebraic set

$$\{(x,\mu): \phi_1 \ge 0, \cdots, \phi_s \ge 0, -\phi_1 \ge 0, \cdots, -\phi_s \ge 0, -\rho_1 \ge 0, \cdots, -\rho_t \ge 0\}.$$

Then by Theorem 2.3.2, there exist SOS polynomials φ , τ_i , ν_i ($i = 1, \dots, s$), and σ_j ($j = 1, \dots, t$) such that

$$1 - (x - z)^T P^{-1}(x - z) = \varphi + \sum_{i=1}^s (\tau_i - \nu_i)\phi_i - \sum_{j=1}^t \sigma_j \rho_j.$$

Let $\lambda_i = \nu_i - \tau_i$. Then we get the result in the theorem.

Remark 7.2.2. If $\{\pm \phi_1, \dots, \pm \phi_s, -\rho_1, \dots, -\rho_t\}$ does not satisfy the constraint qualification condition for Putinar's Theorem, we can add a redundant ball condition like $x^2 + \mu^2 \leq R$ for R sufficiently large. Then Putinar's Theorem can be applied.

Now we are ready to show how to find the minimum ellipsoid bounds. Denote by $R_N[x,\mu]$ the set of polynomials in (x,μ) with degrees at most N. By Theorem 7.2.1, the problem (7.2.6)-(7.2.7) can be relaxed as

$$\hat{E}_N: \min_{\substack{P \in S_{++}^n, z \in \mathbb{R}^n \\ \lambda_i, \sigma_j \in R_N[x,\mu]}} trace(P) \text{ subject to}$$

$$1 - (x - z)^T P^{-1}(x - z) + \sum_{i=1}^s \lambda_i \phi_i + \sum_{j=1}^t \sigma_j \rho_j \succeq_{sos} 0, \quad \sigma_1, \cdots, \sigma_t \succeq_{sos} 0$$

which can be rewritten as

$$\min_{\substack{P \in S_{++}^n, z \in \mathbb{R}^n \\ \lambda_i, \sigma_j \in R_N[x,\mu]}} trace(P)$$
s.t. $1 - \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} I & -z \end{bmatrix}^T P^{-1} \begin{bmatrix} I & -z \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} + \sum_{i=1}^s \lambda_i \phi_i + \sum_{j=1}^t \sigma_j \rho_j \succeq_{sos} 0, \quad \sigma_1, \cdots, \sigma_t \succeq_{sos} 0.$

Now by introducing a new matrix variable Q, this becomes

$$\min_{\substack{Q,P \in S_{++}^n, z \in \mathbb{R}^n \\ \lambda_i, \sigma_j \in R_N[x,\mu]}} trace(P) \quad \text{subject to} \\ 1 - \begin{bmatrix} x \\ 1 \end{bmatrix}^T Q \begin{bmatrix} x \\ 1 \end{bmatrix} + \sum_{i=1}^s \lambda_i(x,\mu)\phi_i + \sum_{j=1}^t \sigma_j(x,\mu)\rho_j \succeq_{sos} 0 \\ \begin{bmatrix} I & -z \end{bmatrix}^T P^{-1} \begin{bmatrix} I & -z \end{bmatrix} \preceq Q, \qquad \sigma_1, \cdots, \sigma_t \succeq_{sos} 0. \end{cases}$$

Taking a Schur complement, this is equivalent to

$$E_N : p_N^* = \min_{\substack{Q, P \in S^n, z \in \mathbb{R}^n \\ \lambda_i, \sigma_j \in R_N[x,\mu]}} trace(P) \qquad \text{subject to}$$
(7.2.8)

$$1 - \begin{bmatrix} x \\ 1 \end{bmatrix}^T Q \begin{bmatrix} x \\ 1 \end{bmatrix} + \sum_{i=1}^s \lambda_i(x,\mu)\phi_i + \sum_{j=1}^t \sigma_j(x,\mu)\rho_j \succeq_{sos} 0$$
(7.2.9)

$$\begin{bmatrix} P & (I & -z) \\ (I & -z)^T & Q \end{bmatrix} \succeq 0, \quad \sigma_1, \cdots, \sigma_t \succeq_{sos} 0.$$
(7.2.10)

The objective is an affine function of P, and the constraints are either LMIs or SOS inequalities, which are also essentially LMIs ([81]). Therefore it can be solved by a standard SDP routine.

As we can see, when the degree N is higher, the ellipsoid bound by solving E_N is tighter. The convergence of \mathcal{E}_N is described as follows.

Theorem 7.2.3. Suppose the polynomial system 7.2.3 satisfies Putinar's constraint qualification condition (1.1.14). Then the trace p_N^* of the ellipsoid \mathcal{E}_N found in E_N converges to trace p^* of the minimum ellipsoid containing the solution set \mathcal{S} when the degree N tends to infinity.

Proof. Let $\mathcal{E}^* = \{x \in \mathbb{R}^n : (x - z^*)^T (P^*)^{-1} (x - z^*) \leq 1\}$ be the minimum ellipsoid containing the solution set \mathcal{S} , with $trace(P^*) = p^*$. Then for arbitrary $\epsilon > 0$, the polynomial $1 - (x - z^*)^T (P^* + \epsilon I_n)^{-1} (x - z^*)$ is strictly positive on the set of (x, μ) defined by (7.2.3). By Theorem 2.3.2, there exist some general polynomials $\lambda_i(x, \mu)(i = 1, \dots, s)$ and SOS polynomials $\sigma_j(x, \mu)(j = 1, \dots, t)$ such that

$$1 - (x - z^*)^T (P^* + \epsilon I_n)^{-1} (x - z^*) + \sum_{i=1}^s \lambda_i \phi_i - \sum_{j=1}^t \sigma_j \rho_j \succeq_{sos} 0.$$

As we showed previously, problems \hat{E}_N and E_N are equivalent formulations. So they have the same optimal objective values. When N is large enough, then in \hat{E}_N we find one feasible solution with objective value $p^* + n\epsilon$. Thus it must be true that $p_N^* \leq p^* + n\epsilon$. Here n is the dimension of x, which is a constant. Since \mathcal{E}^* is minimum, it holds that $p_N^* \geq p^*$. Therefore we have $\lim_{N\to\infty} p_N^* = p^*$. \Box

Last, let us show some examples. All of them are solved via SOSTOOLS [88].

