Finding Representations for Nonnnegative Polynomials on Semialgebraic Sets

Ruchira Datta

Department of Mathematics
University of California
Berkeley, California

April 2nd, 2002

Introducing Polynomial Optimization

- Define a *polynomial objective* f
- We would like to know a lower bound for f
- For example, if $f + t \geq 0$, then $-t$ is such a bound
  - So minimizing is closely related to certifying nonnegativity
- If we have a method of certifying $f + t \geq 0$, then want to find the *least* t that we can with our method
- We could differentiate, but this is impractical

Certifying Nonnegativity

- A homogeneous form in n variables is called *positive semidefinite* (or *psd*) if it takes nonnegative values for all $(x_1, \ldots, x_n) \in \mathbf{R}^n$
- If a polynomial is a *sum of squares* (or *sos*) of other polynomials, it is nonnegative
- So far we haven't constrained the region over which to bound f
- We may restrict it to a *semialgebraic set*, defined by polynomial inequalities:
- $S = \{(x_1, \ldots, x_n) | p_1(x_1, \ldots, x_n) \geq 0, \ldots, p_m(x_1, \ldots, x_n) \geq 0\}$
- $p_1, \ldots, p_m$ are the *defining polynomials* of S
- If a polynomial is a sum of products of the p_i's with nonnegative coefficients, it is nonnegative on S

History: Hilbert

- In 1888, Hilbert showed that every psd ternary quartic form is a sum of three squares of quadratic forms
- At the same time, he proved that some psd forms are *not* the sum of squares (or *sos*) of any other forms
- In his 1900 address to the Paris Congress, he posed as his 17th Problem the following:
  - Suppose f is nonnegative. Does there exist a representation

$$f = \sum_i \frac{p_i^2}{q_i^2}$$

of f as the sum of squares of *rational functions* with real coefficients?

Representations Exist!

- In 1926, Artin answered affirmatively
- In 1964, Krivine showed that if f is *strictly positive* on the semialgebraic set $S = \{\bar{x} : p_1(\bar{x}) \geq 0, \ldots, p_m(\bar{x}) \geq 0\}$, then f can be represented in the form

$$\frac{1 + \sum_{\epsilon \in \{0,1\}^m} (\sum_i a_{\epsilon i} f_{\epsilon i}^2) p_1^{\epsilon_1} \cdots p_m^{\epsilon_m}}{1 + \sum_{\epsilon' \in \{0,1\}^m} (\sum_j b_{\epsilon' i} g_{\epsilon' j}^2) p_1^{\epsilon'_j} \cdots p_m^{\epsilon'_m}}$$

where the $0 \leq a_{\epsilon i}, b_{\epsilon' j} \in \mathbf{R}$ and $f_{\epsilon i}, g_{\epsilon' j} \in \mathbf{R}[\bar{x}]$

Compact Sets Lead To Nicer Representations

- In 1990, Schmüdgen showed that in particular, if S is compact, then f can be represented in the form

$$\sum_{\epsilon \in \{0,1\}^m} \left(\sum_i a_{\epsilon i} f_{\epsilon i}^2\right) p_1^{\epsilon_1} \cdots p_m^{\epsilon_m}$$

- In 1988, Handelman showed that in particular, if S is a polyhedron, then f can be represented in the form

$$\sum_{\epsilon \in \mathbf{N}^m} a_\epsilon p_1^{\epsilon_1} \cdots p_m^{\epsilon_m}$$

Another Nice Representation

- Return to the situation of Schmüdgen's theorem. In 1993, Putinar showed that if all the p_i's have even degree, and their highest degree homogeneous parts do not have common zeros in $\mathbf{R}^n$ other than 0, then f can be represented in the form

$$\sum_j f_{0j}^2 + \sum_{i=1}^m (\sum_j f_{ij}^2) p_i$$

- All these proofs are nonconstructive
- Note that strict positivity is required; so if we are trying to find a lower bound for f by representing $f + t$ in one of the above ways, there may not be a representation which gives us the exact minimum

