Semidefinite programming hierarchies for packing and energy minimization (4/4)

David de Laat (TU Delft)

Summer School on Optimization, Interpolation and Modular Forms,
August 24–28, 2020, EPFL

Topics for the four talks

1. Sums-of-squares hierarchies for polynomial optimization
2. Moment hierarchies for polynomial optimization
3. Packing problems
4. **Energy minimization problems**

Question 1

How can we use the Lasserre hierarchy for 0/1 polynomial optimization problems to derive a hierarchy for energy minimization on S^{n-1}?

Energy minimization on S^{n-1}

Find the minimum of

$$\sum_{1 \le i < j \le N} \frac{1}{\|p_i - p_j\|}$$

over all sets $\{p_1, \ldots, p_N\}$ of N distinct points on S^{n-1}

Energy minimization on S^{n-1}

Find the minimum of

$$\sum_{1 \leq i < j \leq N} \frac{1}{\|p_i - p_j\|}$$

over all sets $\{p_1, \ldots, p_N\}$ of N distinct points on S^{n-1}

For $n = 3$ this is the Thomson problem

Energy minimization on $\{1, \ldots, n\}$

Given a symmetric matrix $W \in \mathbb{R}^{n \times n}$, find the minimum of

$$\sum_{1 \leq i < j \leq N} W_{p_i, p_j}$$

over all sets $\{p_1, \ldots, p_N\}$ of N distinct points in $\{1, \ldots, n\}$

Energy minimization on $\{1, \ldots, n\}$

Given a symmetric matrix $W \in \mathbb{R}^{n \times n}$, find the minimum of

$$\sum_{1 \le i < j \le N} W_{p_i, p_j}$$

over all sets $\{p_1, \ldots, p_N\}$ of N distinct points in $\{1, \ldots, n\}$

As a 0/1 polynomial optimization problem:

$$\min \Big\{ \sum_{1 \le i < j \le N} W_{i,j} x_i x_j : x \in \{0, 1\}^n, \sum_{i=1}^{n} x_i = N \Big\}$$

Moment hierarchy for $0/1$ polynomial optimization

Moment hierarchy for $0/1$ polynomial optimization

$0/1$ polynomial optimization: $P = \inf\{p(x) : x \in \{0,1\}^n, \, x \in S(Q)\}$

Moment hierarchy for $0/1$ polynomial optimization

$0/1$ polynomial optimization: $P = \inf\{p(x) : x \in \{0,1\}^n,\, x \in S(Q)\}$

The moment hierarchy

$$P_t = \inf\Big\{ \sum_S p_S y_S : y_\emptyset = 1,\, M_t^q(y) \succeq 0 \text{ for } q \in \{1\} \cup Q \Big\}$$

Moment hierarchy for $0/1$ polynomial optimization

$0/1$ polynomial optimization: $P = \inf\{p(x) : x \in \{0,1\}^n,\, x \in S(Q)\}$

The moment hierarchy

$$P_t = \inf\Big\{ \sum_S p_S y_S : y_\emptyset = 1,\, M_t^q(y) \succeq 0 \text{ for } q \in \{1\} \cup Q \Big\}$$

Optimize over vectors y indexed by subsets of $\{1, \ldots, n\}$ of size at most $2t$.

Moment hierarchy for $0/1$ polynomial optimization

$0/1$ polynomial optimization: $P = \inf\{p(x) : x \in \{0,1\}^n,\, x \in S(Q)\}$

The moment hierarchy

$$P_t = \inf\Big\{ \sum_S p_S y_S : y_\emptyset = 1,\, M_t^q(y) \succeq 0 \text{ for } q \in \{1\} \cup Q \Big\}$$

Optimize over vectors y indexed by subsets of $\{1, \ldots, n\}$ of size at most $2t$.

Notation: $p(x) = \sum_{S \subseteq \{1,\ldots,n\}} p_S \prod_{i \in S} x_i$

Moment hierarchy for $0/1$ polynomial optimization

$0/1$ polynomial optimization: $P = \inf\{p(x) : x \in \{0,1\}^n,\, x \in S(Q)\}$

The moment hierarchy

$$P_t = \inf\Big\{\sum_S p_S y_S : y_\emptyset = 1,\, M_t^q(y) \succeq 0 \text{ for } q \in \{1\} \cup Q\Big\}$$

Optimize over vectors y indexed by subsets of $\{1,\dots,n\}$ of size at most $2t$.