Example 7.2.4. Consider the following polynomial system of two equations and two inequalities.

$$(1+\mu_1^2)x_1^2 + \mu_2 x_1 x_2 + (1-\mu_2^2)x_2^2 + (\mu_1 + \mu_2)x_1 + (\mu_1 - \mu_2)x_2 - 1 = 0$$
 (7.2.11)

$$(1 - \mu_1^2)x_1^2 + \mu_1 x_1 x_2 + (1 + \mu_2^2)x_2^2 + (\mu_1 - \mu_2)x_1 + (\mu_1 + \mu_2)x_2 - 1 = 0$$
 (7.2.12)

$$\mu_1^2 - \epsilon^2 \le 0, \mu_2^2 - \epsilon^2 \le 0 \quad (7.2.13)$$

where $\epsilon = 0.1$. We formulate the optimization (7.2.8)-(7.2.10) for this polynomial system, and then solve it by SOSTOOLS. In this problem, n = 2, r = 2, D = 4. We choose N = 2 since any nonconstant SOS polynomials have degree at least 2. The resulting 2D-ellipsoid is at the top of Figure 7.3. The asterisks are the solutions (x_1, x_2) when (μ_1, μ_2) are chosen randomly according to the two inequalities. As you can see, the computed ellipsoid is much larger than the set of real solutions. This is because the solution set is not connected.

However, if we want more information about one branch, we can add one more inequality of the form $(x_1-a)^2+(x_2-b)^2 \leq r^2$, where a, b, r are chosen according to the user's interests for the solution region, and then solve the optimization problem again. The role of this new inequality is that it can help to find the ellipsoid bound



Figure 7.3: The ellipsoid for polynomial system (7.2.11)-(7.2.13)

for just one solution component, and it also assures that the Putinar's constraint qualification is satisfied. See Figure 7.4 for the minimum ellipsoid bounds for each component. The left ellipsoid bound is obtained by adding inequality $(x_1 + 0.6)^2 + (x_2 + 0.6)^2 \le 0.6^2$. The right ellipsoid is found by adding inequality $(x_1 - 0.9)^2 + (x_2 - 0.8)^2 \le 0.8^2$.

Example 7.2.5. This example demonstrates how to find a minimum ellipsoid bounding a very elongated set, as indicated in the introduction. Consider the following example:

$$x_1^2 x_2^2 - 2x_1 x_2 + x_2^2 - 3/4 \le 0 (7.2.14)$$

$$x_1^2 - 6x_1 + x_2^2 + 2x_2 - 6 \le 0 (7.2.15)$$

Here n = 2, r = 2, D = 4. We also choose N = 2 as in Example 1. The computed ellipsoid is shown by gray curve in Figure 7.5. The center of the ellipsoid is



Figure 7.4: Ellipsoid bound for each component

 $(4.2970 \ 0.2684)$ and its shape matrix is

$$\begin{bmatrix} 6.6334 & -0.3627 \\ -0.3627 & 0.2604 \end{bmatrix}$$

The short axis is 0.9795 and the long axis is 5.1591. The asterisks are the solutions (x_1, x_2) satisfying the system defined by the above polynomial inequalities. As you can be see, all the asterisks are contained inside the ellipsoid and a few are near the boundary.

7.3 Nearest greatest common divisor

This section discusses the application of minimizing rational polynomials in finding the smallest perturbation of two univariate polynomials that causes them to have a nontrivial GCD, i.e., a common root. We call this probem "finding the nearest GCD" for short.

Let p(z) and q(z) be two monic complex univariate polynomials of degree



Figure 7.5: Ellipsoid bound for polynomial system (7.2.11)-(7.2.15).

 \boldsymbol{m} of the form

$$p(z) = z^m + p_{m-1}z^{m-1} + p_{m-2}z^{m-2} + \dots + p_1z + p_0$$
(7.3.1)

$$q(z) = z^m + q_{m-1}z^{m-1} + q_{m-2}z^{m-2} + \dots + q_1z + q_0.$$
(7.3.2)

Their coefficients p_i, q_j are all complex numbers. When p(z), q(z) have common divisors, their greatest common divisor (GCD) can be computed exactly by using Euclid's algorithm or other refined algorithms [16, 20]. These algorithms assume that all the coefficients of p(z) and q(z) are error-free, and return the exact GCD. However, in practice, it is more interesting to compute the GCD of two polynomials whose coefficients are uncertain. In such situations, we often get the trivial common divisor (the constant polynomial 1) if we apply exact methods like Euclid's algorithm.

For given p(z) and q(z), they may or may not have a common divisor, i.e., a common zero. But we may perturb their coefficients such that the perturbed polynomials have a common divisor, say, z - c. See [48, 49] and [110, §6.4] for a discussion of this problem. The contribution of this section is to solve the associated global optimization problem for rational functions via SOS methods, instead of finding all the real critical points (zero gradient) as suggested in [48, 49].

Throughout this paper, we measure the polynomials p(z), q(z) by $\|\cdot\|_2$ norm of their coefficients, i.e., $\|p\|_2 = \sqrt{\sum_{k=0}^{m-1} |p_k|^2}$, $\|q\|_2 = \sqrt{\sum_{k=0}^{m-1} |q_k|^2}$. The perturbations made to p(z), q(z) are measured similarly. The basic problem in this section is what is the minimum perturbation such that the perturbed polynomials have a common divisor? To be more specific, suppose the perturbed polynomials have the form

$$\hat{p}(z) = z^m + \hat{p}_{m-1} z^{m-1} + \hat{p}_{m-2} z^{m-2} + \dots + \hat{p}_1 z + \hat{p}_0$$
(7.3.3)

$$\hat{q}(z) = z^m + \hat{q}_{m-1} z^{m-1} + \hat{q}_{m-2} z^{m-2} + \dots + \hat{q}_1 z + \hat{q}_0.$$
(7.3.4)

with common zero c, i.e., $\hat{p}(c) = \hat{q}(c) = 0$. The perturbation is measured by

$$\mathcal{N}(c,\hat{p},\hat{q}) = \sum_{i=0}^{m-1} |p_i - \hat{p}_i|^2 + \sum_{j=0}^{m-1} |q_j - \hat{q}_j|^2 = ||p - \hat{p}||_2^2 + ||q - \hat{q}||_2^2.$$

The problem of finding nearest GCD can be formulated as finding (c, \hat{p}, \hat{q}) such that $\mathcal{N}(c, \hat{p}, \hat{q})$ is minimized subject to $\hat{p}(c) = \hat{q}(c) = 0$.

