Semidefinite programming lifts and sparse sums-of-squares

Hamza Fawzi (MIT, LIDS)
Joint work with James Saunderson (UW) and Pablo Parrilo (MIT)

Cornell ORIE, October 2015

Linear optimization

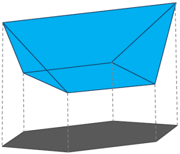
- Central question in optimization is to optimize a linear function ℓ on a convex set C:

$$\min_{x \in C} \ell(x).$$

- Need "good" description of C to solve optimization problem efficiently.

- Interested in linear programming and semidefinite programming descriptions.

Lifts in optimization



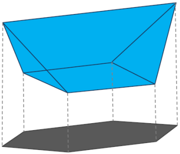
Lifts (a.k.a. extended formulations):

introducing additional variables can simplify an optimization problem dramatically

Lifts in optimization

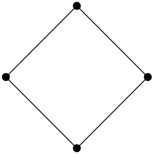
Lifts (a.k.a. extended formulations):

introducing additional variables can simplify an optimization problem dramatically



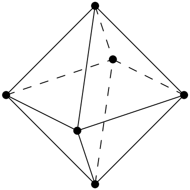
Example ℓ_1 ball

$$P = \{x \in \mathbb{R}^n : \|x\|_1 \leq 1\}$$
$$= \{x \in \mathbb{R}^n : a^T x \leq 1 \ \forall a \in \{-1, 1\}^n\}$$

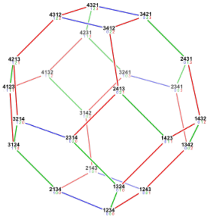


P has 2^n facets, yet we have an efficient description using $2n$ linear inequalities:

$$P = \Big\{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^n \text{ s.t.}$$
$$-y_i \leq x_i \leq y_i \text{ and } \sum_{i=1}^{n} y_i = 1\Big\}.$$



Lifts in optimization: Permutahedron



Source: Wikipedia "Permutahedron"

Permutahedron

$$P = \text{conv}\Big\{(\sigma(1), \dots, \sigma(n)) : \sigma \in \mathfrak{S}_n\Big\} \subseteq \mathbb{R}^n.$$

P has $n!$ vertices and $\Omega(2^n)$ facets.

A simple formulation using n^2 inequalities

$$DS_n = \text{conv(permutation matrices)} = \text{doubly-stochastic matrices.}$$

Then

$$P = \pi(DS_n)$$

where

$$\begin{array}{rcl} \pi : \mathbb{R}^{n\times n} & \to & \mathbb{R}^n \\ M & \mapsto & Mu. \end{array}$$

and $u = (1, 2, \dots, n)$.

Formal definition of lift

Formal definition of LP lift

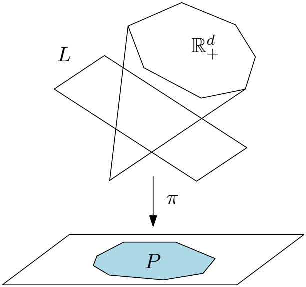
Polytope P has a LP lift of size d if

$$P = \pi(\mathbb{R}^d_+ \cap L)$$

where

- $\pi : \mathbb{R}^d \to \mathbb{R}^n$ linear map
- L affine subspace of $\mathbb{R}^d$

LP extension complexity of P is smallest d such that P has a LP lift of size d.



Formal definition of lift

Formal definition of LP lift

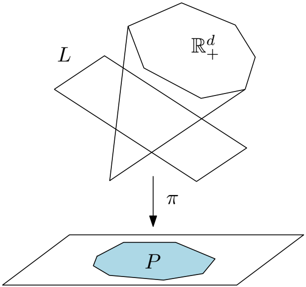
Polytope P has a LP lift of size d if

$$P = \pi(\mathbb{R}^d_+ \cap L)$$

where

- $\pi : \mathbb{R}^d \to \mathbb{R}^n$ linear map
- L affine subspace of $\mathbb{R}^d$

LP extension complexity of P is smallest d such that P has a LP lift of size d.



Examples of LP lifts

- Regular N-gon in $\mathbb{R}^2$ has LP lift of size $O(\log N)$ (Ben-Tal and Nemirovski 2001).
- Permutahedron has LP lift of size $O(n \log n)$ (Goemans 2009).

