The Moment-SOS Hierarchy

Jean B. Lasserre*

LAAS-CNRS, TSE & ANITI, Toulouse, France

Année de l'Optimisation, October 2024

I first would like to express my gratitude to:

– the very active Toulouse optimization research community accross University Toulouse III, N'7, Sup'Aéro, ENAC, INRA, CNRS, TSE, IMT, and ANITI,
– its regular and lively SPOT seminar.

I would also like to celebrate its recent recognition via

- CNRS medals (V. Magron, E. Pauwels)
- IUF (E. Pauwels)
- 2024 (MOS & SIAM) Lagrange Prize (J. Bolte & E. Pauwels)

☞ All that makes Toulouse a very active and important node of the french optimization research network!

At last but not least, my gratitude also to POP (LAAS research team, D. Henrion, M. Korda, V. Magron, and M. Skomra) for fruitful collaboration and friendship.

I first would like to express my gratitude to:

– the very active Toulouse optimization research community accross University Toulouse III, N'7, Sup'Aéro, ENAC, INRA, CNRS, TSE, IMT, and ANITI,
– its regular and lively SPOT seminar.
I would also like to celebrate its recent recognition via

- CNRS medals (V. Magron, E. Pauwels)
- IUF (E. Pauwels)
- 2024 (MOS & SIAM) Lagrange Prize (J. Bolte & E. Pauwels)

☞ All that makes Toulouse a very active and important node of the french optimization research network!

At last but not least, my gratitude also to POP (LAAS research team, D. Henrion, M. Korda, V. Magron, and M. Skomra) for fruitful collaboration and friendship.

I first would like to express my gratitude to:

– the very active Toulouse optimization research community accross University Toulouse III, N'7, Sup'Aéro, ENAC, INRA, CNRS, TSE, IMT, and ANITI,
– its regular and lively SPOT seminar.
I would also like to celebrate its recent recognition via

- CNRS medals (V. Magron, E. Pauwels)
- IUF (E. Pauwels)
- 2024 (MOS & SIAM) Lagrange Prize (J. Bolte & E. Pauwels)

☞ All that makes Toulouse a very active and important node of the french optimization research network!

At last but not least, my gratitude also to POP (LAAS research team, D. Henrion, M. Korda, V. Magron, and M. Skomra) for fruitful collaboration and friendship.

Born in 2000, the Moment-SOS hierarchy :

- ☞ was initially designed to help solve global optimization problems whose criteria and constraints are defined by polynomials (i.e. pbs with an algebraic description).
- ☞ It has since given birth to the field of Polynomial Optimization, now recognized in MSC classification under 90C23

Born in 2000, the Moment-SOS hierarchy :

- ☞ was initially designed to help solve global optimization problems whose criteria and constraints are defined by polynomials (i.e. pbs with an algebraic description).
  - ☞ It has since given birth to the field of Polynomial Optimization, now recognized in MSC classification under 90C23

Born in 2000, the Moment-SOS hierarchy :

- ☞ was initially designed to help solve global optimization problems whose criteria and constraints are defined by polynomials (i.e. pbs with an algebraic description).
  - ☞ It has since given birth to the field of Polynomial Optimization, now recognized in MSC classification under 90C23

However ...

This methodology also applies to solve the
Generalized Moment Problem (GMP)
(with algebraic data) whose list of applications in many area of
Science and Engineering is almost endless ...

A GMP is a problem where:

- The variables are polynomials p and/or measures μ
- The constraints are of the form

$$p(\mathbf{x}) \geq 0, \forall \mathbf{x} \in S; \quad \mathrm{supp}(\mu) \subset \Omega; \quad \int_\Omega f \, d\mu = 0,$$

for some semi-algebraic sets $S, \Omega \subset \mathbb{R}^d$ and polynomial f.

However ...

This methodology also applies to solve the
Generalized Moment Problem (GMP)
(with algebraic data) whose list of applications in many area of Science and Engineering is almost endless ...

A GMP is a problem where:

- The variables are polynomials p and/or measures μ
- The constraints are of the form

$$p(\mathbf{x}) \geq 0\,,\ \forall \mathbf{x} \in S\,; \quad \mathrm{supp}(\mu) \subset \Omega\,; \quad \int_{\Omega} f\, d\mu = 0\,,$$

for some semi-algebraic sets $S, \Omega \subset \mathbb{R}^d$ and polynomial f.

The emergence of the Moment-SOS hierarchy was made possible because of the conjunction of two factors:

- (i) ☞ Powerful Certificates of Positivity in Real Algebraic Geometry in the seventies and nineties (Krivine-Stengle, Handelman, Schmüdgen, Putinar, ...) and their dual facet in Real Analysis (on the **K**-moment problem)
- (ii) ☞ The development of Semidefinite Programming (SDP) as a powerful method (and technology) in optimization with a large number of applications and dedicated software packages. (GloptiPoly, SOSTools, YALMIP, MOSEK, SDPT3, Jump, TSSOS, ...)

The emergence of the Moment-SOS hierarchy was made possible because of the conjunction of two factors:

- (i) ☞ Powerful Certificates of Positivity in Real Algebraic Geometry in the seventies and nineties (Krivine-Stengle, Handelman, Schmüdgen, Putinar, ...) and their dual facet in Real Analysis (on the **K**-moment problem)
- (ii) ☞ The development of Semidefinite Programming (SDP) as a powerful method (and technology) in optimization with a large number of applications and dedicated software packages. (GloptiPoly, SOSTools, YALMIP, MOSEK, SDPT3, Jump, TSSOS, ...)

The emergence of the Moment-SOS hierarchy was made possible because of the conjunction of two factors:

- (i) ☞ Powerful Certificates of Positivity in Real Algebraic Geometry in the seventies and nineties (Krivine-Stengle, Handelman, Schmüdgen, Putinar, ...) and their dual facet in Real Analysis (on the **K**-moment problem)
- (ii) ☞ The development of Semidefinite Programming (SDP) as a powerful method (and technology) in optimization with a large number of applications and dedicated software packages. (GloptiPoly, SOSTools, YALMIP, MOSEK, SDPT3, Jump, TSSOS, ...)

Crucial to link (i) and (ii)

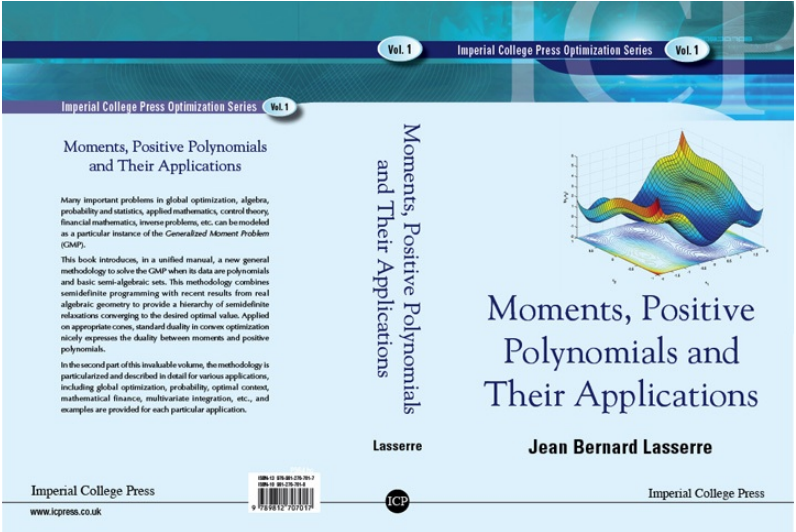
Every Sum-of-Squares polynomial (SOS)
☞ has a semidefinite representation.

