1 Introduction

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

\[
\min f(x) \quad \text{s.t.} \quad x \in S
\]

(1.1)

where \( f(x) \) is a real multivariate polynomial in \( x \in \mathbb{R}^n \) and \( S \) is a feasible set defined by polynomial equalities or inequalities. In this thesis, we do not have any convexity/concavity assumptions on \( f(x) \) or \( S \). The goal is to find the global minimum and global minimizers if any.

Polynomial optimization of form (1.1) is quite general in practical applications. Many NP-hard problems like max cut, matrix copositivity, nonconvex QCQP, partitioning, etc, can be formulated in form of (1.1). Recently, there has been a great deal of work in using sum of squares (SOS) relaxations (also called semidefinite programming (SDP) relaxations in some literature) to solve (1.1). The basic idea is to approximate nonnegative polynomials by sum of squares polynomials, and then relax the intractable nonconvex problem (1.1) to a tractable convex problem which is equivalent to some SDP.

A polynomial \( p(x) \in \mathbb{R}[x_1, \ldots, x_n] \) is nonnegative if \( p(x) \geq 0 \) for all \( x \in \mathbb{R}^n \). A polynomial \( p(x) \in \mathbb{R}[x_1, \ldots, x_n] \) is sum of squares (SOS) if \( p(x) = \sum q_i^2(x) \) for some polynomials \( q_i(x) \in \mathbb{R}[x_1, \ldots, x_n] \). Obviously, if \( p(x) \) is SOS, then \( p(x) \) must be nonnegative. For example, the 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. However, not every nonnegative polynomial is SOS. For instance, the Motzkin polynomial

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

is nonnegative but not SOS. The relation between nonnegative and SOS polynomials is characterized by Hilbert as follows: every nonnegative polynomial having \( n \) variables and degree \( d \) is SOS if and only if \( n = 1 \), or \( d = 2 \), or \( (n, d) = (2, 4) \).

Typically it is intractable to check whether a given polynomial is nonnegative or not. However, it is tractable to check whether a polynomial is SOS or not, which can be done by solving an SDP problem. Let \( p(x) \) be a polynomial of degree \( 2d \). It can be easily shown that \( p(x) \) is SOS if and only if there exists a real symmetric matrix \( X \succeq 0 \) (\( X \) is positive semidefinite) such that

\[
p(x) \equiv m_d(x)^T X m_d(x).
\]

(1.2)

Here \( m_d(x) \) is the column vector of all monomials up to degree \( d \). By comparing coefficients on both sides of (1.2), we know (1.2) is equivalent to an SDP feasibility problem about \( X \):

\[
X \succeq 0, \quad \text{Trace}(A_i X) = b_i, \quad i = 1, \ldots, m.
\]

Here \( A_1, \ldots, A_m \) are constant symmetric matrices, and \( b_1, \ldots, b_m \) are constant scalars. So checking SOS is essentially an SDP problem, which can be solved by numerical algorithms efficiently. There has been extensive work on designing efficient and robust SDP solvers.
Contributions: The contributions of this thesis are in the following aspects: (i) give quantitative convergence analysis of standard SOS methods like Lasserre’s relaxations; (ii) design new numerical methods for solving polynomial optimization which is significantly better in both approximation quality and computational efficiency, like SOS methods based on gradient and KKT ideals; (iii) develop new applications of polynomial optimization, like shape design of transfer functions, perturbation analysis of polynomial systems and sensor network localization.

2 On the Convergence Rate of Lasserre’s Relaxations

A general polynomial optimization has the form

\[
\begin{align*}
\min_{f(x)} & \quad f(x) \\
\text{s.t.} & \quad g_1(x) \geq 0, \ldots, g_m(x) \geq 0
\end{align*}
\]  