Notation: $p(x) = \sum_{S \subseteq \{1,\dots,n\}} p_S \prod_{i \in S} x_i$

Localizing matrices:

$$M_t^q(y)_{J,J'} = \begin{cases} \sum_S q_S\, y_{J \cup J' \cup S} & \text{if } |J \cup J'| \leq 2t - \deg(q), \\ \text{unspecified} & \text{otherwise} \end{cases}$$

for $|J|, |J'| \leq t$

Energy minimization on $\{1, \ldots, n\}$

Formulation as a $0/1$ polynomial optimization problem:

$$\min \Big\{ \sum_{1 \leq i < j \leq N} W_{i,j} x_i x_j : x \in \{0,1\}^n, \sum_{i=1}^{n} x_i = N \Big\}$$

Energy minimization on $\{1, \ldots, n\}$

Formulation as a $0/1$ polynomial optimization problem:

$$\min \Big\{ \sum_{1 \le i < j \le N} W_{i,j} x_i x_j : x \in \{0,1\}^n, \sum_{i=1}^n x_i = N \Big\}$$

So the hierarchy becomes

$$P_t = \inf \Big\{ \sum_{1 \le i < j \le n} W_{i,j} y_{\{i,j\}} : y_\emptyset = 1,\ M_t^q(y) \succeq 0 \text{ for } q \in \{1\} \cup Q \Big\}$$

with $Q = \{q, -q\}$, $q = \sum_{i=1}^n x_i - N$

Energy minimization on $\{1, \ldots, n\}$

Formulation as a $0/1$ polynomial optimization problem:

$$\min \Big\{ \sum_{1 \leq i < j \leq N} W_{i,j} x_i x_j : x \in \{0,1\}^n, \sum_{i=1}^{n} x_i = N \Big\}$$

So the hierarchy becomes

$$P_t = \inf \Big\{ \sum_{1 \leq i < j \leq n} W_{i,j} y_{\{i,j\}} : y_\emptyset = 1,\ M_t^q(y) \succeq 0 \text{ for } q \in \{1\} \cup Q \Big\}$$

with $Q = \{q, -q\}$, $q = \sum_{i=1}^{n} x_i - N$

This reduces to

$$P_t = \inf \Big\{ \sum_{1 \leq i < j \leq N} W_{i,j} y_{\{i,j\}} : y_\emptyset = 1,\ M_t^1(y) \succeq 0,\ M_t^q(y) = 0 \Big\}$$

Localizing matrices:

$$M_t^q(y)_{J,J'} = \begin{cases} \sum_S q_S \, y_{J \cup J' \cup S} & \text{if } |J \cup J'| \leq 2t - \deg(q), \\ \text{unspecified} & \text{otherwise} \end{cases}$$

for $|J|, |J'| \leq t$

Localizing matrices:

$$M_t^q(y)_{J,J'} = \begin{cases} \sum_S q_S \, y_{J \cup J' \cup S} & \text{if } |J \cup J'| \leq 2t - \deg(q), \\ \text{unspecified} & \text{otherwise} \end{cases}$$

for $|J|, |J'| \leq t$

The constraint $M_t^q(y) = 0$ for $q = \sum_{i=1}^n x_i - N$ reduces to

$$N y_S = \sum_{i=1}^n y_{S \cup \{i\}}$$

for all subsets S of $\{1, \ldots, n\}$ of cardinality at most $2t - 1$

Reducing the number of constraints

Lemma Let $t \geq 1$ and $y \in \mathbb{R}^{I_{2t}}$. If

$$y_\emptyset = 1 \quad \text{and} \quad N y_S = \sum_{j=1}^{n} y_{S \cup \{j\}} \quad \text{for all} \quad S \in I_{2t-1},$$

then

$$\sum_{S \in I_{=i}} y_S = \binom{N}{i} \quad \text{for all} \quad 0 \leq i \leq 2t.$$

The hierarchy reduces to

$$\inf \Big\{ \sum_{1 \leq i < j \leq n} W_{i,j} y_{\{i,j\}} : y_\emptyset = 1,\, M_t^1(y) \succeq 0,\, \sum_{S \in I_{=i}} y_S = \binom{N}{i} \text{ for } 0 \leq i \leq 2t \Big\}$$

Finite convergence

This potentially makes the hierarchy weaker, but we can still prove finite convergence to the optimal energy in N steps

Finite convergence

This potentially makes the hierarchy weaker, but we can still prove finite convergence to the optimal energy in N steps

This hierarchy can be generalized from $\{1, \ldots, n\}$ to S^{n-1}, where it still converges in N steps

Question 2

How can we optimize over positive definite kernels?

