Positive polynomials and convergence of LP and SDP relaxations

Markus Schweighofer

Universität Konstanz

ETH Zürich, May 31, 2005

• X_1, \ldots, X_n variables

- X_1, \ldots, X_n variables
- $X := X_1$ when n = 1

- X_1, \ldots, X_n variables
- $X := X_1$ when $n = 1, (X, Y) := (X_1, X_2)$ when n = 2, ...

- X_1, \ldots, X_n variables
- $X := X_1$ when $n = 1, (X, Y) := (X_1, X_2)$ when n = 2, ...
- $\mathbb{R}[X_1, \ldots, X_n]$ polynomial ring

- X_1, \ldots, X_n variables
- $X := X_1$ when $n = 1, (X, Y) := (X_1, X_2)$ when n = 2, ...
- $\mathbb{R}[X_1, \ldots, X_n]$ polynomial ring
- $f \in \mathbb{R}[X_1, \dots, X_n]$ an arbitrary polynomial

- X_1, \ldots, X_n variables
- $X := X_1$ when $n = 1, (X, Y) := (X_1, X_2)$ when n = 2, ...
- $\mathbb{R}[X_1, \ldots, X_n]$ polynomial ring
- $f \in \mathbb{R}[X_1, \dots, X_n]$ an arbitrary polynomial
- $g_1, \ldots, g_m \in \mathbb{R}[X_1, \ldots, X_n]$ polynomials defining...

- X_1, \ldots, X_n variables
- $X := X_1$ when $n = 1, (X, Y) := (X_1, X_2)$ when n = 2, ...
- $\mathbb{R}[X_1, \ldots, X_n]$ polynomial ring
- $f \in \mathbb{R}[X_1, \dots, X_n]$ an arbitrary polynomial
- $g_1, \ldots, g_m \in \mathbb{R}[X_1, \ldots, X_n]$ polynomials defining...
- ... the set $S := \{x \in \mathbb{R}^n \mid g_1(x) \ge 0, \dots, g_m(x) \ge 0\}$

f g_1,\ldots,g_m

n

Optimization

We consider the problem of minimizing f on S.

Optimization

We consider the problem of minimizing f on S. So we want to compute numerically the infimum

$$f^* := \inf\{f(x) \mid x \in S\} \in \mathbb{R} \cup \{\pm \infty\}$$

Optimization

We consider the problem of minimizing f on S. So we want to compute numerically the infimum

$$f^* := \inf\{f(x) \mid x \in S\} \in \mathbb{R} \cup \{\pm \infty\}$$

and, if possible, a minimizer, i.e., an element of the set

 $S^* := \{x^* \in S \mid f(x^*) \le f(x) \text{ for all } x \in S\}.$

L P

Linear Programming

minimize f(x)

subject to $x \in \mathbb{R}^n$ $g_1(x) \ge 0$ \vdots $g_m(x) \ge 0$

where all polynomials f and g_i are linear, i.e., their degree is ≤ 1 . In particular, $S \subseteq \mathbb{R}^n$ is a polyhedron.

Linear Programming

minimize f(x)

subject to $x \in \mathbb{R}^n$ $\begin{pmatrix} g_1(x) & & \\ & \ddots & \\ & & g_m(x) \end{pmatrix} \text{ is psd}$

where all polynomials f and g_i are linear, i.e., their degree is ≤ 1 . In particular, $S \subseteq \mathbb{R}^n$ is a polyhedron.

S D P

minimize f(x)

subject to
$$x \in \mathbb{R}^n$$

$$\begin{pmatrix} g_{11}(x) & \dots & g_{1m}(x) \\ \vdots & \ddots & \vdots \\ & & & & \\ & & & \\ & &$$

where all polynomials f and g_{ij} are linear, i.e., their degree is ≤ 1 .

Semidefinite Programming

minimize f(x)

subject to $x \in \mathbb{R}^n$ $\begin{pmatrix} g_{11}(x) & \dots & g_{1m}(x) \\ \vdots & \ddots & \vdots \\ & & \dots & g_{mm}(x) \end{pmatrix} \text{ is psd}$

where all polynomials f and g_{ij} are linear, i.e., their degree is ≤ 1 .