We can see that $\mathcal{N}(c, \hat{p}, \hat{q})$ is a convex quadratic function in (\hat{p}, \hat{q}) . But the constraints $\hat{p}(c) = \hat{q}(c) = 0$ are nonconvex. However, if the common root c is fixed, the constraints $\hat{p}(c) = \hat{q}(c) = 0$ are linear with respect to (\hat{p}, \hat{q}) , and there is a closed form solution. $\mathcal{N}(c, \hat{p}, \hat{q})$ is a convex quadratic function about (\hat{p}, \hat{q}) . It can be shown that [49] that

$$\min_{(\hat{p},\hat{q}):\hat{p}(c)=\hat{q}(c)=0} \quad \mathcal{N}(c,\hat{p},\hat{q}) = \frac{|p(c)|^2 + |q(c)|^2}{\sum_{i=0}^{m-1} |c^2|^i}.$$

Therefore the problem of finding the nearest GCD become the global optimization of the rational function

$$\min_{c \in \mathbb{C}} \quad \frac{|p(c)|^2 + |q(c)|^2}{\sum_{i=0}^{m-1} |c^2|^i}.$$
(7.3.5)

over the complex plane. Karmarkar and Lakshman [49] proposed the following algorithm to find the nearest GCD:

Algorithm 7.3.1 (Nearest GCD Algorithm, [49]).

Input: Monic polynomials p(z), q(z).

Step 1 Determine the rational function $r(x_1, x_2)$

$$r(x_1, x_2) := \frac{|p(c)|^2 + |q(c)|^2}{\sum_{k=0}^{m-1} (x_1^2 + x_2^2)^k}, \ c = x_1 + \sqrt{-1}x_2$$

Step 2 Solve the polynomial system $\frac{r(x_1, x_2)}{\partial x_1} = \frac{r(x_1, x_2)}{\partial x_1} = 0$. Find all its real solutions inside the box: $-B \leq x_1, x_2 \leq B$ where $B := 5 \max(\|p\|^2, \|q\|^2)$. Choose the one (\hat{x}_1, \hat{x}_2) such that $r(\hat{x}_1, \hat{x}_2)$ is minimum. Let $c := \hat{x}_1 + \sqrt{-1}\hat{x}_2$.

Step 3 Compute the coefficient perturbations

as

$$\lambda_j := \frac{\bar{c}^j p(c)}{\sum_{k=0}^{m-1} |c^2|^k}, \quad \mu_j := \frac{\bar{c}^j q(c)}{\sum_{k=0}^{m-1} |c^2|^k}.$$

Output: The minimum perturbed polynomials with common divisors are returned

$$\hat{p}(z) = z^m + \sum_{k=0}^{m-1} (p_k - \lambda_k) z^k, \ \hat{q}(z) = z^m + \sum_{k=0}^{m-1} (q_k - \mu_k) z^k.$$

The most expensive part in the algorithm above is step 2. Karmarkar and Lakshman [49] proposed to use numerical methods like Arnon and McCallum [2] or Manocha and Demmel [60] to find all the real solutions of a polynomial system inside a box.

However, in practice, it is very expensive to find all the real solutions of a polynomial system inside a box, although a polynomial complexity bound exists as stated in [49]. So in this section, we propose to solve (7.3.5) by SOS relaxations introduced in Chapter 6, instead of finding all the real solutions of a polynomial system. The SOS relaxation of problem (7.3.5) is the following:

$$\begin{split} \sup & \gamma \\ s.t. & f(x_1,x_2) - \gamma(\sum_{i=0}^{m-1}(x_1^2+x_2^2)^i) \text{ is SOS} \end{split}$$
 where $f(x_1,x_2) = |p(x_1+\sqrt{-1}x_2)|^2 + |q(x_1+\sqrt{-1}x_2)|^2.$

In the following examples, we solve this optimization problem via SOS relaxation (6.1.4)-(6.1.6) and its dual (6.1.7)-(6.1.9). In all the examples here, the global minimizers can be extracted and the big ball technique introduced in Section 6.2 is not required.

Example 7.3.2 (Example 2.1,[49]). Consider the following two polynomials

$$p(z) = z^{2} - 6z + 5$$
$$q(z) = z^{2} - 6.30z + 5.72.$$

Solving SOS relaxation (6.1.4)-(6.1.6) and its dual (6.1.7)-(6.1.9), we find the global minimum and extract one minimizer:

$$r^* \approx 0.0121, \quad c^* = x_1^* + \sqrt{-1}x_2^* \approx 5.0971.$$

which are the same as found in [49].

Example 7.3.3. Consider the following two polynomials

$$p(z) = z^{3} - 6z^{2} + 11z - 6$$
$$q(z) = z^{3} - 6.24z^{2} + 10.75z - 6.50.$$

Solving SOS relaxation (6.1.4)-(6.1.6) and its dual (6.1.7)-(6.1.9), , we get a lower bound and extract one point

$$r_{sos}^* \approx 0.0563, \quad (x_1^*, x_2^*) \approx (3.5725, 0.0000).$$

Evaluation of r(x) at x^* shows that $r(x^*) \approx r^*_{sos}$, which implies that $c^* \approx 3.5725$ is a global minimizer for problem (7.3.5).

Example 7.3.4. Consider the following two polynomials

$$p(z) = z^{3} + z^{2} - 2$$
$$q(z) = z^{3} + 1.5z^{2} + 1.5z - 1.25.$$

Solving SOS relaxation (6.1.4)-(6.1.6) and its dual (6.1.7)-(6.1.9), we find the lower bound $r_{sos}^* \approx 0.0643$ and extract two points

$$x^* \approx (-1.0032, 1.1011)$$
 $x^{**} \approx (-1.0032, -1.1011).$

The evaluations of r(x) at x^* and x^{**} show that $r(x^*) = r(x^{**}) \approx r_{sos}^*$, which implies that x^* and x^{**} are both global minimizers. So $c^* = -1.0032 \pm \sqrt{-1} \cdot 1.1011$ are the global minimizers of problem (7.3.5).