Question 2

In the hierarchies for packing problems and for energy minimization we need to optimize over positive definite kernels $K \in \mathcal{C}(X \times X)$ for some space X

Question 2

In the hierarchies for packing problems and for energy minimization we need to optimize over positive definite kernels $K \in \mathcal{C}(X \times X)$ for some space X

A continuous kernel K is positive definite if

$$\sum_{x,y \in C} c_x c_y K(x,y) \geq 0$$

for all finite $C \subseteq X$ and $c \in \mathbb{R}^C$

Question 2

In the hierarchies for packing problems and for energy minimization we need to optimize over positive definite kernels $K \in \mathcal{C}(X \times X)$ for some space X

A continuous kernel K is positive definite if

$$\sum_{x,y \in C} c_x c_y K(x,y) \geq 0$$

for all finite $C \subseteq X$ and $c \in \mathbb{R}^C$

We may assume K is invariant under the symmetry group of the optimization problem

Schoenberg's theorem

If $K \in \mathcal{C}(S^{n-1} \times S^{n-1})$ is an $O(n)$-invariant positive definite kernel, then

$$K(x, y) = \sum_{k=0}^{\infty} c_k P_k^n(x \cdot y)$$

with $c_k \geq 0$, where convergence is uniform absolute.

Schoenberg's theorem

If $K \in \mathcal{C}(S^{n-1} \times S^{n-1})$ is an $O(n)$-invariant positive definite kernel, then

$$K(x, y) = \sum_{k=0}^{\infty} c_k P_k^n(x \cdot y)$$

with $c_k \geq 0$, where convergence is uniform absolute.

What if we wouldn't know about Schoenberg's theorem?

Bochner's theorem

The action of $O(n)$ on S^{n-1} defines an action on $\mathcal{C}(S^{n-1})$:

$$Af(x) = f(A^{-1}x)$$

Bochner's theorem

The action of $O(n)$ on S^{n-1} defines an action on $\mathcal{C}(S^{n-1})$:

$$Af(x) = f(A^{-1}x)$$

A space $S \subseteq \mathcal{C}(S^{n-1})$ is invariant if $Af \in S$ for all $A \in O(n)$ and $f \in S$

Bochner's theorem

The action of $O(n)$ on S^{n-1} defines an action on $\mathcal{C}(S^{n-1})$:

$$Af(x) = f(A^{-1}x)$$

A space $S \subseteq \mathcal{C}(S^{n-1})$ is invariant if $Af \in S$ for all $A \in O(n)$ and $f \in S$

An invariant space $S \subseteq \mathcal{C}(S^{n-1})$ is irreducible it cannot be written as a direct sum of nontrival invariant subspaces

Bochner's theorem

The action of $O(n)$ on S^{n-1} defines an action on $\mathcal{C}(S^{n-1})$:

$$Af(x) = f(A^{-1}x)$$

A space $S \subseteq \mathcal{C}(S^{n-1})$ is invariant if $Af \in S$ for all $A \in O(n)$ and $f \in S$

An invariant space $S \subseteq \mathcal{C}(S^{n-1})$ is irreducible it cannot be written as a direct sum of nontrival invariant subspaces

We can decompose $\mathcal{C}(S^{n-1})$ as a direct sum of irreducible subspaces H_k

Bochner's theorem

The action of $O(n)$ on S^{n-1} defines an action on $\mathcal{C}(S^{n-1})$:

$$Af(x) = f(A^{-1}x)$$

A space $S \subseteq \mathcal{C}(S^{n-1})$ is invariant if $Af \in S$ for all $A \in O(n)$ and $f \in S$

