SOS hierarchies we could have had in the 1920s…

Amir Ali Ahmadi
Princeton, ORFE
Affiliated member of PACM, COS, MAE, CSML

Georgina Hall
Princeton, ORFE-->INSEAD

PRINCETON UNIVERSITY ORFE

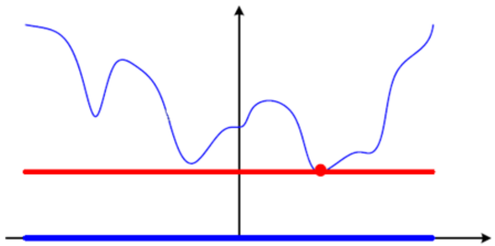
ISMP 2018, Bordeaux

How to prove positivity?

Is $p(x) > 0$ on $\{g_1(x) \geq 0, \ldots, g_m(x) \geq 0\}$?

Why prove positivity?

- (Tight) lower bounds for polynomial minimization problems



- Infeasibility certificates for systems of polynomial inequalities

$$\{g_1(x) \geq 0, g_2(x) \geq 0, \ldots, g_m(x) \geq 0\} \quad \text{empty}$$
$$\Leftrightarrow$$
$$-g_1(x) > 0 \text{ on } \{g_2(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

- Dynamics and control (Lyapunov functions)
- Stats/ML (Georgina's talk)

Positivstellensätze

$p(x) > 0, \forall x \in \mathbb{R}^n$

Artin

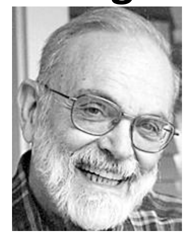
If $p(x) \geq 0, \forall x \in \mathbb{R}^n$,
then $\exists$ sos q s.t. $p \cdot q$ sos.

1927

20th century

1974

Stengle



1991

1993

$p(x) > 0$,
$\forall x \in S = \{x \mid g_i(x) \geq 0\}$

Schmudgen

If $p(x) > 0, \forall x \in S$,
then $p(x) = \sigma_0(x) + \sum_i \sigma_i(x) g_i(x) + \sum_{ij} \sigma_{ij}(x) g_i(x) g_j(x) + \cdots$, where $\sigma_0, \sigma_i, \ldots$ sos

(Requires some compactness assumptions)

Putinar

If $p(x) > 0, \forall x \in S$,
then $p(x) = \sigma_0(x) + \sum_i \sigma_i(x) g_i(x)$,
where σ_0, σ_i are sos

Search for these sos polynomials (when degree is fixed) --->SDP.

Motivation/Outline

"Can we get away with less?"

(in order to produce converging hierarchies of lower bounds for polynomial optimization problems)

Q1: Do we really need Stengle, Schmudgen, Putinar, …?
Can we only use certificates of global positivity? (e.g., Artin's)

Part I: *A meta-theorem*

Q2: Do we really need SDPs (or convex optimization)?

Part II: *An optimization-free Positivstellensatz*

Caveat: Have your theoretical hat on!

A meta-theorem for producing hierarchies

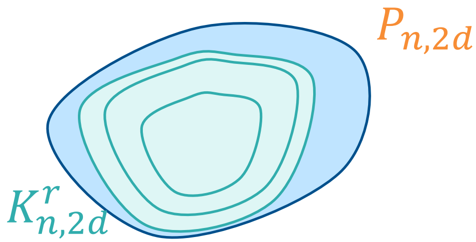
Theorem: Let $\{K^r_{n,2d}\}$ be a sequence of sets of homogeneous polynomials in n variables and of degree $2d$. If

(1) $K^r_{n,2d} \subseteq P_{n,2d}\ \forall r$ and $\exists\ s_{n,2d}$ pd in $K^0_{n,2d}$

(2) $p > 0 \Rightarrow \exists r \in \mathbb{N}$ s.t. $p \in K^r_{n,2d}$

(3) $K^r_{n,2d} \subseteq K^{r+1}_{n,2d}\ \forall r$

(4) $p \in K^r_{n,2d} \Rightarrow p + \epsilon s_{n,2d} \in K^r_{n,2d}, \forall \epsilon \in [0,1]$,



then,

POP	$\min_{x\in\mathbb{R}^n} p(x)$
	$s.t.\ g_i(x) \geq 0, i = 1, \ldots, m$

$2d$ = maximum degree of p, g_i

R: radius of the feasible set

$$\overset{r\uparrow\infty}{=} \text{ opt. val.}$$

$$\max_{\gamma} \gamma$$
$$s.t.\ f_\gamma(z) - \frac{1}{r} s_{n+m+3,4d}(z) \in K^r_{n+m+3,4d}$$

where f_γ is a form in $n + m + 3$ variables of degree $4d$ which can be written down explicitly from p, g_i, R.