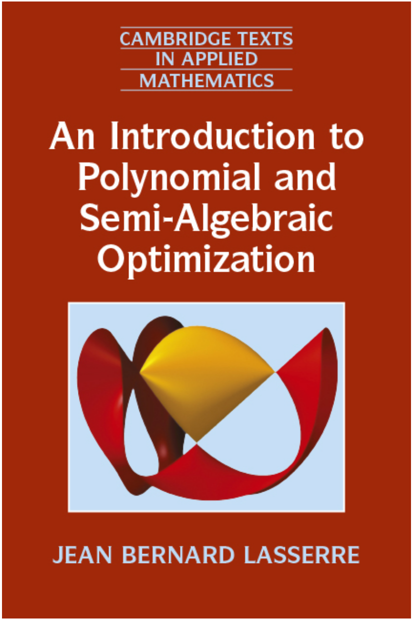
That is, one may efficiently:

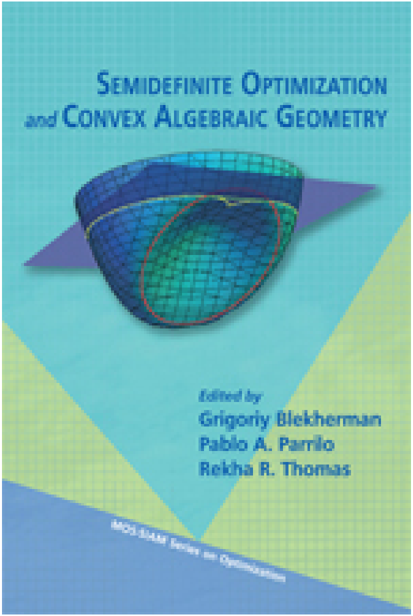
– ☞ detect whether a given polynomial is SOS, and/or

– ☞ impose that a polynomial (as a variable) is SOS,

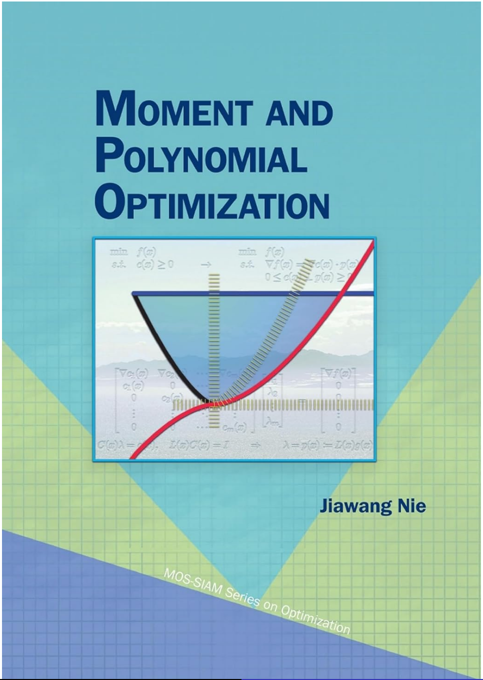
☞ Since 2000 ... several books on this topic (and related topics ...)

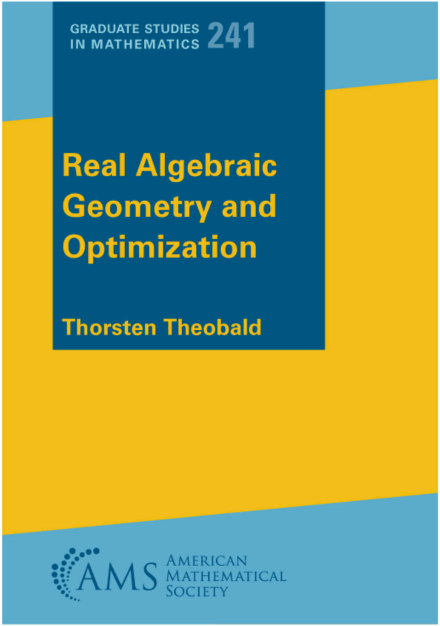

Imperial College Press Optimization Series Vol. 1
Moments, Positive Polynomials and Their Applications
Many important problems in global optimization, algebra, probability and statistics, applied mathematics, control theory, financial mathematics, inverse problems, etc. can be modeled as a particular instance of the Generalized Moment Problem (GMP).
This book introduces, in a unified manual, a new general methodology to solve the GMP when its data are polynomials and basic semi-algebraic sets. This methodology combines semidefinite programming with recent results from real algebraic geometry to provide a hierarchy of semidefinite relaxations converging to the desired optimal value. Applied on appropriate cones, standard duality in convex optimization nicely expresses the duality between moments and positive polynomials.
In the second part of this invaluable volume, the methodology is particularized and described in detail for various applications, including global optimization, probability, optimal control, mathematical finance, multivariate integration, etc., and examples are provided for each particular application.
Imperial College Press
www.icpress.co.uk
Vol. 1
Moments, Positive Polynomials and Their Applications
Lasserre
ICP
Vol. 1
Imperial College Press Optimization Series
Moments, Positive Polynomials and Their Applications
Jean Bernard Lasserre
Imperial College Press

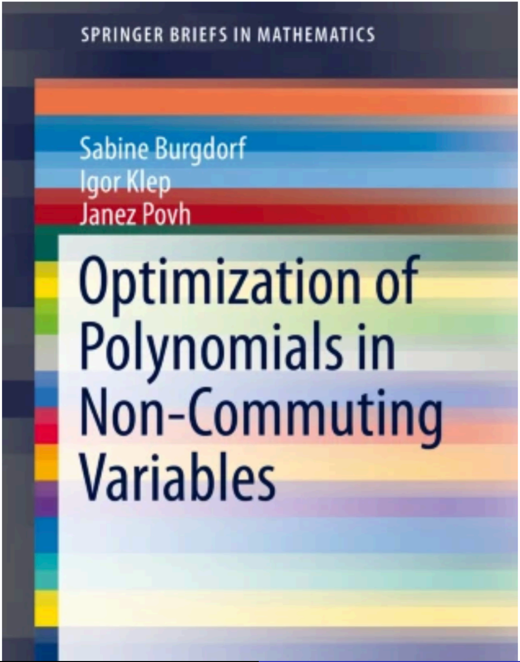

CAMBRIDGE TEXTS IN APPLIED MATHEMATICS
An Introduction to Polynomial and Semi-Algebraic Optimization
JEAN BERNARD LASSERRE

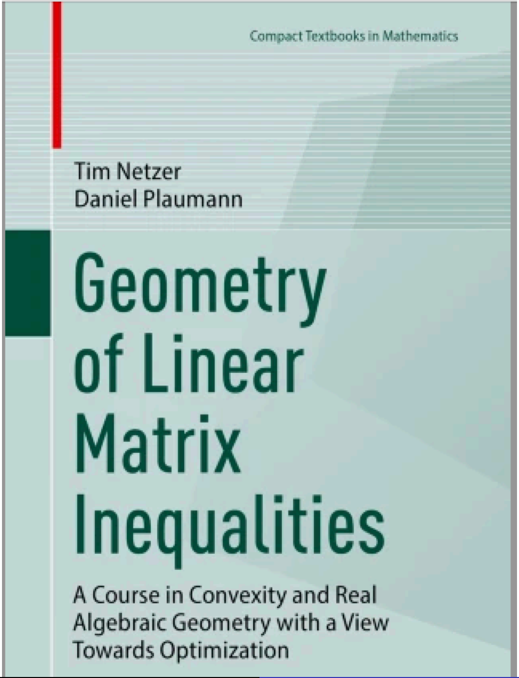

SEMIDEFINITE OPTIMIZATION
and CONVEX ALGEBRAIC GEOMETRY
Edited by
Grigoriy Blekherman
Pablo A. Parrilo
Rekha R. Thomas
MOS-SIAM Series on Optimization

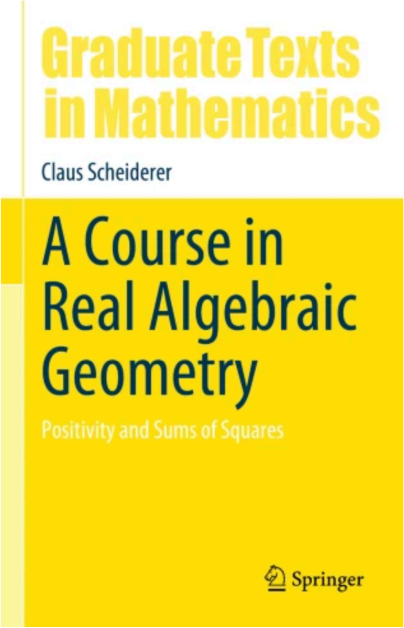

The moment-SOS hierarchy is a powerful methodology that is used to solve the Generalized Moment Problem (GMP) where the list of applications in various areas of Science and Engineering is almost endless. Initially designed for solving polynomial optimization problems (the simplest example of the GMP), it applies to solving any instance of the GMP whose description only involves semi-algebraic functions and sets. It consists of solving a sequence (a hierarchy) of convex relaxations of the initial problem, and each convex relaxation is a semidefinite program whose size increases in the hierarchy.
The goal of this book is to describe in a unified and detailed manner how this methodology applies to solving various problems in different areas ranging from Optimization, Probability, Statistics, Signal Processing, Computational Geometry, Control, Optimal Control and Analysis of a certain class of nonlinear PDEs. For each application, this unconventional methodology differs from traditional approaches and provides an unusual viewpoint. Each chapter is devoted to a particular application, where the methodology is thoroughly described and illustrated on some appropriate examples.
The exposition is kept at an appropriate level of detail to aid the different levels of readers not necessarily familiar with these tools, to better know and understand this methodology.
World Scientific
www.worldscientific.com
Q0252 hc
ISSN 2399-1593
ISBN 978-1-78634-853-1
Vol. 4
The Moment-SOS Hierarchy
Henrion
Korda
Lasserre
Series on Optimization and its Applications – Vol. 4
The Moment-SOS Hierarchy
Lectures in Probability, Statistics, Computational Geometry, Control and Nonlinear PDEs
Didier Henrion
Milan Korda
Jean B. Lasserre
World Scientific