An invariant space $S \subseteq \mathcal{C}(S^{n-1})$ is irreducible it cannot be written as a direct sum of nontrival invariant subspaces

We can decompose $\mathcal{C}(S^{n-1})$ as a direct sum of irreducible subspaces H_k

Here $H_k = \mathrm{span}\{Y_k^j : j = 1, \ldots, d_k\}$ is the space of spherical harmonics of degree k

Bochner's theorem

The action of $O(n)$ on S^{n-1} defines an action on $\mathcal{C}(S^{n-1})$:

$$Af(x) = f(A^{-1}x)$$

A space $S \subseteq \mathcal{C}(S^{n-1})$ is invariant if $Af \in S$ for all $A \in O(n)$ and $f \in S$

An invariant space $S \subseteq \mathcal{C}(S^{n-1})$ is irreducible it cannot be written as a direct sum of nontrival invariant subspaces

We can decompose $\mathcal{C}(S^{n-1})$ as a direct sum of irreducible subspaces H_k

Here $H_k = \mathrm{span}\{Y_k^j : j = 1, \ldots, d_k\}$ is the space of spherical harmonics of degree k

Bochner's theorem says: $K(x, y) = \sum_k c_k \sum_{j=1}^{d_k} Y_k^j(x) Y_k^j(y)$

Bochner's theorem

The action of $O(n)$ on S^{n-1} defines an action on $\mathcal{C}(S^{n-1})$:

$$Af(x) = f(A^{-1}x)$$

A space $S \subseteq \mathcal{C}(S^{n-1})$ is invariant if $Af \in S$ for all $A \in O(n)$ and $f \in S$

An invariant space $S \subseteq \mathcal{C}(S^{n-1})$ is irreducible it cannot be written as a direct sum of nontrival invariant subspaces

We can decompose $\mathcal{C}(S^{n-1})$ as a direct sum of irreducible subspaces H_k

Here $H_k = \mathrm{span}\{Y_k^j : j = 1, \ldots, d_k\}$ is the space of spherical harmonics of degree k

Bochner's theorem says: $K(x, y) = \sum_k c_k \sum_{j=1}^{d_k} Y_k^j(x) Y_k^j(y)$

The addition formula: $\sum_{j=1}^{d_k} Y_k^j(x) Y_k^j(y) = P_k^n(x \cdot y)$

Bochner's theorem

In general we can have irreducible subspaces that are equivalent.

Bochner's theorem

In general we can have irreducible subspaces that are equivalent.

Then we write

$$\mathcal{C}(X) = \bigoplus_{\pi} \bigoplus_{i=1}^{m_\pi} H_{\pi,i}$$

where

$$H_{\pi,i} = \mathrm{span}\{e_{\pi,i,j} : j = 1, \ldots, m_\pi\}$$

Bochner's theorem

In general we can have irreducible subspaces that are equivalent.

Then we write

$$\mathcal{C}(X) = \bigoplus_{\pi} \bigoplus_{i=1}^{m_\pi} H_{\pi,i}$$

where

$$H_{\pi,i} = \mathrm{span}\{e_{\pi,i,j} : j = 1, \ldots, m_\pi\}$$

Bochner's theorem says

$$K(x,y) = \sum_{\pi} \langle C_\pi, Z_\pi(x,y) \rangle,$$

where

$$Z_\pi(x,y)_{i,i'} = \sum_{j=1}^{d_\pi} e_{\pi,i,j}(x) e_{\pi,i',j}(y)$$

Bochner's theorem

In general we can have irreducible subspaces that are equivalent.

Then we write

$$\mathcal{C}(X) = \bigoplus_{\pi} \bigoplus_{i=1}^{m_\pi} H_{\pi,i}$$

where

$$H_{\pi,i} = \operatorname{span}\{e_{\pi,i,j} : j = 1, \ldots, m_\pi\}$$

Bochner's theorem says

$$K(x,y) = \sum_{\pi} \langle C_\pi, Z_\pi(x,y) \rangle,$$

where

$$Z_\pi(x,y)_{i,i'} = \sum_{j=1}^{d_\pi} e_{\pi,i,j}(x) e_{\pi,i',j}(y)$$

Here C_π is a positive semidefinite matrix

Question 3

How do we deal with inequality constraints?