Positive semidefinite matrices and families of vectors

Proposition. A real symmetric $k \times k$ matrix is psd if and only if there are vectors $v_1, \ldots, v_k \in \mathbb{R}^k$ such that

$$M = \begin{pmatrix} \langle v_1, v_1 \rangle & \dots & \langle v_1, v_k \rangle \\ \vdots & & \vdots \\ \langle v_k, v_1 \rangle & \dots & \langle v_k, v_k \rangle \end{pmatrix}$$

• Every linear program (P) has an optimal value P^* .

- Every linear program (P) has an optimal value P^* .
- To every linear program (P), one can define a dual program (D) which is again a linear program.

- Every linear program (P) has an optimal value P^* .
- To every linear program (P), one can define a dual program (D) which is again a linear program.
- If (P) is a minimization problem, then (D) is a maximization problem and weak duality holds:

 $D^* \le P^*$

- Every linear program (P) has an optimal value P^* .
- To every linear program (P), one can define a dual program (D) which is again a linear program.
- If (P) is a minimization problem, then (D) is a maximization problem and weak duality holds:

$$D^* \leq P^*$$

• Strong duality is desired and often holds:

$$D^* = P^*$$

- Every semidefinite program (P) has an optimal value P^* .
- To every semidefinite program (P), one can define a dual program (D) which is again a semidefinite program.
- If (P) is a minimization problem, then (D) is a maximization problem and weak duality holds:

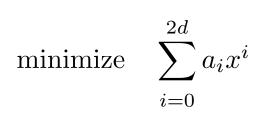
$$D^* \le P^*$$

• Strong duality is desired and often holds:

$$D^* = P^*$$

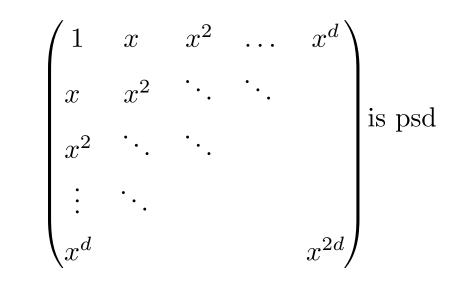
minimize
$$\sum_{i=0}^{2d} a_i x^i$$

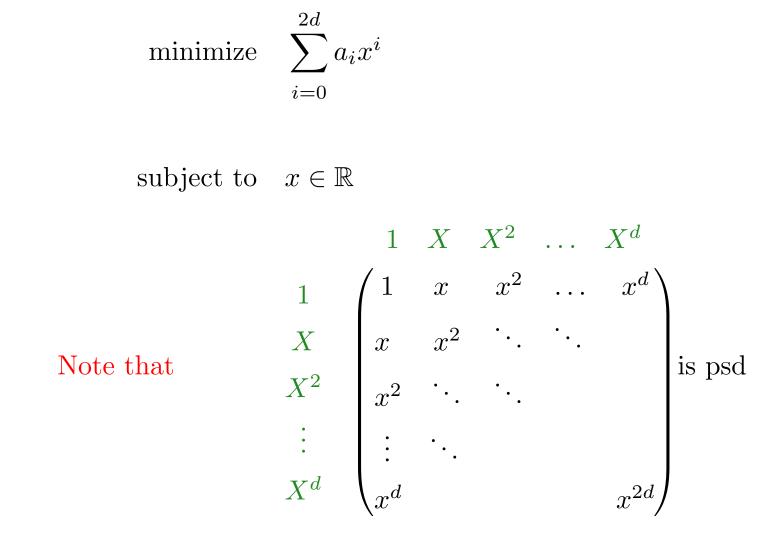
subject to $x \in \mathbb{R}$



subject to $x \in \mathbb{R}$

Note that





(P) minimize
$$\sum_{i=1}^{2d} a_i y_i + a_0$$

Set $f := \sum_{i=0}^{2d} a_i X^i$ and denote by (D) the semidefinite program dual to (P).

$$D^* \le P^* \le f^*.$$

$$D^* \le P^* \le f^*.$$

It turns out that (D) can be interpreted as:

(D) maximize μ subject to $f - \mu$ is sos

$$D^* \le P^* \le f^*.$$

It turns out that (D) can be interpreted as:

(D) maximize μ subject to $f - \mu$ is sos

Proposition. For every $p \in \mathbb{R}[X]$,

 $p \ge 0$ on $\mathbb{R} \implies p$ is a sum of two squares in $\mathbb{R}[X]$.

$$D^* \le P^* \le f^*.$$

It turns out that (D) can be interpreted as:

(D) maximize μ subject to $f - \mu$ is sos

Proposition. For every $p \in \mathbb{R}[X]$,

 $p \ge 0$ on $\mathbb{R} \implies p$ is a sum of two squares in $\mathbb{R}[X]$.