(2.1)

where \( f(x) \) and \( g_i(x) \) are all multivariate polynomials in \( x \in \mathbb{R}^n \). Lasserre proposed the famous SOS relaxation for (2.1): choose an integer \( N > 0 \) and solve the convex program

\[
\begin{align*}
\max_{\gamma} & \quad \gamma \\
\text{s.t.} & \quad f(x) - \gamma \equiv \sigma_0 + \sigma_1 g_1(x) + \cdots + \sigma_m g_m(x) \\
& \quad \deg(\sigma_0), \deg(\sigma_1 g_1), \ldots, \deg(\sigma_m g_m) \leq 2N \\
& \quad \sigma_0, \sigma_1, \ldots, \sigma_m \text{ are SOS}
\end{align*}
\]  

(2.2)

where \( N \) is the relaxation order. The decisions variables in (2.2) are \( \gamma \) and the coefficients of polynomials \( \sigma_i(x) \). The relaxation (2.2) is equivalent to a particular SDP problem. Every \( f_N^* \) is a lower bound of the global minimum \( f^* \). To prove the convergence, a so-called archimedean condition is required, which assumes that there exists a positive scalar \( R > 0 \) and SOS polynomials \( s_0(x), s_1(x), \ldots, s_m(x) \) such that

\[
R - \|x\|_2^2 \equiv s_0(x) + g_1(x)s_1(x) + \cdots + g_m(x)s_m(x).
\]

The archimedean condition implies the feasible set of (2.1) is compact, while the reverse might not be true. However, the archimedean condition can be easily satisfied by adding a redundant constraint like \( R - \|x\|_2^2 \geq 0 \). Under the archimedean condition, Lasserre proved the convergence

\[
\lim_{N \to \infty} f_N^* = f^*.
\]

by applying Putinar’s Positivstellensatz. But it was unknown what is the speed of the convergence. A fundamental question is how good \( f_N^* \) approximates \( f^* \), that is, how fast does \( f_N^* \) approaches \( f^* \). This question was almost completely open since Lasserre proposed this hierarchy of SOS relaxations. A contribution of this thesis is to get specific approximation bounds for the quality of relaxation (2.2). For convenience, denote

\[
S(g_1, \ldots, g_m) = \{ x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_m(x) \geq 0 \}.
\]

Theorem 2.1. Suppose the archimedean condition holds, \( S(g_1, \ldots, g_m) \neq \emptyset \) and \( f \) is a polynomial in \( x \). Then there is

- a constant \( c > 0 \) depending only on \( g_1, \ldots, g_m \) and
- a constant \( c' > 0 \) depending on \( g_1, \ldots, g_m \) and \( f \)
such that for $f^*$ and $f^*_N$ as defined in Lasserre’s relaxations,

$$0 \leq f^* - f^*_N \leq \frac{c'}{\sqrt{\log N}}$$

for all $N \in \mathbb{N}$.

The constants $c, c'$ would be constructible from polynomials $g_1, \ldots, g_m$ and $f$.

The derivation of approximation bounds in Theorem 2.1 is closely related to the question of degree bounds in Putinar’s Positivstellensatz. In 1993, Putinar proved the following milestone:

(Putinar’s Positivstellensatz) Suppose the archimedean condition holds for polynomials $g_1, \ldots, g_m$. Then for every polynomial $f$ which is positive on $S(g_1, \ldots, g_m)$, there exist SOS polynomials $\sigma_0, \sigma_1, \ldots, \sigma_m$ such that

$$f(x) - \gamma = \sigma_0 + \sigma_1(x)g_1(x) + \cdots + \sigma_m(x)g_m(x).$$

This result is the basis for proving the convergence of Lasserre’s relaxations. Since Putinar published his result, very little was known about the degree bounds of $\sigma_0, \sigma_1, \ldots, \sigma_m$ in terms of $f(x)$ and $g_i(x)$. A contribution of this thesis is to prove an explicit degree bound for polynomials $\sigma_i$.