Sums-of-squares

For a polynomial optimization problem

$$P = \inf \left\{ p(x) : x \in S(Q) \right\}, \quad S(Q) = \left\{ x \in \mathbb{R}^n : q(x) \geq 0 \text{ for } q \in Q \right\}$$

we first need to reformulate as

$$P = \sup \left\{ c : p(x) - c \geq 0 \text{ for } x \in S(Q) \right\}$$

to use sums-of-squares

Sums-of-squares

The Delsarte bound:

$$\inf \Big\{ 1 + f(1) : f(u) = \sum_{k=0}^{d} c_k P_k^n(u), \\ c_0, c_1, \ldots, c_d \geq 0, \\ f(u) \ \leq \ -1 \text{ for } u \in [\cos\theta, 1] \Big\}.$$

Sums-of-squares

The Delsarte bound:

$$\inf \Big\{ 1 + f(1) : f(u) = \sum_{k=0}^{d} c_k P_k^n(u), \\ c_0, c_1, \ldots, c_d \geq 0, \\ f(u) \ \leq \ -1 \text{ for } u \in [\cos\theta, 1] \Big\}.$$

Sums-of-squares

The Delsarte bound:

$$\inf \left\{ 1 + f(1) : f(u) = \sum_{k=0}^{d} c_k P_k^n(u), \right.$$
$$c_0, c_1, \ldots, c_d \geq 0,$$
$$\left. 1 - f(u) \geq 0 \text{ for } u \in [\cos\theta, 1] \right\}.$$

Sums-of-squares

The Delsarte bound:

$$\inf \Big\{ 1 + f(1) : f(u) = \sum_{k=0}^{d} c_k P_k^n(u), \\ c_0, c_1, \ldots, c_d \geq 0, \\ 1 - f(u) \geq 0 \text{ for } u \in [\cos\theta, 1] \Big\}.$$

The set $[\cos\theta, 1]$ is semialgebraic, e.g.,

$$[\cos\theta, 1] = \Big\{ x \in \mathbb{R} : x - \cos\theta \geq 0,\ 1 - x \geq 0 \Big\}, \text{ or}$$

$$[\cos\theta, 1] = \Big\{ x \in \mathbb{R} : (x - \cos\theta)(1 - x) \geq 0 \Big\}$$

Sums-of-squares

Here we do not need Putinar's theorem, an older result by Lukács says

$$p(x) \geq 0 \text{ for } x \in [\cos\theta, 1]$$

implies

$$p \in \mathcal{M}_{(\deg(p)-1)/2}(\{x - \cos\theta,\, 1 - x\})$$

if $\deg(p)$ is odd, and

$$p \in \mathcal{M}_{\deg(p)/2-1}(\{(x - \cos\theta)(1 - x)\})$$

if $\deg(p)$ is even

Semidefinite programs with semialgebraic constraints

More generally we can consider semidefinite programs where we have polynomials

- whose coefficients depend linearly on the entries in the positive semidefinite matrix variables, and
- which are nonnegative on semialgebraic sets

Semidefinite programs with semialgebraic constraints

More generally we can consider semidefinite programs where we have polynomials

- whose coefficients depend linearly on the entries in the positive semidefinite matrix variables, and
- which are nonnegative on semialgebraic sets

We can use Putinar's theorem to model these inequality constraints using positive semidefinite matrices

Semidefinite programs with semialgebraic constraints

More generally we can consider semidefinite programs where we have polynomials

- whose coefficients depend linearly on the entries in the positive semidefinite matrix variables, and
- which are nonnegative on semialgebraic sets

We can use Putinar's theorem to model these inequality constraints using positive semidefinite matrices

If these polynomials have symmetries we can use symmetric sums-of-squares which is more efficient
(see the paper by Gatermann and Parrilo)

Computations for energy minimization

Computations for energy minimization

- Setup moment hierarchy by first considering energy minimization on $\{1, \ldots, n\}$ and applying the Lasserre moment hierarchy

Computations for energy minimization

- Setup moment hierarchy by first considering energy minimization on $\{1, \ldots, n\}$ and applying the Lasserre moment hierarchy
- Consider the dual hierarchy where we optimize over continuous, positive definite kernels