7.4 Maximum likelihood optimization

This section discusses another application of polynomial optimization. One important class of problems in statistics and computational biology is maximum likelihood optimization. It can be formulated as

$$\max_{x \in \mathbb{R}^n} \quad \prod_{i=1}^{\prime} f_i(x)^{m_i} \tag{7.4.1}$$

s.t.
$$g_1(x), \cdots, g_\ell(x) \ge 0$$
 (7.4.2)

where $f_i(x), g_j(x)$ are all polynomials in $x \in \mathbb{R}^n$, and m_i are positive integers. Here we assume that each $f_i(x)$ is nonnegative on the feasible set, which is often the case in statistics or computational biology (e.g., $f_i(x)$ represents some probability distribution).

Our goal is to find the global or approximately global solution to (7.4.1)-(7.4.2). Theoretically, the SOS methods can be applied in this problem, since the objective and constraints are all described by polynomials. However, in practice, the exponents m_i are big. It is very common that these integers are hundreds or even thousands. Then SOS methods are too expensive to be implemented, because the reduced SDP is too huge to be solved. So we need cheaper methods, and still want high quality solutions (e.g., approximately global).

Without changing the problem, we take the log of the objective in (7.4.1)-(7.4.2) and get an equivalent problem:

$$\max_{x \in \mathbb{R}^n} \quad \sum_{i=1}^r m_i \log f_i(x) \tag{7.4.3}$$

s.t.
$$g_1(x), \cdots, g_\ell(x) \ge 0.$$
 (7.4.4)

However, the objective is no longer a polynomial, and hence SOS methods can not be applied directly. But moment matrix methods are still applicable.

Suppose $f_i(x)$ has the form $f_i(x) = \sum_{\alpha \in P_i} f_{i,\alpha} x^{\alpha}$, where P_i is the support.

Then we can see that

$$f_i(x) = \sum_{\alpha} f_{i,\alpha} y_{\alpha}$$
 when $y = mon_N(x)$

where $N \ge \deg(f)$. We can also see that $g_j(x) \ge 0$ is the same as

$$g_j(x) \cdot m_{N-d_i}(x) \cdot m_{N-d_i}(x)^T \succeq 0$$

where $d_i = \lceil \deg(g_i)/2 \rceil$.

If we replace each x^{α} by y_{α} , we get the following relaxation:

$$\max_{y=(y_{\alpha})} \sum_{i=1}^{r} m_{i} \log(\sum_{\alpha} f_{i,\alpha} y_{\alpha})$$

s.t. $M_{N-d_{j}}(g_{j} * y) \succeq 0$
 $M_{N}(y) \succeq 0.$

This is still a convex optimization problem, and efficient techniques like interior-point methods are available. Let y^* be the optimal solution to this problem. When the moment matrix $M_N(y^*)$ satisfies the *flat extension* condition, we can extract the maximizer(s) x^* . A very simple choice is $x_i^* = y_{e_i}^*$.

7.5 Sensor network localization

This section shows the application of sum of squares in sensor network localization. The basic description of this problem is as follows. For a sequence of unknown vectors (also called *sensors*) x_1, x_2, \dots, x_n in the Euclidean space $\mathbb{R}^d(d =$ $1, 2, \dots)$, we need find their coordinates such that the distances (not necessarily all) between these sensors and the distances (not necessarily all) to other fixed sensors a_1, \dots, a_m (they are also called *anchors*) are equal to some given numbers. To be more specific, let $\mathcal{A} = \{(i, j) \in [n] \times [n] : ||x_i - x_j||_2 = d_{ij}\}$, and $\mathcal{B} = \{(i, k) \in [n] \times [m] : ||x_i - a_k||_2 = e_{ik}\}$, where d_{ij}, e_{ik} are given distances and $[n] = \{1, 2, \dots, n\}$. Then the problem of sensor network localization is to find vectors $\{x_1, x_2, \dots, x_n\}$ such that $||x_i - x_j||_2 = d_{ij}$ for all $(i, j) \in \mathcal{A}$ and $||x_i - a_k||_2 = e_{ik}$ for all $(i, k) \in \mathcal{B}$.

This task can be formulated as an optimization problem. Let x_1, \dots, x_n be decisions variables, each s_i being a vector in \mathbb{R}^d . Obviously, x_1, \dots, x_m provides the right sensor locations if and only if the optimal value of problem

$$\min_{x_1, \cdots, x_n \in \mathbb{R}^d} \quad \sum_{(i,j) \in \mathcal{A}} \left| \|x_i - x_j\|_2^2 - d_{ij}^2 \right| + \sum_{(i,k) \in \mathcal{B}} \left| \|x_i - a_k\|_2^2 - e_{ik}^2 \right|^2.$$

is zero. This optimization problem is nonconvex, and it is often NP-hard to find global solutions. So approximation methods are of great interest. For example, SDP or second-order cone programming (SOCP) relaxations can be applied to solve the problem approximately. We refer to [10, 106, 112] for work in this area. However, SDP relaxation is very expensive to implement for large problems (e.g., more than 100 sensors). SOCP relaxation is weaker than SDP relaxation, but can solve larger problems.

As we can see, the objective in the above involves absolute values and is not a polynomial. Hence, SOS methods can not be applied. However, if we replace the absolute value by squares, we can get a new optimization problem

$$f(X) := \min_{X = [x_1, \cdots, x_n] \in \mathbb{R}^{d \times n}} \sum_{(i,j) \in \mathcal{A}} \left(\|x_i - x_j\|_2^2 - d_{ij}^2 \right)^2 + \sum_{(i,k) \in \mathcal{B}} \left(\|x_i - a_k\|_2^2 - e_{ik}^2 \right)^2.$$

The good property of this new objective is that it is a quartic polynomial. Therefore, the method of sum of squares is applicable. On the other hand, we must be very careful in applying SOS method to solve this polynomial optimization problem. The total number of decision variables is $n \cdot d$. If we apply SOS relaxation directly, the size of matrix in reduced SDP is $\binom{n \cdot d+4}{4}$, which can be huge for even moderate n and d. For instance, when n = 50and d = 2, this number is as large as

$$\binom{n \cdot d + 4}{4} \ge 10^4.$$

So it seems that the applications of SOS relaxations are very limited in practice. So we can maximize γ such that

$$f(X) - \gamma \equiv \sum_{(i,j) \in \mathcal{A}} \sigma_{ij}(x_i, x_j)$$

where $\sigma_{ij}(x_i, x_j)$ is some SOS polynomial in (x_i, x_j) . If we use this special representation, we can efficiently and accurately solve large scale sensor network localization problems that can not be solved by SDP relaxation. See the following example.