MOMENT AND
POLYNOMIAL
OPTIMIZATION
Jiawang Nie
MOS-SIAM Series on Optimization


GRADUATE STUDIES IN MATHEMATICS 241
Real Algebraic Geometry and Optimization
Thorsten Theobald
AMS AMERICAN MATHEMATICAL SOCIETY


SPRINGER BRIEFS IN MATHEMATICS
Sabine Burgdorf
Igor Klep
Janez Povh
Optimization of Polynomials in Non-Commuting Variables


Compact Textbooks in Mathematics
Tim Netzer
Daniel Plaumann
Geometry of Linear Matrix Inequalities
A Course in Convexity and Real Algebraic Geometry with a View Towards Optimization


Graduate Texts
in Mathematics
Claus Scheiderer
A Course in
Real Algebraic
Geometry
Positivity and Sums of Squares
Springer

200pp | May 2023

Hardcover 978-1-80061-294-5 | **US$88 / £70**

eBook-Individuals 978-1-80061-296-9 | **US$70 / £55**

Visit https://doi.org/10.1142/q0382

World Scientific
Connecting Great Minds

WorldSciNet
ebooks • ejournals • databases

Methodology

- Replace the hard non convex initial problem with an EQUIVALENT Linear Program (a GMP) (but Infinite-dimensional ...)

☞ possible in a very general framework

- In turn replace the infinite-dimensional LP with a nested sequence of finite-dimensional convex relaxations $\cdots$ whose size increases.

☞ possible because problem data are algebraic

☞ whence the name “hierarchy of relaxations”

Methodology

- Replace the hard non convex initial problem with an EQUIVALENT Linear Program (a GMP) (but Infinite-dimensional ...)

  ☞ possible in a very general framework

- In turn replace the infinite-dimensional LP with a nested sequence of finite-dimensional convex relaxations $\cdots$ whose size increases.

  ☞ possible because problem data are algebraic

☞ whence the name "hierarchy of relaxations"

Methodology

- Replace the hard non convex initial problem with an EQUIVALENT Linear Program (a GMP) (but Infinite-dimensional ...)

  ☞ possible in a very general framework

- In turn replace the infinite-dimensional LP with a nested sequence of finite-dimensional convex relaxations $\cdots$ whose size increases.

  ☞ possible because problem data are algebraic

☞ whence the name “hierarchy of relaxations”

Methodology

- Replace the hard non convex initial problem with an EQUIVALENT Linear Program (a GMP) (but Infinite-dimensional ...)

  ☞ possible in a very general framework

- In turn replace the infinite-dimensional LP with a nested sequence of finite-dimensional convex relaxations $\cdots$ whose size increases.

  ☞ possible because problem data are algebraic

☞ whence the name "hierarchy of relaxations"

Solving each convex relaxation reduces to solving an SDP whose size increases in the hierarchy, hence more and more expensive to solve (even if convex) ....!

No free lunch !! ☞ The Curse of dimensionality !

Solving each convex relaxation reduces to solving an SDP whose size increases in the hierarchy, hence more and more expensive to solve (even if convex) ....!

No free lunch !! ☞ The Curse of dimensionality !

However :

- practice reveals fast and often finite convergence
- Moreover sparsity and/or symmetries can be exploited efficiently ... and so

  ☞ many large-scale problems can still be handled ...

The impact of the Moment-SOS hierarchy

On the theoretical part:

- in Mathematics: Real Analysis & Real Algebraic Geometry: ☞ Research on the Moment-Problem, positive polynomials, convex algebraic geometry.
- in Theoretical Computer Science: MOM-SOS is now a basic tool in combinatorial optimization:
  – for hardness of approximation: ☞ has become a meta-algorithm in approximation algorithms.
  – ☞ central for proving/disproving Khot's Unique Games Conjecture, and also e.g., for graph isomorphism, ...
- in Optimization: ☞ Global optimality conditions, rates of convergence in global optimization, finite convergence.

The impact of the Moment-SOS hierarchy

On the theoretical part:

- in Mathematics: Real Analysis & Real Algebraic Geometry: ☞ Research on the Moment-Problem, positive polynomials, convex algebraic geometry.
- in Theoretical Computer Science: MOM-SOS is now a basic tool in combinatorial optimization:
  – for hardness of approximation: ☞ has become a meta-algorithm in approximation algorithms.
  – ☞ central for proving/disproving Khot's Unique Games Conjecture, and also e.g., for graph isomorphism, ...
- in Optimization: ☞ Global optimality conditions, rates of convergence in global optimization, finite convergence.

The impact of the Moment-SOS hierarchy

On the theoretical part:

- in Mathematics: Real Analysis & Real Algebraic Geometry: ☞ Research on the Moment-Problem, positive polynomials, convex algebraic geometry.
- in Theoretical Computer Science: MOM-SOS is now a basic tool in combinatorial optimization:
  – for hardness of approximation: ☞ has become a meta-algorithm in approximation algorithms.
  – ☞ central for proving/disproving Khot's Unique Games Conjecture, and also e.g., for graph isomorphism, ...
- in Optimization: ☞ Global optimality conditions, rates of convergence in global optimization, finite convergence.

The impact of the Moment-SOS hierarchy

On the theoretical part:

- in Mathematics: Real Analysis & Real Algebraic Geometry: ☞ Research on the Moment-Problem, positive polynomials, convex algebraic geometry.
- in Theoretical Computer Science: MOM-SOS is now a basic tool in combinatorial optimization:
  – for hardness of approximation: ☞ has become a meta-algorithm in approximation algorithms.
  – ☞ central for proving/disproving Khot's Unique Games Conjecture, and also e.g., for graph isomorphism, ...
- in Optimization: ☞ Global optimality conditions, rates of convergence in global optimization, finite convergence.

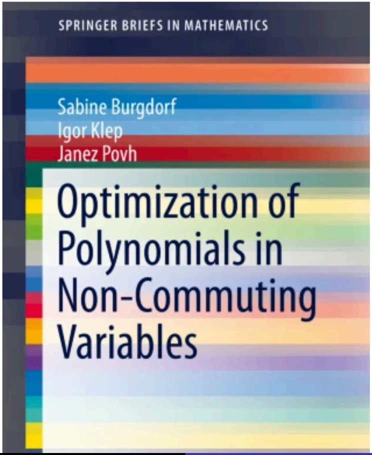
On the applications side

Many important problems in various areas of Science and Engineering can be formulated as specify instances of the Generalized Moment Problem (GMP), and ...

☞ The Moment-SOS hierarchy provides a systematic and convergent numerical scheme to solve such problems ...

In addition,

the commutative and non commutative versions of the MOM-SOS have now become an important and basic computational tool for important problems of Quantum information



For a list of some journal articles on applications of the Moment-SOS hierarchy in such various domains, see e.g.

The Moment-SOS hierarchy: Applications & Related Topics
Acta Numerica **33**, pp. 841–908, 2024.

☞ below is just a partial list of such applications

Control related applications

- Control, Optimal & Stochastic Control, Identification, Approximation of regions of attraction, etc.
- Analysis, control of certain classes of PDEs
- Evaluation of functionals of solutions of PDEs

In Approximation theory, Probability & Statistics

- Approximation: recovery of discontinuous functions from sampling
- Probability: optimal upper bounds under moment conditions (and recently bounds for large deviations), density approximation.
- Optimal Transport: distance between probabilities
- Statistics: Approximate optimal design, Super resolution

In Computer Science:

- Coding in cryptography: error correcting codes
- Packing problems: provide new best bounds (remarkable results by Bachoc, Vallentin, de Laat)
- computations in graphs: Lovász theta number
- Algorithmic Game Theory

In Approximation theory, Probability & Statistics

- Approximation: recovery of discontinuous functions from sampling
- Probability: optimal upper bounds under moment conditions (and recently bounds for large deviations), density approximation.
- Optimal Transport: distance between probabilities
- Statistics: Approximate optimal design, Super resolution

In Computer Science:

- Coding in cryptography: error correcting codes
- Packing problems: provide new best bounds (remarkable results by Bachoc, Vallentin, de Laat)
- computations in graphs: Lovász theta number
- Algorithmic Game Theory

In Computational non linear algebra

- many tensor computations
- some computations in geometry (e.g. volume of semi-algebraic sets)
- computer graphics and geometry processing

In Mathematical finance:

- Portfolio optimization and option pricing
- exit time computation for diffusions.

In Engineering

- Pattern recognition
- Computer vision
  - camera calibration
  - geometric perception in Robotics
- Motion planning in Robotics
- Signal Processing: recovery of sparse signals
- Internet of Things (IoT)
- radar and wireless communication

In Physics: bounding ground state energy

In Chemistry: bounds on stochastic chemical kinetic systems

How does it work?