Kadison-Dubois Representation Theorem: Intro

- In 1983, Becker and Schwarz gave a short algebraic proof of the Kadison-Dubois Representation Theorem (actually due to Krivine)
- *Definition* Let R be a ring containing $\mathbf{R}$. A subset $P \subset R$ is called a preprime iff

$$P + P \subseteq P, PP \subseteq P, \mathbf{R}_+ \subseteq P, -1 \notin P$$

and a preprime is archimedean iff

$$\forall a \in \mathbf{R} \quad \exists n \in \mathbf{N} \quad \text{s.t. } n - a \in P$$

Kadison-Dubois Representation Theorem

- *Theorem* Let R be a ring, $P \subseteq R$ an archimedean preprime and

$$X(P) := \{\phi \in Hom(R, \mathbf{R}) | \phi(P) \subseteq \mathbf{R}_+\}$$

the space of representations. If $f \in R$ satisfies $\phi(f) \geq 0$ for all $\phi \in X(P)$, then for any $n \in \mathbf{N}$, we have $1 + nf \in P$.

Algebraic Proofs of Representation Theorem

- Becker and Schwarz actually prove more than this, but this was all that was needed for:
- Wörmann to give a short algebraic proof of Schmüdgen's theorem in 1998
- In aid, of this, he gave a short algebraic proof of Pólya's theorem

Pólya's Theorem

- *Theorem (Pólya)* Let $f \in \mathbf{R}[\bar{x}]$ be homogeneous and strictly positive on the closed positive orthant minus the origin. Then for some $N \in \mathbf{N}$, the polynomial $(\sum_{i=1}^{n} x_i)^N f$ is a *positive* linear combination of monomials in the x_i's.

- There are bounds on the number N, called the *Pólya exponent*, but these vary inversely with the minimum value of f on S

Schweighofer's Constructive Proof I

- Schweighofer made Wörmann's proof constructive:
- Since S is compact, there is a number R such that $R - \sum_{i=1}^{n} x_i^2$ is strictly positive on S; so, it is sos
- Using the representation of $R - \sum_{i=1}^{n} x_i^2$, we may extend $p_1, \ldots, p_m$ to $q_1, \ldots, q_l$ defining the same set S, such that $\mathbf{R}[q_1, \ldots, q_l] = \mathbf{R}[x]$, and scaled such that $\sum_{i=1}^{l} q_i = 1$

Schweighofer's Constructive Proof II

- Define $\phi\colon \mathbf{R}[Y_1, \ldots, Y_l] \to \mathbf{R}[q_1, \ldots, q_l]$ by $\phi : Y_i \mapsto q_i$
- Then any element of $\phi^{-1}f$ is strictly positive on the variety in $\mathbf{R}^l$ corresponding to $\ker \phi$
- Add a nonnegative multiple s times a sum of squares of elements of $\ker \phi$ to obtain strict positivity on the whole positive orthant $Y_i \geq 0$
- Homogenize by multiplying as necessary by $\sum_{i=1}^{l} Y_i$, which ϕ maps to 1
- Now Pólya's theorem applies

Schweighofer's Algorithm Implemented

- Schweighofer's algorithm gives a Handelman representation on the q_i's, or a and a Schmüdgen representation on the original p_i's
- I have implemented this, solving an LP in s and t
  - s is a nonnegative multiplier of a sum of squares of elements of the Gröbner basis for $\ker \phi$
  - t is a constant added to f to convert certification of nonnegativity to minimization
  - The LP minimizes t

Schweighofer's Algorithm Critiqued

- I have found some problems with it in practice
- The outcome depends heavily on which element of $\phi^{-1} f$ was chosen to start with
- A Handelman-type representation is found; restricting the form of the representation means the size of the representation will be bigger
- In fact, it is even worse: requiring the representation to be homogeneous makes the representation invariably blow up

How To Find A Representation?

- Once we know a representation exists, we can search for the best one
- Sum of squares representations can be found by the Gram matrix method, which can be solved by SDP
- I have implemented this in the unconstrained case, and am working on the extension to Putinar and Schmüdgen representations
- I have implemented search for Handelman representations by LP

Exact vs. Floating-point Computations

- Some of the interest of these methods lies in their ability to give exact, trustworthy answers
- However, all current SDP implementations are based on interior-point methods and use floating-point arithmetic
- In 1994 Gábor Pataki proposed a simplex-type method for solving SDP's for theoretical purposes
- Alternatively, a floating-point implementation could be used as an "oracle", as in some exact LP solvers