Example 7.5.1. We randomly generate test problems which are similar to those given in [11]. First, we generate n = 500 points x_1^*, \dots, x_n^* from the unit square $[-0.5, 0.5] \times [-0.50.5]$. Choose anchors to be four points $(\pm 0.45, \pm 0.45)$. The edge set \mathcal{A} is chosen as follows. Initially set $\mathcal{A} = \emptyset$. Then for each *i* from 1 to 500, compute the set $I_i = \{j \in [500] :, ||x_i^* - x_j^*||_2 \leq 0.3, j \geq i\}$; if $|I_i| \geq 10$, let A_i the subset of I_i consisting of the 10 smallest integers; otherwise, let $A_i = I_i$; then let $\mathcal{A} = \mathcal{A} \cup \{(i, j) : j \in A_i\}$. The edge set \mathcal{B} is chosen such that $\mathcal{B} = \{(i, k) \in [n] \times [m] : ||x_i^* - a_k||_2 \leq 0.3\}$, i.e., every anchor is connected to all the sensors that are within distance 0.3. For every $(i, j) \in \mathcal{A}$ and $(i, k) \in \mathcal{B}$, let the distances be

$$d_{ij} = \|x_i^* - x_j^*\|_2, \quad e_{ik} = \|x_i^* - a_k\|_2.$$

There are no errors in the distances. The computed results are plotted in Figure 7.6. The true sensor locations (denoted by circles) and the computed locations (denoted by stars) are connected by solid lines.



Figure 7.6: 500 sensors, sparse SOS relaxation

From Figure 7.6, we find that all the stars are located inside circles, which implies that SOS relaxation provides high quality locations. The accuracy of the estimated points $\hat{x}_1, \dots, \hat{x}_n$ will be measured by the Root Mean Square Distance (RMSD) which is defined as

RMSD =
$$\left(\frac{1}{n}\sum_{i=1}^{n} \|\hat{x}_i - x_i^*\|_2^2\right)^{\frac{1}{2}}$$
.

The RMSD for this sparse SOS relaxation is $2.9 \cdot 10^{-6}$ (the computed locations will be exact if we ignore rounding errors involved in floating point operations). The interior-point method in *SeDuMi* consumes about 1079 CPU seconds (18 minutes). We generate this random examples 20 times. Every time the RMSD is in the order $O(10^{-6})$ and the CPU time consumed by the sparse SOS relaxation is almost the same.

We refer to [74] for the sparse SOS relaxation for sensor network localization problem.

Bibliography

- B. Alkire and L. Vandenberghe, Convex optimization problems involving finite autocorrelation sequences. Mathematical Programming Series A 93 (2002), 331-359.
- [2] D. Arnon and S. McCallum. A polynomial time algorithm for the topological type of a real algebraic curve. *Journal of Symbolic Computation* 5, 213-236.
- [3] M. Avriel. Nonlinear Programming: Analysis and Methods, Dover Pubns, 2003.
- [4] C. Berg. The multidimensional moment problem and semi-groups, in Moments in Mathematics, H.J. Landau, ed., AMS, Providence, RI, 1980, pp. 110-124.
- [5] D. Bertsekas. Nonlinear Programming, Athena Scientific, 1995, (2nd Edition, 1999).
- [6] S. Basu, R. Pollack and M-F. Roy. Algorithms in Real Algebraic Geometry, Springer, 2003.
- [7] M. Bazaraa, C. Shetty and H. Sherali. Nonlinear Programming: Theory and Algorithms (2nd edition), John Wiley & Sons Inc, 1993.

- [8] E. Becker and R. Neuhaus. Computation of real radicals of polynomial ideals. *Computational algebraic geometry* (Nice, 1992), 1–20, Progress in Mathematics, 109, Birkhäuser, Boston, MA, 1993.
- [9] E. Becker and R. Neuhaus. Computation of real radicals of polynomial ideals.
 II. J. Pure Appl. Algebra 124 (1998) 261–280.
- [10] P. Biswas and Y. Ye. Semidefinite programming for ad hoc wireless sensor network localization. Proc. 3rd IPSN 46-54, 2004.
- [11] P. Biswas, T.C. Liang, K.C. Toh, T.C. Wang and Y. Ye. Semidefinite Programming Approaches for Sensor Network Localization with Noisy Distance Measurements. To appear in to appear in *IEEE Transactions on Automation Science and Engineering*. Website:http://www.stanford.edu/~yyye/.
- [12] G. Blekherman. Volumes of nonnegative polynomials, sums of squares, and powers of linear forms, preprint, arXiv:math.AG/0402158.
- [13] J. Bochnak, M. Coste and M-F. Roy. Real Algebraic Geometry, Springer, 1998.
- [14] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. Linear Matrix Inequalities in System and Control Theory, SIAM, 1994.
- [15] S. Boyd and L. Vandenberghe. Convex optimization, Cambridge University Press, 2003.
- [16] W. Brown and J. Traub. On Euclid's algorithm and the theory of subresultants. J. ACM, 18, 505-514.
- [17] Frank M. Callier, Charles A. Desoer, *Linear System Theory*, Springer-Verlag, New York, 1991.
- [18] G. Calafiore and L. El Ghaoui. Confidence Ellipsoids for Uncertain Linear Equations with Structure. Proc. Conf. Decision and Control, December 1999.
- [19] G. Calafiore and L. El Ghaoui. Ellipsoid Bounds for Uncertain Linear Equations and Dynamical Systems, to appear in *Automatica*, http://robotics.eecs. berkeley.edu/~elghaoui
- [20] G. Collins. Subresultants and reduced polynomial remainder sequences. J. ACM, 14, 128-142.
- [21] D.A. Cox, J.B. Little, and D.O'Shea. Ideals, Varieties and Algorithms: an Introduction to Computational Algebraic Geometry and Commutative Algebra, Second Edition. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1997.
- [22] D.A. Cox, J.B. Little, and D.O'Shea. Using Algebraic Geometry, Graduate Texts in Mathematics, Vol. 185. Springer-Verlag, New York, 1998.
- [23] R.M. Corless, P.M. Gianni, and B.M. Trager. A reorder Schur factorization method for zero-dimensional polynomial systems with multiple roots. *Proc. ACM Int. Symp. Symbolic and Algebraic Computation*, 133-140, Maui, Hawaii, 1997.
- [24] R.E. Curto and L.A. Fialkow. The truncated complex K-moment problem. Trans. Amer. Math. Soc. 352 (2000) 2825–2855.