**Theorem 2.2.** Suppose the archimedean condition holds and $S(g_1, \ldots, g_m) \neq \emptyset$. Then there is a constant $c > 0$ (depending on $g_1, \ldots, g_m$) such that for all polynomials $f$ of degree $d$ with

$$f^* := \min \{ f(x) \mid x \in S(g_1, \ldots, g_m) \} > 0,$$

the SOS polynomials $\sigma_0, \sigma_1, \ldots, \sigma_m$ in Putinar’s Positivstellensatz can be chosen with degrees such that

$$\deg(\sigma_0), \deg(\sigma_1g_1), \ldots, \deg(\sigma_m g_m) \leq c \exp \left( \left( d^2 n^d \| f \| / f^* \right) c \right).$$

In the degree bound of Theorem 2.2, the norm $\| f \|$ is defined as follows: if a polynomial $f(x)$ is presented in the form

$$f(x) = \sum_{(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n, \alpha_1 + \cdots + \alpha_n \leq d} f_{\alpha} x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n},$$

then its norm is defined as

$$\| f \| := \max_{(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n, \alpha_1 + \cdots + \alpha_n \leq d} |f_{\alpha}| \frac{\alpha_1! \cdots \alpha_n!}{|\alpha|!}.$$

In Theorem 2.2, the bound depends on three parameters:

- The polynomials $g_1, \ldots, g_m$, which describes the algebraic geometric property of the feasible set $S(g_1, \ldots, g_m)$,
- the number $n$ of variables and the degree $d$ of polynomial $f$, which measures the “size” of polynomial $f(x)$, and
- the ratio $\| f \| / f^*$, which measures how close $f$ comes to have a zero on $S(g_1, \ldots, g_m)$. 

3
3 SOS Method Based on the Gradient Ideal

A general polynomial optimization is

\[ f^* := \min_{x \in \mathbb{R}^n} f(x) \]  \hspace{1cm} (3.3)

where \( f(x) \) is a real multivariate polynomial in \( x \in \mathbb{R}^n \). As is well-known, problem (3.3) is also NP-hard, even when the degree \( d \) of \( f(x) \) is fixed to be four. The standard SOS relaxation is to find a lower bound for the global minimum \( f^* \) by solving the convex program:

\[ f^*_\text{sos} := \max \gamma \text{ s.t. } f(x) - \gamma \text{ is SOS.} \]  \hspace{1cm} (3.4)

Obviously \( f^*_\text{sos} \leq f^* \). It can be shown that \( f^*_\text{sos} = f^* \) if and only if \( f(x) - f^* \) is SOS. An attractive property of (3.4) is that it is equivalent to an SDP problem and hence can be solved efficiently by numerical methods.

Recently Blekherman showed that, for any fixed even degree \( d \geq 4 \) and large \( n \), there are significantly more nonnegative polynomials than SOS polynomials. So we usually do not expect \( f^*_\text{sos} = f^* \). For the case that \( f^*_\text{sos} < f^* \), how can we find better lower bounds or even the global minimum \( f^* \)? A contribution of this thesis is to construct the so-called gradient SOS method which gives much better approximations. They are based on the observation that the gradient of \( f(x) \) vanishes on the optima. This leads to the gradient SOS method: choose an integer \( N \) and solve the convex program

\[
\begin{align*}
    f^*_{N, \text{grad}} := \max \gamma & \text{ s.t. } \quad f(x) - \gamma \equiv \sigma(x) + \sum_{j=1}^n \phi_j(x) \frac{x_j}{x_j} \\
    & \text{deg}(\sigma), \text{deg} \left( \phi_1 \frac{x_1}{x_1} \right), \ldots, \text{deg} \left( \phi_n \frac{x_n}{x_n} \right) \leq 2N \\
    & \sigma(x) \text{ is SOS.}
\end{align*}
\]  \hspace{1cm} (3.5)

The decision variables in (3.5) are the coefficients of polynomials \( \sigma(x) \) and \( \phi_i(x) \) in stead of \( x \), and (3.5) is equivalent to an SDP problem.

Note that (3.5) essentially requires \( f(x) - \gamma \) to be SOS modulo the gradient ideal:

\[ \mathcal{I}_{\text{grad}}(f) = \left\langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right\rangle. \]  \hspace{1cm} (3.6)

the ideal generated by the partials of \( f(x) \). The ideal \( \mathcal{I}_{\text{grad}}(f) \) is said to be radical if every polynomial \( h(x) \) belongs to \( \mathcal{I}_{\text{grad}}(f) \) if \( \nabla f(x) = 0 \) implies \( h(x) = 0 \). In case \( n = 1 \), radicalness means \( \nabla f(x) = 0 \) has no repeated roots. So \( \mathcal{I}_{\text{grad}}(f) \) being radical is a quite general condition.