Corollary.

$$D^* = P^* = f^*$$

minimize $\sum a_{ij} x^i y^j$ $i+j \leq 4$

subject to $x, y \in \mathbb{R}$

where $a_{ij} \in \mathbb{R} \ (i+j \leq 4)$.

minimize
$$\sum_{i+j \le 4} a_{ij} x^i y^j$$

subject to $x, y \in \mathbb{R}$

Note that

$$\begin{pmatrix} 1 & x & y & x^2 & xy & y^2 \\ x & x^2 & xy & x^3 & x^2y & xy^2 \\ y & xy & y^2 & x^2y & xy^2 & y^3 \\ x^2 & x^3 & x^2y & x^4 & x^3y & x^2y^2 \\ xy & x^2y & xy^2 & x^3y & x^2y^2 & xy^3 \\ y^2 & xy^2 & y^3 & x^2y^2 & xy^3 & y^4 \end{pmatrix}$$
 is psd

where $a_{ij} \in \mathbb{R} \ (i+j \leq 4)$.

minimize
$$\sum_{i+j \le 4} a_{ij} x^i y^j$$

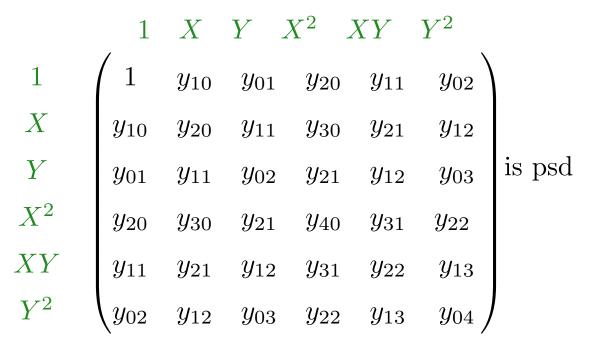
subject to $x, y \in \mathbb{R}$

Note that

where $a_{ij} \in \mathbb{R} \ (i+j \leq 4)$.

(P) minimize
$$\sum_{1 \le i+j \le 4} a_{ij} y_{ij} + a_{00}$$

subject to $y_{ij} \in \mathbb{R} \ (1 \le i + j \le 4)$



where $a_{ij} \in \mathbb{R}$ $(i+j \leq 4)$.

Set $f := \sum_{i+j \leq 4} a_{ij} X^{ij}$ and denote by (D) the semidefinite program dual to (P).

$$D^* \le P^* \le f^*.$$

$$D^* \le P^* \le f^*.$$

It turns out that (D) can be interpreted as:

(D) maximize μ subject to $f - \mu$ is sos

$$D^* \le P^* \le f^*.$$

It turns out that (D) can be interpreted as:

(D) maximize μ subject to $f - \mu$ is sos

Theorem (Hilbert). For every $p \in \mathbb{R}[X, Y]$ of degree ≤ 4 ,

 $p \ge 0$ on $\mathbb{R}^2 \implies p$ is a sum of three squares in $\mathbb{R}[X, Y]$.

David Hilbert: Ueber die Darstellung definiter Formen als Summe von Formenquadraten Math. Ann. XXXII 342-350 (1888)

http://www-gdz.sub.uni-goettingen.de/cgi-bin/digbib.cgi?PPN235181684_0032

$$D^* \le P^* \le f^*.$$

It turns out that (D) can be interpreted as:

(D) maximize μ subject to $f - \mu$ is sos

Theorem (Hilbert). For every $p \in \mathbb{R}[X, Y]$ of degree ≤ 4 ,

 $p \ge 0$ on $\mathbb{R}^2 \implies p$ is a sum of three squares in $\mathbb{R}[X, Y]$.

Corollary. $D^* = P^* = f^*$

David Hilbert: Ueber die Darstellung definiter Formen als Summe von Formenquadraten Math. Ann. XXXII 342-350 (1888)

http://www-gdz.sub.uni-goettingen.de/cgi-bin/digbib.cgi?PPN235181684_0032

• Unfortunately, not every polynomial $p \in \mathbb{R}[X_1, \ldots, X_n]$ with $p \ge 0$ on \mathbb{R}^n is a sum of squares of polynomials.