Semidefinite programming lifts

Semidefinite programming

$$\min \ \mathcal{L}(Y) \quad \text{subject to} \quad Y \in \mathbf{S}_+^d, \quad Y \in L$$

where $\mathbf{S}_+^d =$ cone of $d \times d$ positive semidefinite matrices, $\mathcal{L}$ is a linear function and L affine subspace of $\mathbf{S}^d$. SDP forms a superset of LP.

Semidefinite programming lifts

Semidefinite programming

$$\min \; \mathcal{L}(Y) \quad \text{subject to} \quad Y \in \mathbf{S}^d_+, \quad Y \in L$$

where $\mathbf{S}^d_+ =$ cone of $d \times d$ positive semidefinite matrices, $\mathcal{L}$ is a linear function and L affine subspace of $\mathbf{S}^d$. SDP forms a superset of LP.

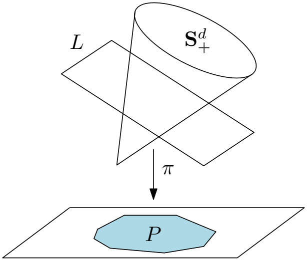
Positive semidefinite lifts

P has SDP lift of size d if

$$P = \pi(\mathbf{S}^d_+ \cap L)$$

where

- π linear map
- L affine subspace of $\mathbf{S}^d$



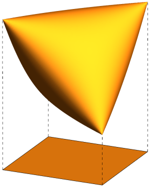
SDP extension complexity of P is smallest d such that P has a SDP lift of size d.

LP lifts vs. SDP lifts

Example The square $P = [-1, 1]^2$:

- **SDP lifts**: P has an SDP lift of size 3:

$$[-1,1]^2 = \left\{ (x_1, x_2) \in \mathbb{R}^2 : \exists u \in \mathbb{R} \begin{bmatrix} 1 & x_1 & x_2 \\ x_1 & 1 & u \\ x_2 & u & 1 \end{bmatrix} \succeq 0 \right\}$$



SDP extension complexity of $[-1, 1]^2$ is 3.

- **LP lifts**: Can show that LP extension complexity of $[-1, 1]^2$ is 4.

LP lifts vs. SDP lifts

Question: How powerful are SDP lifts compared to LP lifts?

Theorem (Fawzi-Saunderson-Parrilo 2015)

There is a family of polytopes $P_d \subset \mathbb{R}^{2d}$ such that

$$\frac{\textit{LP extension complexity of } P_d}{\textit{SDP extension complexity of } P_d} \geq \Omega\left(\frac{d}{\log d}\right) \to +\infty.$$

- Only example known so far of gap between SDP and LP extension complexity.
- Polytopes P_d are highly symmetric and well-studied (trigonometric cyclic polytopes).
- Proof idea relies on finding *sparse* sum-of-squares certificates for facet inequalities.

Constructing lifts using sum-of-squares

- $P \subset \mathbb{R}^n$ polytope, $X =$ extreme points of P
- Facet of P is an affine function ℓ such that

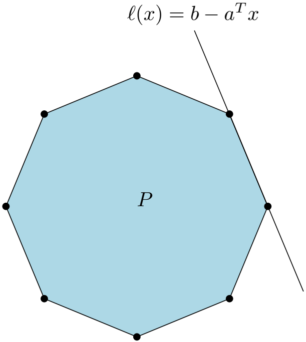
$$\ell(x) \geq 0 \quad \forall x \in X.$$

- Sum-of-squares certificate for $\ell(x)$:

$$\ell(x) = \sum_j h_j(x)^2$$

  for some functions $h_i : X \to \mathbb{R}$

- **SDP lifts** $\leftrightarrow$ **sum-of-squares**: If we can find "small" sum-of-squares certificates for each facet ℓ of P then we get a "small" SDP lift.



Constructing lifts using sum-of-squares

- $P \subset \mathbb{R}^n$ polytope, X = extreme points of P.
- $\mathcal{F}(X, \mathbb{R})$ = space of real-valued functions on X.

Theorem (Lasserre 2010, Gouveia et al. 2011)

Assume there is a subspace V of $\mathcal{F}(X, \mathbb{R})$ such that any facet-defining inequality $\ell(x) \geq 0$ for P can be certified using sum-of-squares in V, i.e., there exist $h_1, \ldots, h_J \in V$:

$$\ell(x) = \sum_{j=1}^{J} h_j(x)^2 \quad \forall x \in X.$$

Then $\operatorname{conv}(X)$ *has an (explicit) SDP lift of size* $\dim V$.

