Lecture 18

Lecturer: Pablo A. Parrilo *Scribe: ???*

In this lecture we introduce Schmüdgen's theorem about the K-moment problem (or equivalently, on the representation of positive polynomials) and describe the basic elements in his proof. This approach combines both algebraic tools (using the Positivstellensatz to prove the boundedness of certain operators) and functional analysis (spectral measures of commuting families of operators and the Hahn-Banach theorem). We will also describe some alternative versions due to Putinar, as well as a related purely functional-analytic result due to Megretski.

For a comprehensive treatment and additional references, we mention [BCR98, Mar00, PD01] among others.

1 Representations of positive polynomials

As we have seen, the Positivstellensatz allows us to obtain certificates of the emptiness of a basic semialgebraic set, explicitly given by polynomials. When looking for bounded degree certificates, this provides a natural hierarchy of SDP-based conditions [Par00, Par03].

What if we want to apply this for the particular case of optimization? As we have seen, it is relatively straightforward to convert a polynomial optimization problem to a one-parameter family of feasibility problems, by considering the sublevel sets, i.e., the sets $\{x \in \mathbb{R}^n \mid f(x) \leq \gamma\}$.

In the general case of constrained problems, however, using the full power of the Psatz will yield conditions that are not linear in the unknown parameter γ (because we need products between the constraints and objective function), and in principle, this presents a difficulty to the direct use of SDP. Notice nevertheless, that the problem is certainly an SDP for any fixed value of γ, and is thus quasiconvex (which is almost as good, except for the fact that we cannot use "standard" SDP solvers to solve it directly, but rather rely on methods such as bisection).

Of course, we can always produce specific families of certificates that are linear in γ, and use them for optimization (e.g., like we did in the copositivity case). However, in general it is unclear whether the desired family is "complete," in the sense that we will be able to prove arbitrarily good bounds on the optimal value as the degree of the polynomials grows to infinity.

1.1 Schmüdgen's theorem

In 1991, Schmüdgen presented a characterization of the moment sequences of measures supported on a compact semialgebraic K (the K-moment problem). As in the one-dimensional case we studied earlier the question is, given an (infinite) sequence of moments, decide whether it actually corresponds to a nonnegative measure with support on a given set K.

His solution combined both real algebraic methods (the Psatz), with some functional analytic tools (reproducing kernel Hilbert spaces, bounded operators, and the spectral theorem).

This characterization of moment sequences can be used, in turn, to produce an explicit description of the set of strictly positive polynomials on a compact semialgebraic set:

Theorem 1 ([Sch91]). *If $p(x)$ is strictly positive on $K = \{x \in \mathbb{R}^n \mid f_i(x) \geq 0\}$, and K is compact, then $p(x) \in \mathbf{cone}\{f_1, \ldots, f_m\}$.*

`expand` ToDo

There are several interesting ideas in the proof; a coarse description follows. The first step is to use the Positivstellensatz to produce an algebraic certificate of the compactness of the set K. Then the given moment sequence (which is a positive definite function on the semigroup of monomials) is used to construct a particular pre-Hilbert space and its completion (namely, the associated reproducing kernel Hilbert space). In this Hilbert space, we consider linear operators T_{x_i} given by multiplication by the coordinate variables, and use the algebraic certificate of compactness to prove that these are bounded. Now, the T_{x_i} are a finite collection of pairwise commuting, bounded, self-adjoint operators, and thus there exists a spectral measure for the family, from which a measure, only supported in K, can be extracted. Finally, a Hahn-Banach (separating hyperplane) argument is used to prove the final result.

1.2 Putinar's approach

The theorem in the previous section requires (in principle) all $2^m - 1$ squarefree products of constraints[1]. Putinar [Put93] presented a modified formulation (under stronger assumptions) for which the representation is *linear* in the constraints. We introduce the following concept:

Definition 2. *Let* $\{f_1, \ldots, f_m\} \subset \mathbb{R}[x]$. *The* preprime *generated by the* f_i*, and denoted by* $\mathbf{preprime}\{f_1, \ldots, f_m\}$ *is the set of all polynomials of the form* $s_0 + s_1 f_1 + \cdots + s_m f_m$*, where all the* s_i *are sums of squares.*

Notice that $\mathbf{preprime}\{f_i\} \subset \mathbf{cone}\{f_i\}$, and that every element of either set takes only nonnegative values on $\{x \in \mathbb{R}^n, f_i(x) \geq 0\}$.