When \( N \to \infty \), (3.5) generates a sequence of lower bounds \( f^*_{N, \text{grad}} \). Its convergence is summarized as follows:

**Theorem 3.1.** Let \( f(x) \) be a polynomial in \( x \in \mathbb{R}^n \) such that its minimum \( f^* \) is attainable at some point. Then \( \lim_{N \to \infty} f^*_{N, \text{grad}} = f^* \). Furthermore, if \( \mathcal{I}_{\text{grad}}(f) \) is radical, then there exists an integer \( N \) such that \( f^*_{N, \text{grad}} = f^* \).  

Define the real gradient variety: \( V^R_{\text{grad}}(f) = \{ u \in \mathbb{R}^n : \nabla f(u) = 0 \} \). The proof of Theorem 3.1 is based on the following result:

**Theorem 3.2.** Let \( f(x) \) be a polynomial. If \( f(x) \) is strictly positive on \( V^R_{\text{grad}}(f) \), then \( f(x) \) is SOS modulo its gradient ideal \( \mathcal{I}_{\text{grad}}(f) \). Furthermore, if \( \mathcal{I}_{\text{grad}}(f) \) is radical and \( f(x) \) is nonnegative over \( V^R_{\text{grad}}(f) \), then \( f(x) \) is SOS modulo \( \mathcal{I}_{\text{grad}}(f) \).
In Theorem 3.2, if \( f(x) \) is nonnegative over \( V_{\text{grad}}^R(f) \) but \( I_{\text{grad}}(f) \) is not radical, then \( f(x) \) might not be SOS modulo \( I_{\text{grad}}(f) \). This is shown by the following counterexample.

**Example 3.3.** Let \( n = 3 \) and \( f(x) = x_1^8 + x_2^8 + x_3^8 + x_1^4x_2^4 + x_1^2x_2^2 + x_3^6 - 3x_1^2x_2^2x_3^2 \). It can be shown that \( f(x) \) is nonnegative on \( V_{\text{grad}}^R(f) \) but not SOS modulo \( I_{\text{grad}}(f) \). However, for arbitrary \( \epsilon > 0 \), it holds that

\[
\begin{align*}
    f(x) + \epsilon \equiv p_1^2 \cdot (p_2^2 + \epsilon) + p_0^2 \cdot \epsilon \cdot (1 + \frac{1}{2\epsilon})^2 & \mod I_{\text{grad}}(f).
\end{align*}
\]

Here \( p_0, p_1, p_2 \) are polynomials having real coefficients independent of \( \epsilon \).

The lack of radicalness of the gradient ideal \( I_{\text{grad}}(f) \) would prevent the gradient SOS method to terminate in finitely many steps. Fortunately, this does not happen often in practice because the ideal \( I_{\text{grad}}(f) \) is generically radical, as shown by the following proposition.

**Proposition 3.4.** For almost every polynomial \( f(x) \), its gradient ideal \( I_{\text{grad}}(f) \) is radical and the gradient variety \( V_{\text{grad}}(f) \) is a finite subset of \( \mathbb{C}^n \).

Usually gradient SOS method gives much better solutions than the standard SOS method. This is illustrated by the following examples.

**Example 3.5.** (i) For \( f(x) = x_1^4x_2^2 + x_1^2x_2^2 + 1 - 3x_1^2x_2^2 \), it holds that \( f^* = 0 > f_{\text{sos}} = -\infty \). However, if we apply (3.5) to minimize \( f(x) \) with \( N = 4 \), we can get \( f_{4,\text{grad}}^* = -6.1463 \cdot 10^{-10} \). The dual optimal solution to (3.5) can be used to obtain the global minimizers \((\pm 1, \pm 1)\).