☞ Illustration for Global Optimization

Consider the polynomial optimization problem:

$$\mathbf{P}: \quad f_{\min} = \min\{ f(\mathbf{x}) : \quad g_j(\mathbf{x}) \geq 0, \; j = 1, \ldots, m \}$$

for some polynomials f, $g_j \in \mathbb{R}[\mathbf{x}]$.

Why Polynomial Optimization?

After all ... **P** is just a particular case of Non Linear Programming (NLP)!

Moment-SOS hierarchy for POLYNOMIAL optimization

Consider the polynomial optimization problem:

$$\mathbf{P}: \quad f_{\min} = \min\{\, f(\mathbf{x}) : \quad g_j(\mathbf{x}) \geq 0,\ j = 1, \ldots, m\,\}$$

for some polynomials f, $g_j \in \mathbb{R}[\mathbf{x}]$.

Why Polynomial Optimization?

After all ... **P** is just a particular case of Non Linear Programming (NLP)!

True!

... if one is interested with a LOCAL optimum only!!

☞ The fact that f, g_j are POLYNOMIALS does not help much!

BUT for GLOBAL Optimization

... the picture is different!

True!

... if one is interested with a LOCAL optimum only!!

☞ The fact that f, g_j are POLYNOMIALS does not help much!

BUT for GLOBAL Optimization

... the picture is different!

True!

... if one is interested with a LOCAL optimum only!!

☞ The fact that f, g_j are POLYNOMIALS does not help much!

BUT for GLOBAL Optimization

... the picture is different!

Remember that for the GLOBAL minimum $f_{\min}$:

$$f_{\min} \;=\; \sup_{\lambda} \{\, \lambda : \; f(\mathbf{x}) - \lambda \,\geq\, 0 \quad \forall \mathbf{x} \in \mathbf{K} \}.$$

(Not true for a LOCAL minimum!)

Equivalently:

$$f_{\min} \;=\; \inf_{\mu} \{ \int_{\mathbf{K}} f\, d\mu : \mu(\mathbf{K}) = 1; \quad \mu \in \mathscr{M}(\mathbf{K})_+ \}$$

- Indeed if $f \geq f_{\min}$ for all $\mathbf{x} \in \mathbf{K}$ then $\int_{\mathbf{K}} f\, d\mu \geq f_{\min}$ for all $\mu \in \mathscr{M}(\mathbf{K})_+$ with $\mu(\mathbf{K}) = 1$.
- On the other hand, for every $\mathbf{x} \in \mathbf{K}$, $f(\mathbf{x}) = \int_{\mathbf{K}} f\, d\delta_{\mathbf{x}}$ with $\delta_x \in \mathscr{M}(\mathbf{K})_+$ (the Dirac measure at $\mathbf{x}$).

Remember that for the GLOBAL minimum $f_{\min}$:

$$f_{\min} \;=\; \sup_{\lambda} \{ \lambda : \; f(\mathbf{x}) - \lambda \geq 0 \quad \forall \mathbf{x} \in \mathbf{K} \}.$$

(Not true for a LOCAL minimum!)

Equivalently:

$$f_{\min} \;=\; \inf_{\mu} \{ \int_{\mathbf{K}} f \, d\mu : \mu(\mathbf{K}) = 1; \quad \mu \in \mathscr{M}(\mathbf{K})_+ \}$$

- Indeed if $f \geq f_{\min}$ for all $\mathbf{x} \in \mathbf{K}$ then $\int_{\mathbf{K}} f \, d\mu \geq f_{\min}$ for all $\mu \in \mathscr{M}(\mathbf{K})_+$ with $\mu(\mathbf{K}) = 1$.
- On the other hand, for every $\mathbf{x} \in \mathbf{K}$, $f(\mathbf{x}) \;=\; \int_{\mathbf{K}} f \, d\delta_{\mathbf{x}}$ with $\delta_{\mathbf{x}} \in \mathscr{M}(\mathbf{K})_+$ (the Dirac measure at $\mathbf{x}$).

Remember that for the GLOBAL minimum $f_{\min}$:

$$f_{\min} \quad = \quad \sup_{\lambda} \{ \lambda : \; f(\mathbf{x}) - \lambda \geq 0 \quad \forall \mathbf{x} \in \mathbf{K} \}.$$

(Not true for a LOCAL minimum!)

Equivalently:

$$f_{\min} \quad = \quad \inf_{\mu} \{ \int_{\mathbf{K}} f \, d\mu : \mu(\mathbf{K}) = 1; \quad \mu \in \mathscr{M}(\mathbf{K})_+ \}$$

- Indeed if $f \geq f_{\min}$ for all $\mathbf{x} \in \mathbf{K}$ then $\int_{\mathbf{K}} f \, d\mu \geq f_{\min}$ for all $\mu \in \mathscr{M}(\mathbf{K})_+$ with $\mu(\mathbf{K}) = 1$.
- On the other hand, for every $\mathbf{x} \in \mathbf{K}$, $f(\mathbf{x}) = \int_{\mathbf{K}} f \, d\delta_{\mathbf{x}}$ with $\delta_{\mathbf{x}} \in \mathscr{M}(\mathbf{K})_+$ (the Dirac measure at $\mathbf{x}$).

Remember that for the GLOBAL minimum $f_{\min}$:

$$f_{\min} \quad = \quad \sup_{\lambda} \{ \lambda : \ f(\mathbf{x}) - \lambda \geq 0 \quad \forall \mathbf{x} \in \mathbf{K} \}.$$

(Not true for a LOCAL minimum!)

Equivalently:

$$f_{\min} \quad = \quad \inf_{\mu} \{ \int_{\mathbf{K}} f \, d\mu : \mu(\mathbf{K}) = 1; \quad \mu \in \mathscr{M}(\mathbf{K})_+ \}$$

• Indeed if $f \geq f_{\min}$ for all $\mathbf{x} \in \mathbf{K}$ then $\int_{\mathbf{K}} f \, d\mu \geq f_{\min}$ for all $\mu \in \mathscr{M}(\mathbf{K})_+$ with $\mu(\mathbf{K}) = 1$.

• On the other hand, for every $\mathbf{x} \in \mathbf{K}$, $f(\mathbf{x}) \;=\; \int_{\mathbf{K}} f \, d\delta_{\mathbf{x}}$ with $\delta_x \in \mathscr{M}(\mathbf{K})_+$ (the Dirac measure at $\mathbf{x}$).

Therefore: (Real Analysis view)

$$f_{\min} = \inf_{\mu} \left\{ \int_{\mathbf{K}} f\, d\mu \,:\, \mu(\mathbf{K}) = 1; \quad \mu \in \mathscr{M}(\mathbf{K})_+ \right\}$$
$$= \inf_{\mu} \left\{ \langle f, \mu \rangle \,:\, \langle 1, \mu \rangle = 1;\ \mu \geq 0; \quad \mu \in \mathscr{M}(\mathbf{K}) \right\}$$

☞ A "primal" LP on measures on **K**

but also: (Real Algebraic view)

$$f_{\min} = \sup_{\lambda} \left\{ \lambda \,:\, f(\mathbf{x}) - \lambda \geq 0 \quad \forall \mathbf{x} \in \mathbf{K} \right\}.$$
$$= \sup_{\lambda} \left\{ \lambda \,:\, f - \lambda \in \mathscr{P}(\mathbf{K})_+ \right\}$$

☞ Its "dual" LP on polynomials positive on **K**

☞ Two Infinite-dimensional LPs in DUALITY!

Therefore: (Real Analysis view)

$$f_{\min} = \inf_{\mu} \left\{ \int_{\mathbf{K}} f\, d\mu : \mu(\mathbf{K}) = 1; \quad \mu \in \mathscr{M}(\mathbf{K})_+ \right\}$$
$$= \inf_{\mu} \left\{ \langle f, \mu \rangle : \langle 1, \mu \rangle = 1; \ \mu \geq 0; \quad \mu \in \mathscr{M}(\mathbf{K}) \right\}$$

☞ A "primal" LP on measures on **K**

but also: (Real Algebraic view)

$$f_{\min} = \sup_{\lambda} \left\{ \lambda : \ f(\mathbf{x}) - \lambda \geq 0 \quad \forall \mathbf{x} \in \mathbf{K} \right\}.$$
$$= \sup_{\lambda} \left\{ \lambda : \ f - \lambda \in \mathscr{P}(\mathbf{K})_+ \right\}$$

☞ Its "dual" LP on polynomials positive on **K**

☞ Two Infinite-dimensional LPs in DUALITY!