What is f_γ? + Proof idea (1/2)

$$\min_{x\in\mathbb{R}^n} p(x)$$
$$s.t.\ g_i(x) \geq 0, i = 1, \ldots, m$$

$2d$ = maximum degree of p, g_i
R: radius of the feasible set

$$f_\gamma(x,s,y) := \left(\gamma y^{2d} - y^{2d}p(x/y) - s_0^2 y^{2d-2}\right)^2 + \sum_{i=1}^m \left(y^{2d} g_i(x/y) - s_i^2 y^{2d-2}\right)^2$$
$$+ \left((R + \sum_{i=1}^m \eta_i + \beta + \gamma)^d y^{2d} - (\sum_{i=1}^n x_i^2 + \sum_{i=0}^m s_i^2)^d - s_{m+1}^{2d}\right)^2$$

$degree\ 4d\ and\ in\ n+m+3\ variables\ (x_1,\ldots,x_n,s_0,\ldots,s_m,s_{m+1},y)$

$p(x) > \gamma$ on $\{x \mid g_i(x) \geq 0\} \Leftrightarrow f$ is positive definite (pd)

PRINCETON UNIVERSITY ORFE

Proof idea (2/2)

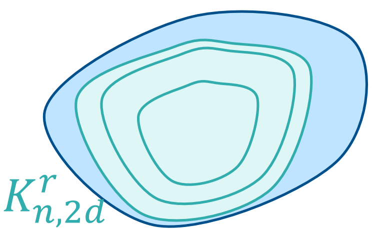
$p(x) > \gamma$ on $\{x \mid g_i(x) \geq 0\} \Leftrightarrow f$ is positive definite (pd)

(1) $K^r_{n,2d} \subseteq P_{n,2d}\ \forall r$ and $\exists\, s_{n,2d}$ pd in $K^0_{n,2d}$

(2) $p > 0 \Rightarrow \exists r \in \mathbb{N}$ s.t. $p \in K^r_{n,2d}$

(3) $K^r_{n,2d} \subseteq K^{r+1}_{n,2d}\ \forall r$

(4) $p \in K^r_{n,2d} \Rightarrow p + \epsilon s_{n,2d} \in K^r_{n,2d}, \forall \epsilon \in [0,1]$



POP	$\min_{x \in \mathbb{R}^n} p(x)$
	$s.t.\ g_i(x) \geq 0, i = 1, \ldots, m$

$$\overset{r \uparrow \infty}{=} \text{ opt. val.}$$

$$\max_{\gamma} \gamma$$
$$s.t.\ f_\gamma(z) - \frac{1}{r} s_{n+m+3,4d}(z) \in K^r_{n+m+3,4d}$$

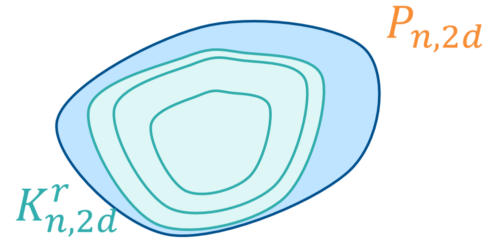
Families of cones that satisfy (1)-(4)

(1) $K_{n,2d}^r \subseteq P_{n,2d}$ $\forall r$ and $\exists$ $s_{n,2d}$ pd in $K_{n,2d}^0$

(2) $p > 0 \Rightarrow \exists r \in \mathbb{N}$ s.t. $p \in K_{n,2d}^r$

(3) $K_{n,2d}^r \subseteq K_{n,2d}^{r+1}$ $\forall r$

(4) $p \in K_{n,2d}^r \Rightarrow p + \epsilon s_{n,2d} \in K_{n,2d}^r, \forall \epsilon \in [0,1]$