Theorem 3 ([Put93]). *Consider a set* $K = \{x \in \mathbb{R}^n \,|\, f_i(x) \geq 0\}$*, such that there exists a* $q \in \mathbf{preprime}\{f_1, \ldots, f_m\}$ *with* $\{x \in \mathbb{R}^n, q(x) \geq 0\}$ *compact (this implies that* K *itself is compact). Then,* $p(x) > 0$ *on* K *if and only if* $p(x) \in \mathbf{preprime}\{f_1, \ldots, f_m\}$.

Notice that here, the polynomial q serves as an algebraic certificate of the compactness of K, so in this case the Psatz is not needed.

Putinar's theorem was used by Lasserre to define a hierarchy of semidefinite relaxations for polynomial optimization, based on the dual moment interpretation [Las01].

1.3 An elementary argument for the n-sphere

`Complete` ToDo

1.4 Tradeoffs

In principle (and often, in practice) there is a tradeoff between how "expressive" our family of certificates is, the quality of the resulting bounds, and the complexity of finding proofs.

On one extreme, the most general method is the Psatz, as it encapsulates pretty much every possible "algebraic deduction," and will certainly provide the strongest bounds, since it includes the other techniques as special cases. For optimization, Schmüdgen's theorem provides the advantages

[1] Recall that in practice, this may not be a issue at all, since the restriction on the degree of the certificates imposes a strict limit on how many products can be included.

of a linear representation, although (possibly) at the cost of having a large number of products between the constraints. Finally, the Putinar approach has a reduced number of constraints (and thus, SOS multipliers), although the obtained bounds can potentially be much weaker than the previous ones.

In the end, the decision concerning what approach to use should be dictated by the available computational resources, i.e., the size of the SDPs that we can solve in a reasonable time. It is not difficult to produce examples with significant gaps between the corresponding bounds; see for instance [Ste96] for a particularly simple example, that is trivial for the Psatz, but for which either the Schmüdgen or Putinar representations need large degree refutations.

Example 4. *In [Ste96], Stengle presented an interesting example to assess the computational requirements of Schmüdgen's theorem. His concrete example was to find a representation certifying the nonnegativity of $f(x) := 1 - x^2$ over $g(x) := (1 - x^2)^3 \geq 0$.*

The Positivstellansatz gives a very simple certificate of this property, or equivalently, the emptiness of the set $\{g(x) \geq 0, -f(x) \geq 0, zf(x) - 1 = 0\}$ (where we have used, as before, Rabinowitch's trick). Indeed, we have the identity:

$$z^4 \cdot (-f) \cdot g + (zf - 1) \cdot (z^3 f^3 + z^2 f^2 + zf + 1) = -1.$$

Using a simple argument, Stengle proved in [Ste96], that no representation of the form (1) exists when $\gamma = 0$.

$$(1 - x^2) + \gamma = Q(x) + P(x)(1 - x^2)^3, \tag{1}$$

where $Q(x), P(x)$ are sums of squares. To see this, evaluate this expression at $x = \pm 1$.

Furthermore, he has shown that $\gamma \to 0$, the degrees of P, Q satisfying the identity necessarily have to go to infinity, and provided the bounds $O(\gamma^{-\frac{1}{2}}) \leq deg(P) \leq O(\gamma^{-\frac{1}{2}} \log \frac{1}{\gamma})$.

As an interesting aside, it can be shown that the optimal solution of this problem can be exactly computed:

Theorem 5. *Let the degree of $P(x)$ be equal to $4N$. Then, the optimal solution that minimizes γ in (1) has:*

$$\gamma_N^* = \frac{1}{(2N+2)^2 - 1}, \quad P(x) = p(x)^2, \quad Q(x) = q(x)^2$$

where

$$\begin{aligned} p(x) &= 2(N+1)\ {}_2F_1(-N, N+2\,;\ \tfrac{1}{2}; x^2) \\ q(x) &= \frac{1}{\gamma_N^*}\, x\ {}_2F_1(-N-1, N+1\,;\ \tfrac{3}{2}; x^2) \end{aligned}$$

and ${}_2F_1(a, b; c, x)$ is the standard Gauss hypergeometric function [AS64, Chapter 15].

1.5 Trigonometric case

Megretski [Meg03] analyzed the case of trigonometric polynomials, and gave a simple argument for the existence of certain distinguished representations. We introduce the following notation: let $\mathbb{T}_n = \{z \in \mathbb{C}^n, |z_i| = 1\}$ be the n-dimensional torus, P_n is the set of multivariate Laurent polynomials, and $RP_n \subset P_n$ are the Laurent polynomials that are real-valued on $\mathbb{T}_n$.