- Unfortunately, not every polynomial $p \in \mathbb{R}[X_1, \ldots, X_n]$ with $p \ge 0$ on \mathbb{R}^n is a sum of squares of polynomials.
- The first explicit example was found in 1967 by Motzkin:

$$p := X^4 Y^2 + X^2 Y^4 - 3X^2 Y^2 + 1$$

- Unfortunately, not every polynomial $p \in \mathbb{R}[X_1, \ldots, X_n]$ with $p \ge 0$ on \mathbb{R}^n is a sum of squares of polynomials.
- The first explicit example was found in 1967 by Motzkin:

$$p := X^4 Y^2 + X^2 Y^4 - 3X^2 Y^2 + 1$$

• In fact, there is even no $N \in \mathbb{N}$ such that p + N is a sum of squares in $\mathbb{R}[X, Y, Z]$.

- Unfortunately, not every polynomial $p \in \mathbb{R}[X_1, \dots, X_n]$ with $p \ge 0$ on \mathbb{R}^n is a sum of squares of polynomials.
- The first explicit example was found in 1967 by Motzkin:

$$p := X^4 Y^2 + X^2 Y^4 - 3X^2 Y^2 + 1$$

- In fact, there is even no $N \in \mathbb{N}$ such that p + N is a sum of squares in $\mathbb{R}[X, Y, Z]$.
- Described method always yields certified lower bounds, but they might by $-\infty$:

$$-\infty \le D^* = P^* \le f^*$$

- Unfortunately, not every polynomial $p \in \mathbb{R}[X_1, \ldots, X_n]$ with $p \ge 0$ on \mathbb{R}^n is a sum of squares of polynomials.
- The first explicit example was found in 1967 by Motzkin:

$$p := X^4 Y^2 + X^2 Y^4 - 3X^2 Y^2 + 1$$

- In fact, there is even no $N \in \mathbb{N}$ such that p + N is a sum of squares in $\mathbb{R}[X, Y, Z]$.
- Described method always yields certified lower bounds, but they might by $-\infty$:

$$-\infty \le D^* = P^* \le f^*$$

• But there are a lot of remedies...

For simplicity, we suppose m = 1 and write $g := g_1$ (technical difficulties which are however not very serious otherwise), i.e.

$$S = \{ x \in \mathbb{R}^n \mid g(x) \ge 0 \}.$$

For simplicity, we suppose m = 1 and write $g := g_1$ (technical difficulties which are however not very serious otherwise), i.e.

$$S = \{ x \in \mathbb{R}^n \mid g(x) \ge 0 \}.$$

Now we get a sequence $(P_k)_{2k\geq d}$ of relaxations such that

$$D_k^* \le P_k^* \le f^*$$
 and $\lim_{k \to \infty} D_k^* = \lim_{k \to \infty} P_k^* = f^*.$

Jean Lasserre: Global optimization with polynomials and the problem of moments SIAM J. Optim. **11**, No. 3, 796–817 (2001)

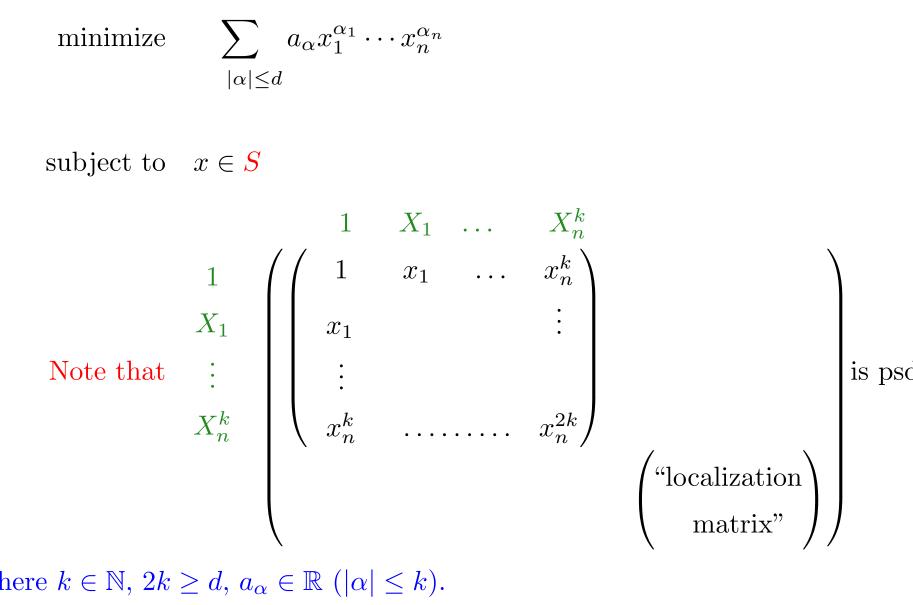
minimize
$$\sum_{|\alpha| \le d} a_{\alpha} x_1^{\alpha_1} \cdots x_n^{\alpha_n}$$