Constructing lifts using sum-of-squares

- $P \subset \mathbb{R}^n$ polytope, X = extreme points of P.
- $\mathcal{F}(X, \mathbb{R})$ = space of real-valued functions on X.

Theorem (Lasserre 2010, Gouveia et al. 2011)

Assume there is a subspace V of $\mathcal{F}(X, \mathbb{R})$ such that any facet-defining inequality $\ell(x) \geq 0$ for P can be certified using sum-of-squares in V, i.e., there exist $h_1, \ldots, h_J \in V$:

$$\ell(x) = \sum_{j=1}^{J} h_j(x)^2 \quad \forall x \in X.$$

Then $\operatorname{conv}(X)$ has an (explicit) SDP lift of size $\dim V$.

$\rightarrow$ This theorem reduces the problem of constructing SDP lifts to studying sum-of-squares certificates of facets of P.

Lasserre SDP lifts

- Lasserre SDP lift works by *degree*: certificates of facets ℓ of the form

$$\ell(x) = \sum_j h_j(x)^2$$

where $\deg h_j \leq k$. This often results in large SDP lifts (e.g., regular polygons).

Lasserre SDP lifts

- Lasserre SDP lift works by *degree*: certificates of facets ℓ of the form

$$\ell(x) = \sum_j h_j(x)^2$$

  where $\deg h_j \leq k$. This often results in large SDP lifts (e.g., regular polygons).

- Idea to construct smaller lifts: Look instead for *sparse* sum-of-squares certificates, i.e.,

$$\ell(x) = \sum_j h_j(x)^2$$

  where h_j are *sparse* polynomials.

Lasserre SDP lifts

- Lasserre SDP lift works by *degree*: certificates of facets ℓ of the form

$$\ell(x) = \sum_j h_j(x)^2$$

  where $\deg h_j \leq k$. This often results in large SDP lifts (e.g., regular polygons).

- Idea to construct smaller lifts: Look instead for *sparse* sum-of-squares certificates, i.e.,

$$\ell(x) = \sum_j h_j(x)^2$$

  where h_j are *sparse* polynomials.

Key question: Given a nonnegative function that is sparse, can we write it as a sum-of-squares of sparse functions?

Sparse sum-of-squares on finite abelian group

Setting: Functions on a finite abelian group $G \to$ Natural basis to measure sparsity of functions, namely *Fourier basis* of G.

- $G = \mathbb{Z}_N \to$ usual Fourier basis (complex exponentials)
- $G = \{-1, 1\}^n \to$ Fourier analysis on the hypercube (square-free monomials).

Sparse sum-of-squares on finite abelian group

Setting: Functions on a finite abelian group $G \to$ Natural basis to measure sparsity of functions, namely *Fourier basis* of G.

- $G = \mathbb{Z}_N \to$ usual Fourier basis (complex exponentials)
- $G = \{-1, 1\}^n \to$ Fourier analysis on the hypercube (square-free monomials).

Main result (informal) Given G finite abelian group and $\mathcal{S} \subseteq \hat{G}$, we give a method to construct $\mathcal{T} \subseteq \hat{G}$ such that the following holds:

> *any nonnegative function $f : G \to \mathbb{R}_+$ supported on $\mathcal{S}$ has a sum-of-squares certificate supported on $\mathcal{T}$, i.e., $f(x) = \sum_j |h_j(x)|^2$ where support$(h_j) \subseteq \mathcal{T}$.*

Method involves constructing “nice” chordal covers of the Cayley graph Cay$(\hat{G}, \mathcal{S})$.

Sparse sum-of-squares on finite abelian group

Setting: Functions on a finite abelian group $G \to$ Natural basis to measure sparsity of functions, namely *Fourier basis* of G.

- $G = \mathbb{Z}_N \to$ usual Fourier basis (complex exponentials)
- $G = \{-1, 1\}^n \to$ Fourier analysis on the hypercube (square-free monomials).

Main result (informal) Given G finite abelian group and $\mathcal{S} \subseteq \hat{G}$, we give a method to construct $\mathcal{T} \subseteq \hat{G}$ such that the following holds:

> *any nonnegative function $f : G \to \mathbb{R}_+$ supported on $\mathcal{S}$ has a sum-of-squares certificate supported on $\mathcal{T}$, i.e., $f(x) = \sum_j |h_j(x)|^2$ where support$(h_j) \subseteq \mathcal{T}$.*

Method involves constructing “nice” chordal covers of the Cayley graph Cay$(\hat{G}, \mathcal{S})$.