Theorem 6 ([Meg03]). *Let $\{F, Q_1, \ldots, Q_m\} \subset RP_n$, such that $F(z) > 0$ for all $z \in \mathbb{T}_n$ satisfying $Q_1(z) = \cdots = Q_m(z) = 0$. Then there exist $V_1, \ldots, V_r \in P_n$, $H_1, \ldots, H_m \in RP_n$, such that*

$$F(z) = \sum_{i=1}^{r} |V_i(z)|^2 + \sum_{j=1}^{m} H_j(z) Q_i(z).$$

By splitting into real and imaginary part, this corresponds to a special kind of (standard) polynomials, and a compact semialgebraic set (so in principle, any of the previous theorems would apply). However, this result exploits the complex structure for a more concise representation.

This theorem deals only with the equality case (no inequalities), and the feasible set is compact (since so it $\mathbb{T}^n$). It essentially states that a positive polynomial is a sum of squares modulo the ideal generated by the Q_i. Recall we have proved similar results in the zero-dimensional case, and this theorem naturally generalizes these.

Megretski's proof is purely functional-analytic, the main tools being Bochner's theorem and Hahn-Banach. If $(G, +)$ is an Abelian group, a function $\phi : G \to \mathbb{C}$ is *positive definite* if the matrix $[\phi(g_j - g_k)]_{j,k=1}^n$ is Hermitian positive semidefinite for all finite subsets $\{g_1, \ldots, g_n\} \subseteq G$. Bochner's theorem, an important result in harmonic analysis, characterizes positive definite functions in terms of the nonnegativity of their Fourier transform:

Theorem 7 (Bochner). *A function $\phi : G \to \mathbb{C}$ is positive definite if and only if there exists a nonnegative measure μ on the dual group $\hat{G}$ such that*

$$\phi(g) = \int_{\hat{G}} \rho(g) d\mu(\rho).$$

Recall that the elements of the dual group $\hat{G}$ are the *characters* (i.e., the group homomorphisms $\rho : G \to \mathbb{T}$, satisfying $\rho(g_1 + g_2) = \rho(g_1)\rho(g_2)$). Thus, Bochner's theorem says that every positive definite function on an Abelian group is a convex combination of characters. In our case, the group $G = \mathbb{Z}^n$ can be identified with the set of monomials, and the corresponding dual group is $\hat{G} = \mathbb{T}_n$.

In simplified terms, one reason why trigonometric (or Laurent) polynomials are somewhat "easier" than the general case is because in this case monomials form a *group*, as opposed to the *semigroup* structure of regular monomials. For the group case, the corresponding theory is the classical harmonic analysis on abelian groups (e.g., [Rud90]); while for semigroups there is the newer, but well-developed characterizations of positive definite functions on (Abelian) semigroups; see for instance [BCR84].

We also mention that there are "purely algebraic" versions of these theorems, that do not use functional analytic ideas (e.g., [Mar00]). Roughly, the role played by the compactness of K in proving the boundedness of the operators T_{x_i} is replaced with a property called *Archimedeanity* of the corresponding preorder.

References

[AS64] M. Abramowitz and I.A. Stegun, editors. *Handbook of Mathematical Functions.* Dover, 1964.

[BCR84] C. Berg, J. P. R. Christensen, and P. Ressel. *Harmonic analysis on semigroups*, volume 100 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1984.

[BCR98] J. Bochnak, M. Coste, and M-F. Roy. *Real Algebraic Geometry.* Springer, 1998.

[Las01] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817, 2001.

[Mar00] M. Marshall. *Positive polynomials and sums of squares.* Dottorato de Ricerca in Matematica. Dept. di Mat., Univ. Pisa, 2000.

[Meg03] A. Megretski. Positivity of trigonometric polynomials. In *Proceedings of the 42^{th} IEEE Conference on Decision and Control*, pages 3814–3817, 2003.

[Par00] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* PhD thesis, California Institute of Technology, May 2000. Available at `http://resolver.caltech.edu/CaltechETD:etd-05062004-055516`.

[Par03] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Prog.*, 96(2, Ser. B):293–320, 2003.

[PD01] A. Prestel and C. N. Delzell. *Positive polynomials: from Hilbert's 17th problem to real algebra.* Springer Monographs in Mathematics. Springer, 2001.

[Put93] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.*, 42(3):969–984, 1993.

[Rud90] W. Rudin. *Fourier analysis on groups.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1990.

[Sch91] K. Schmüdgen. The K-moment problem for compact semialgebraic sets. *Math. Ann.*, 289:203–206, 1991.

[Ste96] G. Stengle. Complexity estimates for the Schmüdgen Positivstellensatz. *J. Complexity*, 12(2):167–174, 1996.