- [25] E. De Klerk. Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications. *Applied Optimization Series*, Volume 65. Kluwer Academic Publishers, March 2002, 300 pp., ISBN 1-4020-0547-4.
- [26] E. De Klerk, M. Laurent and P. Parrilo. A PTAS for the minimization of polynomials of fixed degree over the simplex. To appear in *Theoretical Computer Science*.
- [27] C. Delzell and A. Prestel, *Positive Polynomials*, Monographs in Mathematics, Springer, Berlin, 2001.
- [28] J. Demmel. Applied numerical linear algebra. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.
- [29] D. Eisenbud, Commutative Algebra with a View Toward Algebraic Geometry, Graduate Texts in Mathematics, Vol. 150. Springer-Verlag, New York, 1995.
- [30] D. Eisenbud, with Daniel R. Grayson, Michael Stillman, and Bernd Sturmfels (Eds.)). Computations in Algebraic Geometry with Macaulay 2, Algorithms and Computation in Mathematics. Springer-Verlag, 2002.
- [31] L. Faybusovich, On Nesterov's approach to semi-infinite programming, Acta Applicandae Mathematicae 74 (2002), 195-215.
- [32] C.A. Floudas and P.M. Pardalos. A collection of test problems for constrained global optimization algorithms, Lecture Notes in Computer Science, 455, Springer-Verlag, Berlin, 1990.
- [33] E. Fortuna, P. Gianni, and B. Trager: Derivations and radicals of polynomial

ideals over fields of arbitrary characteristic. Computer algebra (London, ON, 2001). J. Symbolic Comput. **33** (2002), no. 5, 609–625.

- [34] M. R. Garey and D. S. Johnson. Computers and Intractability: A guide to the theory of NP-completeness. W. H. Freeman and Company, 1979.
- [35] Y. Genin, Y. Hachez, Yu. Nesterov, P. Van Dooren, "Convex Optimization over Positive Polynomials and filter design", *Proceedings UKACC Int. Conf. Control* 2000, page SS41, 2000.
- [36] Y. Genin, Y. Hachez, Yu. Nesterov, P. Van Dooren, "Optimization problems over positive pseudopolynomial matrices", SIAM Journal on Matrix Analysis and Applications 25 (2003), 57-79.
- [37] D. Grigoriev and N. N. Vorobjov, Jr. Solving systems of polynomial inequalities in subexponential time. J. Symbolic Comput., 5(1-2):37-64, 1988.
- [38] D. Handelman. Representing polynomials by positive linear functions on compact convex polyhedra. Pac. J. of Math. 132, No.1, 35-62 (1988).
- [39] B. Hanzon and D. Jibetean: Global minimization of a multivariate polynomial using matrix methods. *Journal of Global Optimization*, 27:1-23, 2003.
- [40] D. Henrion and J. Lasserre. GloptiPoly: Global optimization over polynomials with Matlab and SeDuMi. ACM Trans. Math. Soft., 29:165-194, 2003.
- [41] D. Henrion and J. Lasserre. Detecting global optimality and extracting solutions in GloptiPoly. In *Positive Polynomials in Control*, D. Henrion and A. Garulli, eds., Lecture Notes on Control and Information Sciences, Springer Verlag, 2005.

- [42] J. W. Helton. "Positive" noncommutative polynomials are sums of squares. Ann. of Math. (2) 156 (2002), no. 2, 675–694.
- [43] R. Horst, P. Pardalos and N. Thoai. Introduction to global optimization, second edition, Kluwer Academic Publishers, 2000.
- [44] T. Jacobi and A. Prestel. Distinguished representations of strictly positive polynomials. J. Reine Angew. Math. 532, 223–235 (2001)
- [45] D. Jibetean and M. Laurent. Semidefinite approximations for global unconstrained polynomial optimization. To appear in the SIAM Journal on Optimization. Website: http://www.cwi.nl/~monique.
- [46] T.Kailath and A.H. Sayed, "Displacement Structure: theory and applications", SIAM Rev. 37(1995), 297-386.
- [47] S. Kim, M. Kojima and H. Waki. Generalized Lagrangian Duals and Sums of Squares Relaxations of Sparse Polynomial Optimization Problems. To appear in SIAM Journal on Optimization, September 2003. Revised July 2004.
- [48] N. Karmarkar and Y. Lakshman. Approximate polynomial greatest common divisors and nearest singular polynomials. Proceedings ISSAC'96 (Zurich, Switzerland). New York, ACM Press.
- [49] N. Karmarkar and Y. Lakshman. On approximate GCDs of univariate polynomials. Journal of Symbolic Computation (1998) 26, 653-666.
- [50] M. Kojima, S. Kim and H. Waki. Sparsity in Sums of Squares of Polynomials. June 2003, Revised June 2004. *Mathematical Programming* Vol.103 (1) 45-62.

- [51] T. Krick and A. Logar. An algorithm for the computation of the radical of an ideal in the ring of polynomials. Applied algebra, algebraic algorithms and error-correcting codes (New Orleans, LA, 1991), 195–205, Lecture Notes in Comput. Sci., 539, Springer, Berlin, 1991.
- [52] J. Lasserre. Global optimization with polynomials and the problem of moments. SIAM J. Optim., 11 (2001), No. 3, 796–817.
- [53] J. Lasserre. A sum of squares approximation of nonnegative polynomials. SIAM J. Optim., Vol. 16 (2006), No. 3, 751-765.
- [54] J. Lasserre. S.o.s. approximation of polynomials nonnegative on a real algebraic set. LAAS report #04480; August 2004.
- [55] M. Laurent. Semidefinite representations for finite varieties. To appear in Math. Prog., 2004. Website: http://www.cwi.nl/~monique.
- [56] M. Laurent. Moment matrices and optimization over polynomials A survey on selected topics. Preprint, September 2005. Website: http://www.cwi.nl/ ~monique.
- [57] J. Löfberg. YALMIP: A MATLAB toolbox for rapid prototyping of optimization problems. http://control.ee.ethz.ch/~joloef/yalmip.php.
- [58] D. Luenberger. Linear and Nonlinear Programming (2nd edition), Kluwer Academic Publisher, 2003.
- [59] Lukacs. Verschäarfung der ersten Mittelwersatzes der Integralrechnung für rationale Polynome, Math. Zeitschrift, 2, 229-305, 1918.