Consequence for lifts: Allows us to construct SDP lifts of moment polytope $\mathcal{M}(G, \mathcal{S})$ of size $|\mathcal{T}|$.

Application 1: Degree d polynomials on $\mathbb{Z}_N$

$$TC_{N,2d} = \operatorname{conv}\Big\{\left(\cos\left(\tfrac{2\pi x}{N}\right), \sin\left(\tfrac{2\pi x}{N}\right), \ldots, \cos\left(\tfrac{2\pi dx}{N}\right), \sin\left(\tfrac{2\pi dx}{N}\right)\right) : x \in \{0, 1, 2 \ldots, N-1\}\Big\} \subset \mathbb{R}^{2d}$$

Trigonometric cyclic polytope (corresponds to $G = \mathbb{Z}_N$ and $\mathcal{S} = \{-d, \ldots, d\}$).

Application 1: Degree d polynomials on $\mathbb{Z}_N$

$$TC_{N,2d} = \text{conv}\Big\{\left(\cos\left(\tfrac{2\pi x}{N}\right), \sin\left(\tfrac{2\pi x}{N}\right), \ldots, \cos\left(\tfrac{2\pi dx}{N}\right), \sin\left(\tfrac{2\pi dx}{N}\right)\right) : \\ x \in \{0, 1, 2 \ldots, N-1\}\Big\} \subset \mathbb{R}^{2d}$$

Trigonometric cyclic polytope (corresponds to $G = \mathbb{Z}_N$ and $\mathcal{S} = \{-d, \ldots, d\}$).

Using main result (with good choice of chordal cover):

- If d divides N then $TC_{N,2d}$ has a PSD lift of size $\leq 3d\log_2(N/d)$.
- Case $N = d^2$ gives gap between SDP and LP lifts
  - SDP lift of size $O(d\log(d))$
  - LP lift must have size $\geq \Omega(d^2)$
    (lower bound due to Fiorini et al. for d-neighborly polytopes)

Consequence 2: Quadratic polynomials on $\{-1,1\}^n$

Conjecture (Laurent 2003): If

$$f(x) = a_0 + \sum_{i<j} a_{ij} x_i x_j \quad \text{non-negative } \forall x \in \{-1,1\}^n$$

then f is a sum of squares of polynomials of degree at most $\lceil n/2 \rceil$.

- Laurent (2003): degree at least $\lceil n/2 \rceil$ necessary
- Blekherman, Gouveia, Pfeiffer (2014): true if allow multipliers

Consequence 2: Quadratic polynomials on $\{-1, 1\}^n$

Conjecture (Laurent 2003): If

$$f(x) = a_0 + \sum_{i<j} a_{ij} x_i x_j \quad \text{non-negative } \forall x \in \{-1, 1\}^n$$

then f is a sum of squares of polynomials of degree at most $\lceil n/2 \rceil$.

- Laurent (2003): degree at least $\lceil n/2 \rceil$ necessary
- Blekherman, Gouveia, Pfeiffer (2014): true if allow multipliers

In our language:

- Group: $G = \{-1, 1\}^n \cong \mathbb{Z}_2^n$
- Characters: $\chi_S(x) = \prod_{i \in S} x_i$ (square-free polynomials)
- non-negative functions with support $\mathcal{S} = \{S : |S| \in \{0, 2\}\}$

Good choices in main result $\rightarrow$ prove Laurent's conjecture

Conclusion

Summary

- Used sparse sums-of-squares to construct lifts that are smaller than the those from the Lasserre construction.
- Allowed us to give the first example of a polytope with a gap between SDP lifts and LP lifts.
- Allowed us to prove conjecture of Laurent.

Questions:

- All the lifts we produce respect the symmetry of the polytope P (they are *equivariant*). Does breaking symmetry help in reducing the size of lifts? (for LP lifts it does).
- Lower bounds?

For more information: preprint `arXiv:1503.01207`

Conclusion

Summary

- Used sparse sums-of-squares to construct lifts that are smaller than the those from the Lasserre construction.
- Allowed us to give the first example of a polytope with a gap between SDP lifts and LP lifts.
- Allowed us to prove conjecture of Laurent.

Questions:

- All the lifts we produce respect the symmetry of the polytope P (they are *equivariant*). Does breaking symmetry help in reducing the size of lifts? (for LP lifts it does).
- Lower bounds?

For more information: preprint `arXiv:1503.01207`

Thank you!