Therefore: (Real Analysis view)

$$
\begin{aligned}
f_{\min} &= \inf_{\mu} \left\{ \int_{\mathbf{K}} f\, d\mu \,:\, \mu(\mathbf{K}) = 1; \quad \mu \in \mathscr{M}(\mathbf{K})_+ \right\} \\
&= \inf_{\mu} \left\{ \langle f, \mu \rangle \,:\, \langle 1, \mu \rangle = 1;\ \mu \geq 0; \quad \mu \in \mathscr{M}(\mathbf{K}) \right\}
\end{aligned}
$$

☞ A "primal" LP on measures on **K**

but also: (Real Algebraic view)

$$
\begin{aligned}
f_{\min} &= \sup_{\lambda} \left\{ \lambda \,:\, f(\mathbf{x}) - \lambda \geq 0 \quad \forall \mathbf{x} \in \mathbf{K} \right\}. \\
&= \sup_{\lambda} \left\{ \lambda \,:\, f - \lambda \in \mathscr{P}(\mathbf{K})_+ \right\}
\end{aligned}
$$

☞ Its "dual" LP on polynomials positive on **K**

☞ Two Infinite-dimensional LPs in DUALITY!

Remember ...

This infinite-dimensional LP (primal or dual) is our first step in the methodology of the Moment-SOS hierarchy, where one replaces the initial hard problem with an infinite-dimensional LP (a GMP)!

Notice how simple is the corresponding GMP!

$$\inf_{\mu} \{ \langle f, \mu \rangle : \langle 1, \mu \rangle = 1; \ \mu \geq 0; \quad \mu \in \mathscr{M}(\mathbf{K}) \}$$

1 equality constraint !

☞ Global Optimization is the simplest instance of the GMP

Remember ...

This infinite-dimensional LP (primal or dual) is
our first step in the methodology of the Moment-SOS hierarchy,
where one replaces
the initial hard problem with an infinite-dimensional LP (a GMP)!

Notice how simple is the corresponding GMP!

$$\inf_{\mu} \{ \langle f, \mu \rangle : \langle 1, \mu \rangle \; = 1; \; \mu \geq 0; \quad \mu \in \mathscr{M}(\mathbf{K}) \}$$

1 equality constraint !

☞ Global Optimization is the simplest instance of the GMP and so to compute (or approximate) $f_{\min}$ ...

☞ one needs to handle EFFICIENTLY the difficult constraint

$$f(\mathbf{x}) - \lambda \geq 0 \quad \forall \mathbf{x} \in \mathbf{K}, \quad (f - \lambda \in \mathscr{P}(\mathbf{K})_+)$$

i.e. one needs

☞ TRACTABLE CERTIFICATES of POSITIVITY on **K**

for the polynomial $\mathbf{x} \mapsto f(\mathbf{x}) - \lambda$!

and so to compute (or approximate) $f_{\min}$ ...

☞ one needs to handle EFFICIENTLY the difficult constraint

$$f(\mathbf{x}) - \lambda \geq 0 \quad \forall \mathbf{x} \in \mathbf{K}, \quad (f - \lambda \in \mathscr{P}(\mathbf{K})_+)$$

i.e. one needs

☞ TRACTABLE CERTIFICATES of POSITIVITY on **K**

for the polynomial $\mathbf{x} \mapsto f(\mathbf{x}) - \lambda$!

or ...

one needs to handle EFFICIENTLY measures supported on $\mathbf{K}$, to solve the primal LP on measures ... (e.g., via its moments)

☞ i.e., one needs TRACTABLE CERTIFICATES for a real sequence $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}}$ to be moments of a measure μ on $\mathbf{K}$, i.e.,

$$\exists \mu : \quad y_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu, \quad \forall \alpha \in \mathbb{N}^n.$$

or ...

one needs to handle EFFICIENTLY measures supported on **K**, to solve the primal LP on measures ... (e.g., via its moments)

☞ i.e., one needs TRACTABLE CERTIFICATES for a real sequence $\mathbf{y} = (y_\alpha)_{\alpha\in\mathbb{N}}$ to be moments of a measure μ on **K**, i.e.,

$$\exists\mu : \quad y_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu \,, \quad \forall \alpha \in \mathbb{N}^n .$$

REAL ALGEBRAIC GEOMETRY helps!!!!

Indeed, POWERFUL CERTIFICATES OF POSITIVITY EXIST!

Moreover .... and importantly,

Such certificates are amenable to PRACTICAL COMPUTATION!

($\star$ Stronger Positivstellensatzë exist for analytic functions but (so far) are useless from a computational viewpoint.)

REAL ALGEBRAIC GEOMETRY helps!!!!

Indeed, POWERFUL CERTIFICATES OF POSITIVITY EXIST!

Moreover .... and importantly,

Such certificates are amenable to PRACTICAL COMPUTATION!

($\star$ Stronger Positivstellensatzë exist for analytic functions but (so far) are useless from a computational viewpoint.)

REAL ALGEBRAIC GEOMETRY helps!!!!

Indeed, POWERFUL CERTIFICATES OF POSITIVITY EXIST!

Moreover .... and importantly,

Such certificates are amenable to PRACTICAL COMPUTATION!

($\star$ Stronger Positivstellensatzë exist for analytic functions but (so far) are useless from a computational viewpoint.)

Most of such certificates of positivity in Real Algebraic geometry have a dual facet in Real Analysis on the **K**-Moment problem.

The facet on the **K**-Moment problem

☞ is crucial to solve the LP on measures

while

the dual facet on positive polynomials

☞ is crucial to solve the LP on polynomials.

Most of such certificates of positivity in Real Algebraic geometry have a dual facet in Real Analysis on the **K**-Moment problem.

The facet on the **K**-Moment problem
☞ is crucial to solve the LP on measures
while
the dual facet on positive polynomials
☞ is crucial to solve the LP on polynomials.

SOS-based certificate

A polynomial p is a sum-of-squares (SOS) if and only if

$$p(\mathbf{x}) = \sum_{k=1}^{s} q_k(\mathbf{x})^2, \quad \forall \mathbf{x} \in \mathbb{R}^n,$$

for some polynomials q_k.

☞ Detecting whether a given polynomial p is SOS can be done efficiently by solving a SEMIDEFINITE PROGRAM

☞ A SEMIDEFINITE PROGRAM (SDP) is a CONIC, CONVEX OPTIMIZATION PROBLEM that can be solved EFFICIENTLY (up to arbitrary fixed precision)

SOS-based certificate

A polynomial p is a sum-of-squares (SOS) if and only if

$$p(\mathbf{x}) = \sum_{k=1}^{s} q_k(\mathbf{x})^2, \quad \forall \mathbf{x} \in \mathbb{R}^n,$$

for some polynomials q_k.

☞ Detecting whether a given polynomial p is SOS can be done efficiently by solving a SEMIDEFINITE PROGRAM

☞ A SEMIDEFINITE PROGRAM (SDP) is a CONIC, CONVEX OPTIMIZATION PROBLEM that can be solved EFFICIENTLY (up to arbitrary fixed precision)

$$\text{Let } \mathbf{K} := \{\, \mathbf{x} : \; g_j(\mathbf{x}) \,\geq\, 0, \quad j = 1, \ldots, m \,\}$$

be compact (with $g_1(\mathbf{x}) = M - \|\mathbf{x}\|^2$, so that $\mathbf{K} \subset \mathbf{B}(0, M)$).

Theorem (Putinar's Positivstellensatz)

If $f \in \mathbb{R}[\mathbf{x}]$ is strictly positive ($f > 0$) on $\mathbf{K}$ then:

$$\dagger \quad f(\mathbf{x}) \,=\, \sigma_0(\mathbf{x}) + \sum_{j=1}^{m} \sigma_j(\mathbf{x})\, g_j(\mathbf{x}), \qquad \forall \mathbf{x} \in \mathbb{R}^n,$$

for some SOS polynomials $(\sigma_j) \subset \mathbb{R}[\mathbf{x}]$.

$$\text{Let } \mathbf{K} := \{ \mathbf{x} : \; g_j(\mathbf{x}) \geq 0, \quad j = 1, \ldots, m \}$$

be compact (with $g_1(\mathbf{x}) = M - \|\mathbf{x}\|^2$, so that $\mathbf{K} \subset \mathbf{B}(0, M)$).

Theorem (Putinar's Positivstellensatz)

If $f \in \mathbb{R}[\mathbf{x}]$ is strictly positive ($f > 0$) on $\mathbf{K}$ then:

$$\dagger \quad f(\mathbf{x}) \; = \; \sigma_0(\mathbf{x}) + \sum_{j=1}^{m} \sigma_j(\mathbf{x}) \, g_j(\mathbf{x}), \qquad \forall \mathbf{x} \in \mathbb{R}^n,$$

for some SOS polynomials $(\sigma_j) \subset \mathbb{R}[\mathbf{x}]$.

However ... In Putinar's theorem

... nothing is said on the DEGREE of the SOS polynomials (σ_j)!