Examples:

"Artin cones":

$$A_{n,2d}^r = \{p \mid p \cdot q \text{ is sos for some sos } q \text{ of degree } 2r\}$$

"Reznick cones":

$$R_{n,2d}^r = \{p \mid p \cdot (x_1^2 + \cdots + x_n^2)^r \text{ is sos}\}$$

(both lead to SDP-based hierarchies for polynomial optimization)

Enough to produce a converging hierarchy for POPs

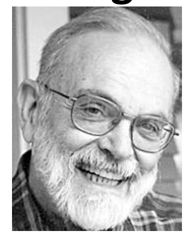
$p(x) > 0, \forall x \in \mathbb{R}^n$

Artin

If $p(x) \geq 0, \forall x \in \mathbb{R}^n$,
then $\exists$ sos q s.t. $p \cdot q$ sos.

1927

Stengle



1974

20th century

1991

1993

$p(x) > 0$,
$\forall x \in S = \{x \mid g_i(x) \geq 0\}$

Schmudgen

(Requires some compactness assumptions)

Putinar

If $p(x) > 0, \forall x \in S$,
then $p(x) = \sigma_0(x) + \sum_i \sigma_i(x) g_i(x) + \sum_{ij} \sigma_{ij}(x) g_i(x) g_j(x) + \cdots$, where $\sigma_0, \sigma_i, \ldots$ sos

If $p(x) > 0, \forall x \in S$,
then $p(x) = \sigma_0(x) + \sum_i \sigma_i(x) g_i(x)$,
where σ_0, σ_i are sos

Part II

An optimization-free Positivstellensatz (1/2)

$$p(x) > 0, \forall x \in \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \dots, m\}$$

$2d$ =maximum degree of p, g_i

$\Updownarrow$ Under compactness assumptions, i.e., $\{x \mid g_i(x) \geq 0\} \subseteq B(0, R)$

$\exists\, r \in \mathbb{N}$ such that

$$\left(f(v^2 - w^2) - \frac{1}{r}\left(\sum_i (v_i^2 - w_i^2)^2\right)^d + \frac{1}{2r}\left(\sum_i (v_i^4 + w_i^4)\right)^d\right) \cdot \left(\sum_i v_i^2 + \sum_i w_i^2\right)^{r^2}$$

has **nonnegative coefficients**,

where f is a form in $n + m + 3$ variables and of degree $4d$, which can be explicitly written from p, g_i and R.

PRINCETON UNIVERSITY ORFE

An optimization-free Positivstellensatz (2/2)

$$p(x) > 0 \text{ on } \{x \mid g_i(x) \geq 0\} \Leftrightarrow$$

$$\exists r \in \mathbb{N} \text{ s.t. } \left(f(v^2 - w^2) - \frac{1}{r}\left(\sum_i (v_i^2 - w_i^2)^2\right)^d + \frac{1}{2r}\left(\sum_i (v_i^4 + w_i^4)\right)^d \right) \cdot \left(\sum_i v_i^2 + \sum_i w_i^2\right)^{r^2}$$

has ≥ 0 coefficients

- $p(x) > 0$ on $\{x \mid g_i(x) \geq 0\} \Leftrightarrow f$ is pd
- **Result by Polya (1928):**

f **even** and pd $\Rightarrow \exists r \in \mathbb{N}$ such that $f(z) \cdot \left(\sum_i z_i^2\right)^r$ has nonnegative coefficients.

- Make $f(z)$ even by considering $f(v^2 - w^2)$. We lose positive definiteness of f with this transformation.
- Add the positive definite term $\frac{1}{2r}\left(\sum_i (v_i^4 + w_i^4)\right)^d$ to $f(v^2 - w^2)$ to make it positive definite. **Apply Polya's result.**
- The term $-\frac{1}{r}\left(\sum_i (v_i^2 - w_i^2)^2\right)^d$ ensures that the converse holds as well.