- [60] D. Manocha and J. Demmel. Algorithms for interesting parametric and implicit curves I: simple intersections. ACM Trans. Graphics, 13, 73-100.
- [61] A.A. Markov. Lecture notes on functions with the least deviation from zero, 1906. Reprinted in *Markov A.A. Selected Papers* (ed. N. Achiezer), GosTechIzdat, 244-291, 1948, Moscow(in Russian).
- [62] M. Marshall. Optimization of polynomial functions, Canad. Math. Bull., 46 (2003) 575–587.
- [63] M. Marshall: Representation of non-negative polynomials with finitely many zeros, to appear in Annales de la Faculté des Sciences de Toulouse http://math.usask.ca/~marshall/
- [64] K.G. Murty and S.N.Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39:117-129, 1987.
- [65] S. Nash and A. Sofer. *Linear and Nonlinear Programming*, McGraw-Hill Book Co Ltd, 2000.
- [66] Yu. Nesterov and A. Nemirovsky, "interior-point polynomial algorithms in convex programming", SIAM Studies in Applied Mathematics, vol. 13, Society of Industrial and Applied Mathematics(SIAM), Philadelphia, PA, 1994.
- [67] Y. Nesterov. Squared functional systems and optimization problems. *High Per*formance Optimization (H.Frenk et al., eds), Kluwer Academic Publishers, 2000, pp.405-440.
- [68] Y. Nesterov and A. Nemirovskii. Interior-point polynomial algorithms in convex programming, SIAM, 1994.

- [69] J. Nie and J. W. Demmel. Shape Optimization of Transfer Functions. *Multiscale optimization methods and applications*, eds. W. Hager, P. Pardalos, S. Huang, etc., p. 313-326, Springer series on nonconvex optimization and its application series, 2005.
- [70] J. Nie and J. W. Demmel. Minimum ellipsoid bounds for solutions of polynomial systems via sum of squares. *Journal of Global Optimization* (2005) 33: 511-525.
- [71] J. Nie, J. Demmel and B. Sturmfels. Minimizing Polynomials via Sum of Squares over the Gradient Ideal. *Mathematical Programming, Series A*, Vol. 106 (2006), No. 3, 587-606.
- [72] J. Nie and M. Schweighofer. On the complexity of Putinar's Positivstellensatz. Preprint. arXiv: math.AG/0510309.
- [73] J. Demmel, J. Nie and V. Powers. Representations of Positive Polynomials on Non-compact Semialgebraic Sets via KKT Ideals. Preprint, 2005. Website: http://www.math.berkeley.edu/~njw.
- [74] J. Nie. Sum of squares methods for sensor network localization. Preprint, 2006. arXiv:math.OC/0605652.
- [75] Jorge Nocedal and Stephen J. Wright. Numerical Optimization, Springer Series in Operations Research, Springer-Verlag, New York, 1999.
- [76] P. Pardalos and J. Rosen. Constrained Global Optimization: Algorithms and Applications, Springer-Verlag, Lecture Notes in Computer Science 268 (1987).

- [77] C. Floudas and P. Pardalos. Collection of Test Problems for Constrained Global Optimization Algorithms, Springer-Verlag, Lecture Notes in Computer Science 455 (1990)
- [78] R. Horst, P. Pardalos and N. Thoai. Introduction to Global Optimization, Second Edition, Nonconvex optimization and its applications, Vol. 48, Kluwer Academic Publishers, 2000.
- [79] P. Parrilo. Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization, Ph.D Thesis, California Institute of Technology, 2000.
- [80] P. Parrilo and B. Sturmfels. Minimizing polynomial functions, Proceedings of the DIMACS Workshop on Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science (March 2001), (eds. S. Basu and L. Gonzalez-Vega), American Mathematical Society, 2003, pp. 83– 100.
- [81] P. Parrilo. Semidefinite Programming relaxations for semialgebraic problems. Mathematical Programming, Ser. B 96 (2003), No. 2, 293–320.
- [82] E. Polak. "Optimization: Algorithms and Consistent Approximations". Applied Mathematical Sciences, Vol. 124, Springer, New York, 1997.
- [83] G. Pólya. Collected Papers, Volume 2, 309-313, Cambridge, MIT Press (1974).
- [84] G. Pólya and G. Szegö, Problems and Theorems in Analysis II, Springer-Verlag, New York, 1976

- [85] Victoria Powers and Bruce Reznick, "Polynomials That are Positive on an Interval", Transactions of the American Mathematical Society, vol. 352, No. 10, pp. 4677-4692, 2000.
- [86] V. Powers and B. Reznick. Polynomials positive on unbounded rectangles. In *Positive Polynomials in Control*, Springer Lecture Notes in Control and Information Sciences, Vol. 312, 2005.
- [87] V. Powers and B. Reznick. A new bound for Pólya's Theorem with applications to polynomials positive on polyhedra, J. Pure Appl. Algebra 164, No. 1-2, 221-229, 2001.
- [88] S. Prajna, A. Papachristodoulou and P. Parrilo. SOSTOOLS User's Guide.
 Website: http://www.mit.edu/~parrilo/SOSTOOLS/.
- [89] P. Parrilo. An explicit construction of distinguished representations of polynomials nonnegative over finite sets, IfA Technical Report AUT02-02, March 2002.
- [90] A. Prestel. Bounds for representations of polynomials positive on compact semialgebraic sets. *Fields Inst. Commun.* 32, 253–260 (2002)
- [91] M. Putinar. Positive polynomials on compact semi-algebraic sets, Ind. Univ. Math. J. 42 (1993) 203–206.
- [92] M. Putinar. On Hermitian polynomials optimization. Preprint., Departement of Mathematics, University of California at Santa Babara, 2005.
- [93] J. Renegar. On the computational complexity and geometry of the first-order

theory of the reals. Part I-III: Introduction and Preliminaries. J. Symbolic Comput., 13(3):255-352, 1992.