(ii) For \( f(x) = x_1^4 + x_1^2 + x_2^6 - 3x_1^2x_2^2 \), it holds \( f^* = 0 > f_{\text{sos}} = -729/4096 \). However, if we apply (3.5) to minimize \( f(x) \) with \( N = 4 \), we get \( f_{4,\text{grad}}^* = -9.5415 \cdot 10^{-12} \). The dual optimal solution to (3.5) can be used to get the global minimizers \((0, 0), (\pm 1, \pm 1)\).

People might think the gradient SOS method (3.5) would be much more expensive to solve than the standard SOS method (3.4), since (3.5) involves more variables. However, this is not the typical case in numerical implementation. The performance of the gradient SOS method (3.5) was tested on a class of randomly generated polynomials of the form:

\[
f(x) = x_1^d + \cdots + x_n^d + p(x),
\]

where \( d \) is an even integer and \( p(x) \) is a multivariate polynomial in \( x \in \mathbb{R}^n \) of degree \( \leq d - 1 \) whose coefficients are randomly generated in Gaussian distribution. The polynomials in this family are guaranteed to have a finite global minimum which is attainable at some point, and were proposed by Parrilo and Sturmfels in their original paper as testing the performance of the standard SOS relaxation. For relaxation order \( N = d/2 \) (\( d \) is the even degree of polynomial \( f(x) \)), the gradient SOS relaxation (3.5) almost always gives the exact global minimum value (up to machine precision). However, more interestingly, the gradient SOS method (3.5) takes less CPU time than the standard SOS relaxation (3.4) does, taking about \( 3/4 \) of the CPU time consumed by (3.4). This experimentation shows that the gradient SOS method (3.5) is also computationally competitive when compared with the standard SOS method (3.4). This is very surprising since (3.5) uses more variables and the resulting SDP has bigger size. One possible reason for this observation is that adding gradients improves the conditioning of the resulting SDPs and makes the interior-point algorithms solving the SDPs converge faster.
4 SOS Method based on KKT ideal

The general polynomial optimization has the form

\[
\begin{align*}
\text{min} & \quad f(x) \\
\text{s.t.} & \quad g_i(x) = 0, \quad i = 1, \ldots, s, \\
& \quad h_j(x) \geq 0, \quad j = 1, \ldots, t.
\end{align*}
\]

(4.1)

where \( f(x), g_i(x), h_j(x) \) are all multivariate polynomials in \( x \in \mathbb{R}^n \). Lasserre’s relaxation would be applied to solve (4.1) approximately. When the feasible set of (4.1) is compact and the archimedean condition holds, Lasserre’s relaxations converge. Generally we do not expect that Lasserre’s relaxations converge in finitely many steps, and the speed might be very slow as shown in Section 2. When the feasible set of (4.1) is noncompact, Lasserre’s relaxations usually do not converge.

Similar to the gradient SOS method, better numerical methods can be obtained by using optimality conditions, i.e., the Kuhn-Karush-Tucker (KKT) system:

\[
F := \nabla f(x) + \sum_{i=1}^{s} \lambda_i \nabla g_i(x) - \sum_{j=1}^{t} \nu_j \nabla h_j(x) = 0,
\]

\[
\nu_j h_j(x) = 0, \quad j = 1, \ldots, t,
\]

\[
g_i(x) = 0, \quad i = 1, \ldots, s,
\]

\[
h_j(x) \geq 0, \nu_j \geq 0, \quad j = 1, \ldots, t.
\]

(4.2)

The basic idea is as follows: find the maximum lower bound \( \gamma \) such that \( f(x) - \gamma \) has certain SOS representations modulo the KKT ideal. To describe the method, define the truncated KKT ideal and truncated preorder and linear cones

\[
I_{N,KKT} = \left\{ \sum_{k=1}^{n} \phi_k F_k + \sum_{i=1}^{s} \varphi_i g_i + \sum_{j=1}^{t} \psi_j \nu_j h_j \mid \deg(\phi_k F_k), \deg(\varphi_i g_i), \deg(\psi_j \nu_j h_j) \leq 2N \right\}
\]

\[
P_{N,KKT} = \left\{ \sum_{i=0}^{t} \sigma \theta_i h_i \mid \deg(\sigma \theta_i h_i) \leq 2N \right\}
\]

\[
M_{N,KKT} = \left\{ \sigma_0 \sum_{j=1}^{t} \sigma_j h_j \mid \deg(\sigma_0), \deg(\sigma_j) \leq 2N \right\} + I_{N,KKT}.
\]