BUT ... GOOD news ..!!

☞ Testing whether † holds
for some SOS $(\sigma_j) \subset \mathbb{R}[\mathbf{x}]$ with a degree bound,

is SOLVING an SDP!

However ... In Putinar's theorem

... nothing is said on the DEGREE of the SOS polynomials (σ_j)!

BUT ... GOOD news ..!!

☞ Testing whether † holds
for some SOS $(\sigma_j) \subset \mathbb{R}[\mathbf{x}]$ with a degree bound,

is SOLVING an SDP!

Dual side: The K-moment problem

Given a real sequence $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, does there exist a Borel measure μ on $\mathbf{K}$ such that

$$\dagger \quad y_\alpha = \int_{\mathbf{K}} x_1^{\alpha_1} \cdots x_n^{\alpha_n} \, d\mu, \qquad \forall \alpha \in \mathbb{N}^n \quad ?$$

If yes then y is said to have
a representing measure supported on $\mathbf{K}$.

Let $\mathbf{K} := \{ \mathbf{x} : \; g_j(\mathbf{x}) \; \geq \; 0, \quad j = 1, \ldots, m \}$

be compact (with $g_1(\mathbf{x}) = M - \|\mathbf{x}\|^2$, so that $\mathbf{K} \subset \mathbf{B}(0, M)$).

Theorem (Dual side of Putinar's Theorem)

A sequence $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, has a representing measure supported on $\mathbf{K}$ IF AND ONLY IF for every $d = 0, 1, \ldots$

$$(\star) \qquad \mathbf{M}_d(y) \succeq 0 \quad \text{and} \quad \mathbf{M}_d(g_j\,\mathbf{y}) \succeq 0, \quad j = 1, \ldots, m.$$

☞ The real symmetric matrix $\mathbf{M}_2(y)$ is called the MOMENT MATRIX associated with the sequence y

☞ The real symmetric matrix $\mathbf{M}_d(g_j\,\mathbf{y})$ is called the LOCALIZING MATRIX associated with the sequence y and the polynomial g_j.

$$\text{Let } \mathbf{K} := \{\, \mathbf{x} : \; g_j(\mathbf{x}) \geq 0, \quad j = 1, \ldots, m \,\}$$

be compact (with $g_1(\mathbf{x}) = M - \|\mathbf{x}\|^2$, so that $\mathbf{K} \subset \mathbf{B}(0, M)$).

Theorem (Dual side of Putinar's Theorem)

A sequence $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, has a representing measure supported on $\mathbf{K}$ IF AND ONLY IF for every $d = 0, 1, \ldots$

$$(\star) \qquad \mathbf{M}_d(y) \succeq 0 \quad \textit{and} \quad \mathbf{M}_d(g_j\, \mathbf{y}) \succeq 0, \quad j = 1, \ldots, m.$$

- ☞ The real symmetric matrix $\mathbf{M}_2(y)$ is called the MOMENT MATRIX associated with the sequence y
- ☞ The real symmetric matrix $\mathbf{M}_d(g_j\, \mathbf{y})$ is called the LOCALIZING MATRIX associated with the sequence y and the polynomial g_j.

$$\text{Let } \mathbf{K} := \{\, \mathbf{x} : \; g_j(\mathbf{x}) \geq 0, \quad j = 1, \ldots, m \,\}$$

be compact (with $g_1(\mathbf{x}) = M - \|\mathbf{x}\|^2$, so that $\mathbf{K} \subset \mathbf{B}(0, M)$).

Theorem (Dual side of Putinar's Theorem)

A sequence $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, has a representing measure supported on $\mathbf{K}$ IF AND ONLY IF for every $d = 0, 1, \ldots$

$$(\star) \qquad \mathbf{M}_d(y) \succeq 0 \quad \text{and} \quad \mathbf{M}_d(g_j\,\mathbf{y}) \succeq 0, \quad j = 1, \ldots, m.$$

☞ The real symmetric matrix $\mathbf{M}_2(y)$ is called the MOMENT MATRIX associated with the sequence y

☞ The real symmetric matrix $\mathbf{M}_d(g_j\,\mathbf{y})$ is called the LOCALIZING MATRIX associated with the sequence y and the polynomial g_j.

Let $\mathbf{K} := \{\,\mathbf{x} : \; g_j(\mathbf{x}) \,\geq\, 0, \quad j = 1, \ldots, m\,\}$

be compact (with $g_1(\mathbf{x}) = M - \|\mathbf{x}\|^2$, so that $\mathbf{K} \subset \mathbf{B}(0, M)$).

Theorem (Dual side of Putinar's Theorem)

A sequence $\mathbf{y} = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, *has a representing measure supported on* $\mathbf{K}$ *IF AND ONLY IF for every* $d = 0, 1, \ldots$

$$(\star) \qquad \mathbf{M}_d(y) \succeq 0 \quad \text{and} \quad \mathbf{M}_d(g_j\,\mathbf{y}) \succeq 0, \quad j = 1, \ldots, m.$$

☞ The real symmetric matrix $\mathbf{M}_2(y)$ is called the MOMENT MATRIX associated with the sequence y

☞ The real symmetric matrix $\mathbf{M}_d(g_j\,\mathbf{y})$ is called the LOCALIZING MATRIX associated with the sequence y and the polynomial g_j.

Remarkably:

The Necessary & Sufficient conditions $(\star)$ for existence of a representing measure are stated only in terms of

☞ countably many LINEAR MATRIX INEQUALITIES (LMI) on the sequence y !

(nowhere in the conditions is mentioned the unknown representing measure)

☞ An LMI "$\mathbf{A}(\mathbf{x}) \succeq 0$" reads "$\mathbf{A}_0 + \sum_{i=1}^{n} \mathbf{A}_i \, \mathbf{x}_i \succeq 0$" where the $\mathbf{A}_i$'s are real symmetric matrices and "$\succeq 0$" means "is psd".

☞ In addition: $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}(\mathbf{x}) \succeq 0\}$ is a CONVEX SET (called a spectrahedron)

Remarkably:

The Necessary & Sufficient conditions ($\star$) for existence of a representing measure are stated only in terms of

☞ countably many LINEAR MATRIX INEQUALITIES (LMI) on the sequence y !

(nowhere in the conditions is mentioned the unknown representing measure)

☞ An LMI "$\mathbf{A}(\mathbf{x}) \succeq 0$" reads "$\mathbf{A}_0 + \sum_{i=1}^{n} \mathbf{A}_i \, \mathbf{x}_i \succeq 0$" where the $\mathbf{A}_i$'s are real symmetric matrices and "$\succeq 0$" means "is psd".

☞ In addition: $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}(\mathbf{x}) \succeq 0\}$ is a CONVEX SET (called a spectrahedron)

Remarkably:

The Necessary & Sufficient conditions $(\star)$ for existence of a representing measure are stated only in terms of

☞ countably many LINEAR MATRIX INEQUALITIES (LMI) on the sequence y !

(nowhere in the conditions is mentioned the unknown representing measure)

☞ An LMI "$\mathbf{A}(\mathbf{x}) \succeq 0$" reads "$\mathbf{A}_0 + \sum_{i=1}^{n} \mathbf{A}_i \, \mathbf{x}_i \succeq 0$" where the $\mathbf{A}_i$'s are real symmetric matrices and "$\succeq 0$" means "is psd".

☞ In addition: $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}(\mathbf{x}) \succeq 0\}$ is a CONVEX SET (called a spectrahedron)

ALGEBRAIC SIDE

POSITIVITY ON K

$$f(x) = \sum_\alpha f_\alpha \, x^\alpha$$

$$f > 0 \text{ on } K \,?$$

CHARACTERIZE THOSE f

ALGEBRAIC SIDE

POSITIVITY ON K

$$f(x) = \sum_\alpha f_\alpha x^\alpha$$

$f \geq 0$ on K ?

CHARACTERIZE THOSE f

FUNCTIONAL ANALYSIS

THE K-MOMENT PROBLEM

$$y = (y_\alpha), \quad \alpha \in \mathbb{N}^n$$

$$y_\alpha \overset{?}{=} \int_K x^\alpha \, d\mu \quad \forall \alpha$$

for some μ

CHARACTERIZE THOSE y

DUALITY $\langle f, y \rangle = \sum_\alpha f_\alpha y_\alpha$

• In fact, polynomials NONNEGATIVE ON A SET $\mathbf{K} \subset \mathbb{R}^n$ are ubiquitous. They also appear in many important applications (outside optimization),

. . . modeled as

particular instances of the so called
Generalized Moment Problem, among which:
Probability, Optimal and Robust Control, Game theory, Signal processing, multivariate integration, etc.

☞ Whence the list of applications of the Moment-SOS hierarchy mentioned at the beginning.