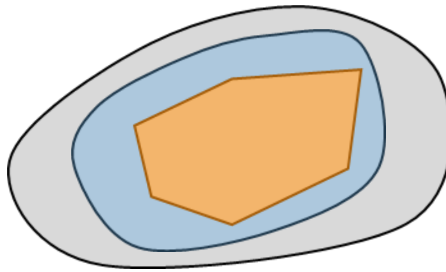
As a corollary, gives LP/SOCP-based converging hierarchies…

LP/SOCP-based alternatives to SOS

Sum of squares **(sos)**	$p(x) = z(x)^T Q z(x), Q \succcurlyeq 0$	**SDP**

DD cone $:= \{Q \mid Q_{ii} \geq \sum_{j \neq i} \lvert Q_{ij} \rvert, \forall i\}$

PSD cone$:= \{Q \mid Q \succcurlyeq 0\}$



SDD cone $:= \{Q \mid \exists$ **diagonal** D **with** $D_{ii} > 0$ **s.t.** $DQD\ dd\}$

Diagonally dominant sum of squares **(dsos)**	$p(x) = z(x)^T Q z(x), Q\ diagonally\ dominant\ (\mathrm{dd})$	**LP**
Scaled diagonally dominant sum of squares **(sdsos)**	$p(x) = z(x)^T Q z(x), Q\ scaled\ diagonally\ dominant\ (\mathrm{sdd})$	**SOCP**

[AAA, Majumdar]

LP and SOCP-based converging hierarchies for POPs

$$\min_{x \in \mathbb{R}^n} p(x)$$
$$s.t.\ g_i(x) \geq 0, i = 1, \ldots, m$$

$2d =$ maximum degree of p, g_i

opt. val. $=$

Under compactness assumptions, i.e., $\{x \mid g_i(x) \geq 0\} \subseteq B(0, R)$
For large enough r

$$\sup_{\gamma} \gamma$$

s.t. $\left(f_\gamma(v^2 - w^2) - \frac{1}{r}\left(\sum_i (v_i^2 - w_i^2)^2\right)^d + \frac{1}{2r}\left(\sum_i (v_i^4 + w_i^4)\right)^d\right) \cdot q(v, w)$ is s/dsos

$q(v, w)$ is s/dsos of degree $2r$,

where f_γ is a form in $n + m + 3$ variables and of degree $4d$

PRINCETON UNIVERSITY ORFE

The main takeaway

- Positivstellensatze by Artin (1927) and Polya (1928) are enough to produce a converging hierarchy of lower bounds for polynomial optimization problems with a bounded feasible set.
- The latter only involves polynomial multiplication.

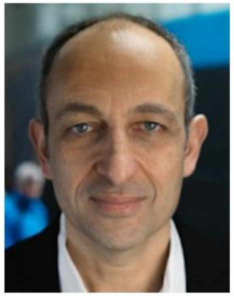
Want to know more? http://aaa.princeton.edu

You are cordially invited…

Princeton Day of Optimization
September 28, 2018

http://orfe.princeton.edu/pdo/

D. Bertsimas



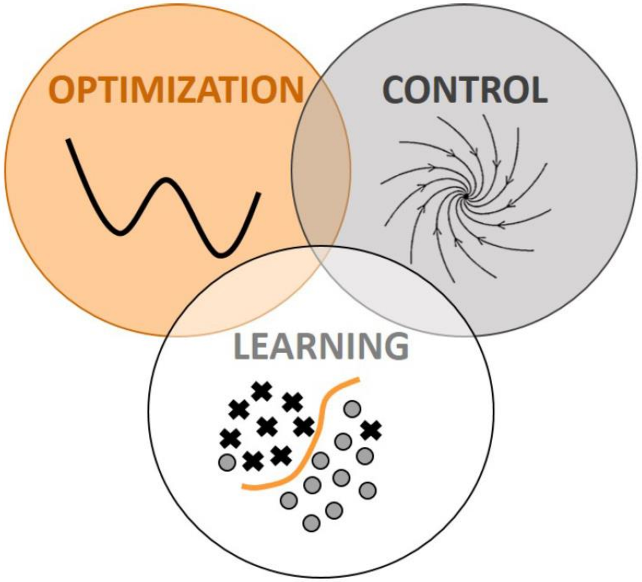
M. Dahleh

E. Hazan

P. Parrilo

B. Recht

K. Scheinberg



PRINCETON UNIVERSITY ORFE

Thank you.