Based on \( P_{N,KKT} \) and \( M_{N,KKT} \), two sequences \( \{p_N^*\}, \{f_N^*\} \) of lower bounds for the global minimum \( f^* \) can be obtained as follows:

\[
p_N^* \quad (\text{resp. } f_N^*) := \max_{\gamma} \quad \gamma \\
\quad \text{s.t. } \quad f(x) - \gamma \in P_{N,KKT} \quad (\text{resp. } M_{N,KKT}).
\]

(4.3)

The (4.3) is called the KKT SOS relaxation, or the KKT SOS method. Note that (4.3) is also equivalent to some SDP problems and hence can be solved efficiently. The sequences \( \{f_N^*\}_{N=1}^{\infty} \) and \( \{p_N^*\}_{N=1}^{\infty} \) are monotonically increasing and satisfy \( f_N^* \leq p_N^* \leq f^* \). Their convergence is as follows:

**Theorem 4.1.** Assume \( f^* = f(x^*) \) at some KKT point \( x^* \). Then \( \lim_{N \to \infty} p_N^* = f^* \). Furthermore, if in addition the KKT ideal is radical, then there exists \( N \in \mathbb{N} \) such that \( p_N^* = f^* \).

The convergence of the lower bounds \( \{f_N^*\} \) to \( f^* \) are not guaranteed. A counterexample is

\[
\min \quad (x_3 - x_1^2 x_2)^2 - 1
\]

\[
\text{s.t. } \quad 1 - x_1^2 \geq 0, \quad x_2 \geq 0, \quad x_3 - x_2 - 1 \geq 0.
\]
However, if the KKT ideal satisfies the archimedean condition, the we also have \( \lim_{N \to \infty} f_N^* = f^* \).

Usually the KKT SOS relaxation (4.3) gives much better approximations than Lasserre’s relaxations. This is illustrated by the following example:

**Example 4.2.** Consider the nonconvex quadratic optimization

\[
\begin{align*}
\min \quad & x_1^2 + x_2^2 \\
\text{s.t.} \quad & x_2^2 - 1 \geq 0, \quad x_1^2 \pm Mx_1x_2 - 1 \geq 0
\end{align*}
\]

where \( M \) is a positive constant. Simple calculation shows that the global minimum is

\[
f^* = \frac{1}{2} (M^2 + M\sqrt{M^2 + 4}) + 2.
\]

If we apply the standard Lasserre’s relaxation, it can be shown that the maximum possibly obtainable lower bound is 2. When \( M \to \infty \), the quality of Lasserre’s relaxation would be arbitrarily bad. However, if we apply the KKT gradient method, we can get \( f_N^* = p_N^* = f^* \), i.e., it converges in five steps. The KKT system plays a crucial role in this example.

## 5 Minimizing Rational Functions

An important problem is the multivariate rational polynomial optimization of the form

\[
\begin{align*}
r^* := \min \quad & \frac{f(x)}{g(x)} \\
\text{s.t.} \quad & h_1(x) \geq 0, \ldots, h_m(x) \geq 0
\end{align*}
\]  

where \( f(x), g(x), h_i(x) \) are all real multivariate polynomials in \( x \in \mathbb{R}^n \). Denote by \( r^* \) the global minimum objective value. Without loss of generality, assume that \( g(x) \) is nonnegative in the feasible set, since otherwise we can replace \( f(x)g(x) \) by \( f(x)g^2(x) \). SOS methods can be generalized to solve (5.1), but there are some special features different from the polynomial case. The difficulty appears when the minimum occurs on the common zeros of \( f(x) \) and \( g(x) \).