- [94] J. Renegar. Mathematical View of Interior-Point Methods in Convex Optimization, SIAM, 2001.
- [95] B. Reznick. Some concrete aspects of Hilbert's 17th problem. In *Contempo*rary Mathematics, volume 253, pages 251-272. American Mathematical Society, 2000.
- [96] R. Rockafellar. Convex Analysis. Princeton University Press, second edition, 1970.
- [97] C. Scheiderer. Distinguished representations of non-negative polynomials. Journal of Algebra 289 (2005), 558-573.
- [98] K. Schmüdgen. The K-moment problem for compact semialgebraic sets, Math. Ann. 289 (1991), 203–206.
- [99] M. Schweighofer. An algorithmic approach to Schmüdgen's Positivstellensatz, J. Pure Appl. Algebra 166, No. 3, 307–319 (2002)
- [100] Markus Schweighofer. On the complexity of Schmüdgen's Positivstellensatz, Journal of Complexity 20, 529-543, 2004.
- [101] Markus Schweighofer. Optimization of polynomials on compact semialgebraic sets. SIAM Journal on Optimization 15, No. 3, 805-825, 2005.
- [102] M. Schweighofer. Certificates for nonnegativity of polynomials with zeros on

compact semialgebraic sets. *Manuscripta Mathematica* **117**, No. 4, 407 - 428 (2005)

- [103] M. Schweighofer. Global optimization of polynomials using gradient tentacles and sums of squares. Preprint, 2005. Website: http://www.math. uni-konstanz.de/~schweigh
- [104] Shafarevich. Basic algebraic geometry. Die Grundlehren der mathematischen Wissenschaften. Band 213. Springer-Verlag, 1974.
- [105] N.Z. Shor. Quadrtic Optimization Problems. Soviet J. Comput. Systems Sci., 25 (1987), pp. 1-11.
- [106] A. Man-cho So and Y. Ye. The theory of semidefinite programming for sensor network localization. To appear in *Math. Prog.*. Website: http://www. stanford.edu/~yyye.
- [107] A. J. Sommese and C. W. Wampler. The Numberical Solution of Systems of Polynomials, Singapore: World Scientific, 2005.
- [108] G. Stengle. A nullstellensatz and a positivstellensatz in semi-algebraic geometry. Math. Ann. 207, 87-97 (1974).
- [109] G. Stengle, Complexity estimates for the Schmüdgen Positivstellensatz. J. Complexity 12(1996), 167-174. MR 97d:14080
- [110] H. Stetter. Numerical Polynomial Algebra. Society of Industrial and Applied Mathematics, 2004.

- [111] J.F. Sturm. SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, Optimization Methods and Software, 11&12(1999)625-653.
- [112] P. Tseng. Second-order cone programming relaxation of sensor network localization, August, 2005, submitted to SIAM J. Optim.. Website: http: //www.math.washington.edu/~tseng
- [113] L. Vandenberghe and S. Boyd. Semidefinite Programming, SIAM Review 38 (1996) 49-95.
- [114] H. Waki, S. Kim, M. Kojima and M. Muramatsu. Sums of Squares and Semidefinite Programming Relaxations for Polynomial Optimization Problems with Structured Sparsity. October 2004. Revised February 2005. website: http://www.is.titech.ac.jp/~kojima/sdp.html.
- [115] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. Handbook of semidefinite programming. Kluwer's Publisher, 2000.
- [116] S.-P. Wu, S. Boyd, and L. Vandenberghe, "FIR filter design via spectral factorization and convex optimization", *Applied and Computational Control, Signals* and Circuits, B. Datta, ed., Birkhäuser, 1997, ch.2, pp.51-81.

Glossary of Notations

- $\mathbb{R}:$ the field of real numbers
- $\mathbb{C}:$ the field of complex numbers
- $\mathbb{N}:$ the set of nonnegative integers
- \mathbb{Z} : the ring of integers
- \mathbb{R}^n : Euclidean Space of dimension n
- \mathbb{R}^n_+ : The nonnegative orthant of \mathbb{R}^n
- $\mathbb{R}[x]$: the ring of polynomials in (x_1, \cdots, x_n)
- $\mathbb{R}^{m\times n} {:}$ the vector space of matrices with dimension $m\times n$
- \mathcal{S}^n : the vector space of symmetric matrices with size n
- $A \succeq 0$: the symmetric matrix A is positive semidefinite
- $A \succ 0$: the symmetric matrix A is positive definite

- $A \bullet B$: the inner product of two matrices of same dimensions defined as trace $(A^T B)$
- \mathcal{S}^n_+ : the positive semidefinite cone of \mathcal{S}^n
- \mathcal{S}_{++}^n : the positive definite interior of \mathcal{S}_{+}^n
- $\Sigma \mathbb{R}[x]^2$: the cone of SOS polynomials
- $x = (x_1, \cdots, x_n)$: a *n*-dimensional vector
- $p(x)\colon$ polynomial evaluated at the vector $x\in\mathbb{R}^n$

deg(p): the degree of polynomial p(x)

supp(p): the support of polynomial p(x)

 $p(x) \succeq q(x)$: the polynomial p(x) - q(x)is SOS

GCD: greatest common divisors

S: a basic closed semialgebraic set

- $\mathcal{M}(S)$: the quadratic module associated with S
- \mathcal{P}_{KKT} : the preorder cone associated with KKT system
- \mathcal{M}_{KKT} : the quadratic module associated with KKT system
- $\mathcal{M}(S)_N$: the subset of $\mathcal{M}(S)$ with de-

gree at most N in each summand

- $\mathcal{M}(y)$: the moment matrix induced by multi-indexed vector y
- $\mathcal{M}_N(y)$: the $\binom{N+n}{n}$ -th leading submatrix of M(y)
- $\mathcal{M}_N(g * y)$: the moment matrix induced by multi-indexed vector $\hat{y} = (\sum_{\beta} g_{\beta} y_{\alpha+\beta})$ where $g(x) = \sum_{\beta} g_{\beta} x^{\alpha}$.