How does it work? (algebraic facet)

Remember that for the GLOBAL minimum $f_{\min}$:

$$f_{min} \quad = \quad \sup_{\lambda} \{ \lambda : \; f(\mathbf{x}) - \lambda \geq 0 \quad \forall \mathbf{x} \in \mathbf{K} \}.$$

Then for each d solve:

$$\rho_d = \sup_{\lambda, \sigma_j} \{ \lambda : \; f(\mathbf{x}) - \lambda = \sigma_0(\mathbf{x}) + \sum_{j=1}^{m} \sigma_j(\mathbf{x}) \, g_j(\mathbf{x}), \; \forall \mathbf{x} \in \mathbb{R}^n$$

$$\deg(\sigma_j \, g_j) \leq 2d \,, \quad j = 0, \ldots, m \}$$ ☞ SDP !

☞ $\rho_d \leq \rho_{d+1} \leq f_{min}$ for all d and $\rho_d \uparrow f_{min}$ as $d \to \infty$.

How does it work? (Real Analysis facet)

Alternatively and equivalently, remember that:

$$f_{min} = \inf_{\mu} \left\{ \int_{\mathbf{K}} f \, d\mu \; : \; \mu(\mathbf{K}) = 1 \right\}$$

Then for each d solve:

$$\rho_d^* = \inf_{y} \{ \langle y, f \rangle : \quad (\text{think of } \textstyle\int f d\mu)$$

$$\left.\begin{aligned} y_0 &= 1 \\ \mathbf{M}_d(y) &\succeq 0 \\ \mathbf{M}_{d-t_j}(g_j\, y) &\succeq 0 \quad \forall j = 1, \ldots, m \end{aligned}\right\} \Leftarrow y_\alpha = \int_{\mathbf{K}} x^\alpha d\mu \}$$

☞ where $y = (y_\alpha)_{\alpha \in \mathbb{N}^n_{2d}}$ is a real sequence of pseudo-moments up to degree $2d$.

Theorem (Lass 2000)

☞ $\rho_d \leq \rho_d^* \leq f_{min}$ *for all d and* $\rho_d^* \uparrow f_{min}$ *as* $d \to \infty$.

Moreover, generically $\rho_d^* = f_{min}$ *and one may extract global minimizers from the optimal (truncated moment) solution* y^*.

☞ In fact ... FINITE CONVERGENCE is generic!

Theorem (Marshall, Nie)

Let $\mathbf{x}^* \in \mathbf{K}$ *be a global minimizer of*

$$\mathbf{P}: \quad f^* = \min\{f(\mathbf{x}) : g_j(\mathbf{x}) \geq 0, \; j = 1, \ldots, m\}.$$

and assume that:

(i) *The gradients* $\{\nabla g_j(\mathbf{x}^*)\}$ *are linearly independent,*

(ii) *Strict complementarity holds (*$\lambda_j^* \, g_j(\mathbf{x}^*) = 0$ *for all j.)*

(iii) *Second-order sufficiency conditions hold at* $(\mathbf{x}^*, \lambda^*) \in \mathbf{K} \times \mathbb{R}_+^m$.

Then $f(\mathbf{x}) - f^* = \sigma_0^*(\mathbf{x}) + \sum_{j=1}^{m} \sigma_j^*(\mathbf{x}) g_j(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^n$, *for some SOS polynomials* $\{\sigma_j^*\}$.

Moreover, the conditions (i)-(ii)-(iii) HOLD GENERICALLY!

Theorem (Marshall, Nie)

Let $\mathbf{x}^* \in \mathbf{K}$ *be a global minimizer of*

$$\mathbf{P}: \quad f^* = \min\{f(\mathbf{x}) : g_j(\mathbf{x}) \geq 0, \, j = 1, \ldots, m\}.$$

and assume that:

(i) *The gradients* $\{\nabla g_j(\mathbf{x}^*)\}$ *are linearly independent,*

(ii) *Strict complementarity holds* ($\lambda_j^* \, g_j(\mathbf{x}^*) = 0$ *for all* j.)

(iii) *Second-order sufficiency conditions hold at* $(\mathbf{x}^*, \lambda^*) \in \mathbf{K} \times \mathbb{R}_+^m$.

Then $f(\mathbf{x}) - f^* = \sigma_0^*(\mathbf{x}) + \sum_{j=1}^{m} \sigma_j^*(\mathbf{x}) g_j(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^n$, *for some SOS polynomials* $\{\sigma_j^*\}$.

Moreover, the conditions (i)-(ii)-(iii) HOLD GENERICALLY!

Theorem (Marshall, Nie)

Let $\mathbf{x}^* \in \mathbf{K}$ *be a global minimizer of*

$$\mathbf{P}: \quad f^* = \min\{f(\mathbf{x}) : g_j(\mathbf{x}) \geq 0, \; j = 1, \ldots, m\}.$$

and assume that:

(i) *The gradients* $\{\nabla g_j(\mathbf{x}^*)\}$ *are linearly independent,*

(ii) *Strict complementarity holds (*$\lambda_j^* \, g_j(\mathbf{x}^*) = 0$ *for all j.)*

(iii) *Second-order sufficiency conditions hold at* $(\mathbf{x}^*, \lambda^*) \in \mathbf{K} \times \mathbb{R}_+^m$.

Then $f(\mathbf{x}) - f^* = \sigma_0^*(\mathbf{x}) + \sum_{j=1}^m \sigma_j^*(\mathbf{x}) g_j(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^n$, *for some SOS polynomials* $\{\sigma_j^*\}$.

Moreover, the conditions (i)-(ii)-(iii) HOLD GENERICALLY!

Hence, FINITE CONVERGENCE

$\rho_d = \rho_d^* = f_{min}$ for some d

is generic!

In particular, at every global minimizer $\mathbf{x}^* \in \mathbf{K}$:

$$\nabla f(\mathbf{x}^*) = \sum_{j=1}^{m} \underbrace{\sigma_j^*(\mathbf{x}^*)}_{\lambda_j^* \geq 0} \nabla g_j(\mathbf{x}^*)$$

and so

- One obtains a GLOBAL OPTIMALITY condition à la Karush-Kuhn-Tucker (KKT) with SOS multipliers σ_j^* (instead of SCALARS in KKT)
- Valid for non-convex and discrete optimization !

Hence, FINITE CONVERGENCE

$$\rho_d = \rho_d^* = f_{min} \text{ for some } d$$

is generic!

In particular, at every global minimizer $\mathbf{x}^* \in \mathbf{K}$:

$$\nabla f(\mathbf{x}^*) = \sum_{j=1}^{m} \underbrace{\sigma_j^*(\mathbf{x}^*)}_{\lambda_j^* \geq 0} \nabla g_j(\mathbf{x}^*)$$

and so

- One obtains a GLOBAL OPTIMALITY condition à la Karush-Kuhn-Tucker (KKT) with SOS multipliers σ_j^* (instead of SCALARS in KKT)
- Valid for non-convex and discrete optimization !

☞ In KKT conditions, $\nabla f(\mathbf{x}^*) = \sum_{j=1}^{m} \lambda_j^* \nabla g_j(\mathbf{x}^*)$ at a local minimizer $\mathbf{x}^*$. BUT except in the CONVEX case,

$$\underbrace{f(\mathbf{x}) - f^* - \sum_{j=1}^{m} \lambda_j^* g_j(\mathbf{x})}_{\textit{Lagrangian}} \not\geq 0\,, \quad \forall \mathbf{x} \in \mathbb{R}^n\,,$$

whereas in Putinar's certificate of global optimality:

$$f(\mathbf{x}) - f^* - \sum_{j=1}^{m} \sigma_j^*(\mathbf{x}) g_j(\mathbf{x}) \quad (= \sigma_0^*(\mathbf{x})) \geq 0, \quad \forall \mathbf{x} \in \mathbb{R}^n\,.$$

That is, any global minimizer of f on $\mathbf{K}$ is a global optimizer of the extended Lagrangian

$$\mathbf{x} \mapsto f(\mathbf{x}) - f^* - \sum_{j=1}^{m} \sigma_j^*(\mathbf{x}) g_j(\mathbf{x})\,, \quad \mathbf{x} \in \mathbb{R}^n$$

on the whole $\mathbb{R}^n$.