The following SOS relaxation solves (5.1) approximately: choose a positive integer \( N \) and solve the SOS program

\[
\begin{align*}
r_N^* := \max \quad & \gamma \\
\text{s.t.} \quad & f(x) - \gamma g(x) \equiv \sigma_0(x) + \sum_{i=1}^m \sigma_i(x)h_i(x) \\
& \deg(\sigma_0), \deg(\sigma_1g_1), \ldots, \deg(\sigma_mg_m) \leq 2N \\
& \sigma_0, \sigma_1, \ldots, \sigma_m \text{ are SOS.}
\end{align*}
\]

(5.2)

For every fixed \( N \), the SOS program (5.2) is equivalent to some SDP problem and hence can be solved efficiently. It can be shown that \( \{r_N^*\} \) is a monotonically increasing sequence of lower bounds for \( r^* \). To prove the convergence of \( r_N^* \) to \( r^* \), we need the archimedean condition: there exist a scalar \( R > 0 \) and SOS polynomials \( s_0(x), s_1(x), \ldots, s_m(x) \) such that

\[
R - \|x\|^2_2 \equiv s_0(x) + s_1(x)h_1(x) + \cdots + s_m(x)h_m(x).
\]

**Theorem 5.1.** Assume the archimedean condition holds and the minimum \( r^* \) is finite. If \( f(x) = g(x) = 0 \) has no solutions in the feasible set, then \( \lim_{N \to \infty} r_N^* = r^* \).
In Theorem 5.1, the assumption that \( f(x) \) and \( g(x) \) have no common zeros in the feasible set can not be removed. For a counterexample, consider problem \((n = 1)\)

\[
\min_{x \in \mathbb{R}} \quad \frac{1 + x}{(1 - x^2)^2} \quad \text{s.t.} \quad (1 - x^2)^3 \geq 0.
\]

The global minimum is \( r^* = \frac{27}{52} \) and the minimizer is \( x^* = -\frac{1}{3} \). However, for any \( N \), (5.2) does not give any finite lower bound for \( r^* \).

However, in case of two variables \((n = 2)\) and under some regularity assumptions, we can still prove the convergence of \( r_N^* \) to \( r^* \) even if \( f(x) \) and \( g(x) \) have common zeros in the feasible set.

**Theorem 5.2.** Let \( n = 2 \) and \( S \) be the feasible set of (5.1). Let \( Z(f, g) = \{u \in S : f(u) = g(u) = 0\} \) and \( \Theta \) 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 \( C_i = \{x \in \mathbb{C}^2 : h_i(x) = 0\} \) is reduced and no two of them share an irreducible component. No point in \( \Omega \) is a singular point of the curve \( C_1 \cup \cdots \cup C_m \).

(ii) Each point of \( \Omega \) 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 \( r_N^* = r^* \).

### 6 Applications of Polynomial Optimization

The following applications of polynomial optimization are discussed in this thesis:

**Shape optimization of transfer functions:** In linear control theory, an important problem is to design a system such that its transfer function

\[
H(s) = d + c^T (sI - A)^{-1} b
\]

has the desired shape when \( s \) is restricted to be purely imaginary numbers. Here \( A, b, c, d \) are matrices/vectors of proper dimensions. The Positivstellensatz can be applied successfully in this application by solving certain SDP problems.

**Perturbation analysis for polynomial systems:** The solution set of a multivariate polynomial system changes when its parameters are perturbed in some neighborhood. One practical technique is to find the smallest ellipsoid to contain all the perturbed solutions. The can be achieved by solving some SDP problems using Positivstellensatz. It gives much better perturbation analysis than traditional approaches.

**Sensor network localization:** The problem is as follows: for a sequence of unknown vectors (called sensors) \( x_1, x_2, \cdots, x_n \) in \( \mathbb{R}^k \) \((k = 1, 2, \cdots)\), 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 \( a_1, \cdots, a_m \) (called anchors) are equal to some given numbers. The problem is equivalent to solve the a quartic polynomial optimization problem of the form

\[
\min_{x_1, \cdots, x_n \in \mathbb{R}^{d \times n}} \quad \sum_{(i,j) \in A} (\|x_i - x_j\|_2^2 - d_{ij}^2)^2 + \sum_{(i,k) \in B} (\|x_i - a_k\|_2^2 - e_{ik}^2)^2.
\]

The above polynomial is very special and sparse. The sparse SOS relaxations can be applied successfully to find sensor locations in large scale problems.