Computations for energy minimization

- Setup moment hierarchy by first considering energy minimization on $\{1, \ldots, n\}$ and applying the Lasserre moment hierarchy
- Consider the dual hierarchy where we optimize over continuous, positive definite kernels
- The first step is Yudin's bound, so consider the second step

Computations for energy minimization

- Setup moment hierarchy by first considering energy minimization on $\{1, \ldots, n\}$ and applying the Lasserre moment hierarchy
- Consider the dual hierarchy where we optimize over continuous, positive definite kernels
- The first step is Yudin's bound, so consider the second step
- Use harmonic analysis to efficiently parameterize the space of positive definite kernels by positive semidefinite matrices

Computations for energy minimization

- Setup moment hierarchy by first considering energy minimization on $\{1, \ldots, n\}$ and applying the Lasserre moment hierarchy
- Consider the dual hierarchy where we optimize over continuous, positive definite kernels
- The first step is Yudin's bound, so consider the second step
- Use harmonic analysis to efficiently parameterize the space of positive definite kernels by positive semidefinite matrices
- Use symmetric sums-of-squares characterizations to deal with the polynomial inequality constraints

Computations for energy minimization

- Setup moment hierarchy by first considering energy minimization on $\{1, \ldots, n\}$ and applying the Lasserre moment hierarchy
- Consider the dual hierarchy where we optimize over continuous, positive definite kernels
- The first step is Yudin's bound, so consider the second step
- Use harmonic analysis to efficiently parameterize the space of positive definite kernels by positive semidefinite matrices
- Use symmetric sums-of-squares characterizations to deal with the polynomial inequality constraints
- Solve the resulting semidefinite program on a computer (sdpa-qd or sdpa-gmp because we need high precision)

Computations for energy minimization

- Setup moment hierarchy by first considering energy minimization on $\{1, \ldots, n\}$ and applying the Lasserre moment hierarchy
- Consider the dual hierarchy where we optimize over continuous, positive definite kernels
- The first step is Yudin's bound, so consider the second step
- Use harmonic analysis to efficiently parameterize the space of positive definite kernels by positive semidefinite matrices
- Use symmetric sums-of-squares characterizations to deal with the polynomial inequality constraints
- Solve the resulting semidefinite program on a computer (sdpa-qd or sdpa-gmp because we need high precision)
- The Thomson problem for 5 particles on S^2 was solved by Schwartz in 2015, but not using a certificate. Yudin's bound and the three-point bound by Cohn and Woo are not sharp here.

Computations for energy minimization

- Setup moment hierarchy by first considering energy minimization on $\{1, \ldots, n\}$ and applying the Lasserre moment hierarchy
- Consider the dual hierarchy where we optimize over continuous, positive definite kernels
- The first step is Yudin's bound, so consider the second step
- Use harmonic analysis to efficiently parameterize the space of positive definite kernels by positive semidefinite matrices
- Use symmetric sums-of-squares characterizations to deal with the polynomial inequality constraints
- Solve the resulting semidefinite program on a computer (sdpa-qd or sdpa-gmp because we need high precision)
- The Thomson problem for 5 particles on S^2 was solved by Schwartz in 2015, but not using a certificate. Yudin's bound and the three-point bound by Cohn and Woo are not sharp here.
- The second step of the hierarchy is numerically sharp (it gets at least 28 decimal digits of the ground state energy correct)

Advertisement

PhD position in Delft: www.daviddelaat.nl/vacancy.html
Application deadline: 1 October 2020

Papers

Jean B. Lasserre, **Global Optimization with Polynomials and the Problem of Moments**, SIAM J. Optim., 11(3), 796–817, 2001

Monique Laurent, **A Comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre Relaxations for 0-1 Programming**, Math. Oper. Res. Vol. 28, No. 3 (Aug., 2003), pp. 470-496

Karin Gatermann, Pablo A. Parrilo, **Symmetry groups, semidefinite programs, and sums of squares**, Journal of Pure and Appl. Algebra, Vol. 192, No. 1-3, pp. 95-128, 2004

David de Laat, Frank Vallentin, **A semidefinite programming hierarchy for packing problems in discrete geometry**, Math. Program., Ser. B 151 (2015)

David de Laat, **Moment methods in energy minimization: New bounds for Riesz minimal energy problems** Trans. Amer. Math. Soc. (2019)