☞ In KKT conditions, $\nabla f(\mathbf{x}^*) = \sum_{j=1}^{m} \lambda_j^* \nabla g_j(\mathbf{x}^*)$ at a local minimizer $\mathbf{x}^*$. BUT except in the CONVEX case,

$$\underbrace{f(\mathbf{x}) - f^* - \sum_{j=1}^{m} \lambda_j^* g_j(\mathbf{x})}_{\text{Lagrangian}} \not\geq 0\,, \quad \forall \mathbf{x} \in \mathbb{R}^n\,,$$

whereas in Putinar's certificate of global optimality:

$$f(\mathbf{x}) - f^* - \sum_{j=1}^{m} \sigma_j^*(\mathbf{x}) g_j(\mathbf{x}) \quad (= \sigma_0^*(\mathbf{x})) \geq 0, \quad \forall \mathbf{x} \in \mathbb{R}^n\,.$$

That is, any global minimizer of f on $\mathbf{K}$ is a global optimizer of the extended Lagrangian

$$\mathbf{x} \mapsto f(\mathbf{x}) - f^* - \sum_{j=1}^{m} \sigma_j^*(\mathbf{x}) g_j(\mathbf{x})\,, \quad \mathbf{x} \in \mathbb{R}^n$$

on the whole $\mathbb{R}^n$.

Observe that:

Except for compactness, no other assumption on **K**!

What matters is an ALGEBRAIC DESCRIPTION of **K**

via equality and/or inequality constraints !

☞ Hence **K** can be non convex, non-connected.
For instance, binary 0/1 variables are modelled via the simple quadratic equality constraint as $x_i^2 - x_i = 0$

Observe that:

Except for compactness, no other assumption on **K**!

What matters is an ALGEBRAIC DESCRIPTION of **K**

via equality and/or inequality constraints !

☞ Hence **K** can be non convex, non-connected.

For instance, binary $0/1$ variables are modelled via the simple quadratic equality constraint as $x_i^2 - x_i = 0$

An important observation

**The MOMENT-SOS approach
is a GENERAL PURPOSE method**

AIMING AT SOLVING NP-hard PROBLEMS

... and ANY GENERAL PURPOSE approach

should have the HIGHLY DESIRABLE feature

to behave efficiently for problems considered " EASY"!

Otherwise ... would you buy such a package?

An important observation

The MOMENT-SOS approach is a GENERAL PURPOSE method

AIMING AT SOLVING NP-hard PROBLEMS

... and ANY GENERAL PURPOSE approach

should have the HIGHLY DESIRABLE feature

to behave efficiently for problems considered " EASY"!

Otherwise ... would you buy such a package?

An important observation

The MOMENT-SOS approach
is a GENERAL PURPOSE method

AIMING AT SOLVING NP-hard PROBLEMS

... and ANY GENERAL PURPOSE approach

should have the HIGHLY DESIRABLE feature

to behave efficiently for problems considered " EASY"!

Otherwise ... would you buy such a package?

Optimization problems

$$f^* = \min \{ f(x) \,:\, g_j(x) \geq 0, \quad j = 1, \ldots, m \}$$

where f and $-g_j$ are convex,

are considered EASY as they can be solved efficiently by appropriate methods (e.g. using logarithmic barrier method).

A polynomial $f \in \mathbb{R}[X]$ is SOS-CONVEX

if its Hessian $\nabla^2 f(x)$ factors as $L(x)\, L(x)^T$ for some matrix polynomial $L \in \mathbb{R}[X]^{n \times p}$ (for some p).

Optimization problems

$$f^* = \min \{ f(x) \,:\, g_j(x) \geq 0, \quad j = 1, \ldots, m \}$$

where f and $-g_j$ are convex,

are considered EASY as they can be solved efficiently by appropriate methods (e.g. using logarithmic barrier method).

A polynomial $f \in \mathbb{R}[X]$ is SOS-CONVEX

if its Hessian $\nabla^2 f(x)$ factors as $L(x)\, L(x)^T$ for some matrix polynomial $L \in \mathbb{R}[X]^{n \times p}$ (for some p).

Optimization problems

$$f^* = \min\{f(x) \,:\, g_j(x) \geq 0, \quad j = 1,\ldots,m\}$$

where f and $-g_j$ are convex,

are considered EASY as they can be solved efficiently by appropriate methods (e.g. using logarithmic barrier method).

A polynomial $f \in \mathbb{R}[X]$ is SOS-CONVEX

if its Hessian $\nabla^2 f(x)$ factors as $L(x)\, L(x)^T$ for some matrix polynomial $L \in \mathbb{R}[X]^{n\times p}$ (for some p).

A remarkable property of the MOM-SOS hierarchy: I

When solving the optimization problem

$$\mathbf{P}: \quad f^* = \min \{ f(\mathbf{x}) : \; g_j(\mathbf{x}) \geq 0, \; j = 1, \ldots, m \}$$

one does NOT distinguish between CONVEX, CONTINUOUS NON CONVEX, and 0/1 (and DISCRETE) problems! A boolean variable x_i is modelled via the equality constraint "$x_i^2 - x_i = 0$".

Only an algebraic description of the problem matters.
and then ☞ all problems are treated the same!

A remarkable property of the MOM-SOS hierarchy: I

When solving the optimization problem

$$\mathbf{P}: \quad f^* = \min\{f(\mathbf{x}) : \; g_j(\mathbf{x}) \geq 0, \; j = 1, \ldots, m\}$$

one does NOT distinguish between CONVEX, CONTINUOUS NON CONVEX, and 0/1 (and DISCRETE) problems! A boolean variable x_i is modelled via the equality constraint "$x_i^2 - x_i = 0$".

Only an algebraic description of the problem matters.
and then ☞ all problems are treated the same!

In Non Linear Programming (NLP),

modeling a 0/1 variable with the polynomial equality constraint

"$x_i^2 - x_i = 0$"

and applying a standard descent algorithm would be considered "stupid"!

Current practice:

☞ Each class of problems has its own *ad hoc* tailored algorithms.

In Non Linear Programming (NLP),

modeling a 0/1 variable with the polynomial equality constraint

$$“x_i^2 - x_i = 0"$$

and applying a standard descent algorithm would be considered “stupid"!

Current practice:

☞ Each class of problems has its own *ad hoc* tailored algorithms.

Even though the moment-SOS approach DOES NOT SPECIALIZE to each class of problems:

- It recognizes the class of (easy) SOS-convex problems as FINITE CONVERGENCE occurs at the FIRST relaxation in the hierarchy.
- Finite convergence also occurs for general convex problems and generically for non convex problems
- $\rightarrow$ (NOT true for the LP-hierarchy.)
- The SOS-hierarchy dominates other lift-and-project hierarchies (i.e. provides the best lower bounds) for hard 0/1 combinatorial optimization problems! The Computer Science community talks about a META-Algorithm.

Even though the moment-SOS approach DOES NOT SPECIALIZE to each class of problems:

- It recognizes the class of (easy) SOS-convex problems as FINITE CONVERGENCE occurs at the FIRST relaxation in the hierarchy.
- Finite convergence also occurs for general convex problems and generically for non convex problems
- $\rightarrow$ (NOT true for the LP-hierarchy.)
- The SOS-hierarchy dominates other lift-and-project hierarchies (i.e. provides the best lower bounds) for hard 0/1 combinatorial optimization problems! The Computer Science community talks about a META-Algorithm.

Even though the moment-SOS approach DOES NOT SPECIALIZE to each class of problems:

- It recognizes the class of (easy) SOS-convex problems as FINITE CONVERGENCE occurs at the FIRST relaxation in the hierarchy.
- Finite convergence also occurs for general convex problems and generically for non convex problems
- $\rightarrow$ (NOT true for the LP-hierarchy.)
- The SOS-hierarchy dominates other lift-and-project hierarchies (i.e. provides the best lower bounds) for hard 0/1 combinatorial optimization problems! The Computer Science community talks about a META-Algorithm.

Even though the moment-SOS approach DOES NOT SPECIALIZE to each class of problems:

- It recognizes the class of (easy) SOS-convex problems as FINITE CONVERGENCE occurs at the FIRST relaxation in the hierarchy.
- Finite convergence also occurs for general convex problems and generically for non convex problems
- $\rightarrow$ (NOT true for the LP-hierarchy.)
- The SOS-hierarchy dominates other lift-and-project hierarchies (i.e. provides the best lower bounds) for hard 0/1 combinatorial optimization problems! The Computer Science community talks about a META-Algorithm.

Even though the moment-SOS approach DOES NOT SPECIALIZE to each class of problems:

- It recognizes the class of (easy) SOS-convex problems as FINITE CONVERGENCE occurs at the FIRST relaxation in the hierarchy.
- Finite convergence also occurs for general convex problems and generically for non convex problems
- $\rightarrow$ (NOT true for the LP-hierarchy.)
- The SOS-hierarchy dominates other lift-and-project hierarchies (i.e. provides the best lower bounds) for hard 0/1 combinatorial optimization problems! The Computer Science community talks about a META-Algorithm.

A remarkable property: II

FINITE CONVERGENCE of the SOS-hierarchy is GENERIC!

... and provides a GLOBAL OPTIMALITY CERTIFICATE,

the analogue for the NON CONVEX CASE of the
KKT-OPTIMALITY conditions in the CONVEX CASE!

Thank You!