subject to $x \in S$

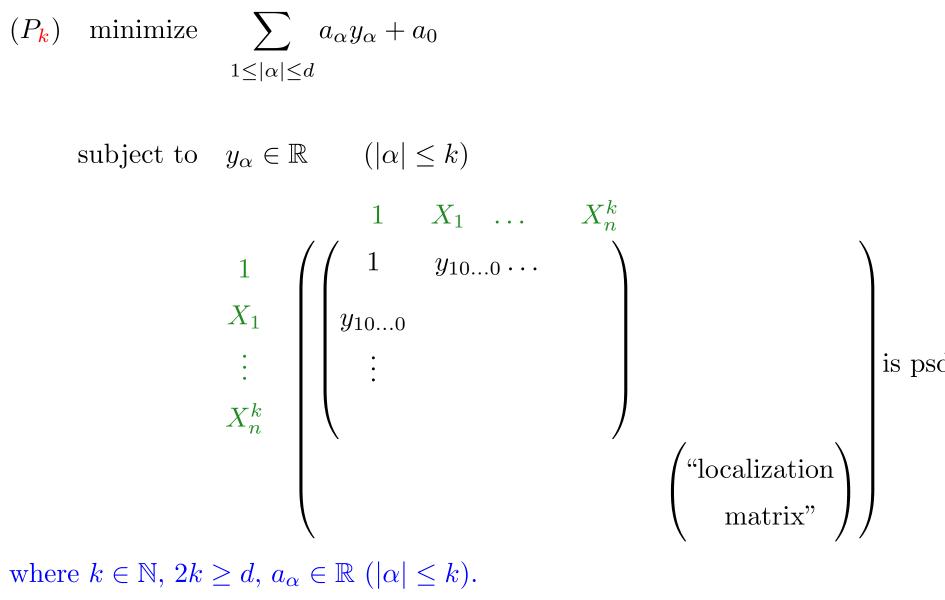
where $k \in \mathbb{N}, 2k \ge d, a_{\alpha} \in \mathbb{R} \ (|\alpha| \le k).$

$$\begin{array}{ll} \text{minimize} & \sum_{|\alpha| \le d} a_{\alpha} x_{1}^{\alpha_{1}} \cdots x_{n}^{\alpha_{n}} \\ \text{subject to} & x \in S \\ \\ \text{Note that} & \begin{pmatrix} \left(\begin{array}{cccc} 1 & x_{1} & \dots & x_{n}^{k} \\ x_{1} & & \vdots \\ \vdots & & \\ x_{n}^{k} & \dots & x_{n}^{2k} \end{pmatrix} \\ & & & \begin{pmatrix} \text{"localization} \\ \text{matrix"} \end{pmatrix} \end{pmatrix} \text{is pset} \end{array}$$

where $k \in \mathbb{N}, \ 2k \ge d, \ a_{\alpha} \in \mathbb{R} \ (|\alpha| \le k).$



where $k \in \mathbb{N}, 2k \geq d, a_{\alpha} \in \mathbb{R} \ (|\alpha| \leq k).$



where $k \in \mathbb{N}, 2k \geq d, a_{\alpha} \in \mathbb{R} \ (|\alpha| \leq k).$

Theorem (Schmüdgen, Putinar, ...) If f > 0 on S, then f = s + gt for sums of squares s, t in $\mathbb{R}[X_1, \ldots, X_n]$.

Theorem (Schmüdgen, Putinar, ...) If f > 0 on S, then f = s + gt for sums of squares s, t in $\mathbb{R}[X_1, \ldots, X_n]$.

Corollary (Lasserre). $(D_k^*)_{k \in \mathcal{N}}$ and $(P_k^*)_{k \in \mathcal{N}}$ are increasing sequences that converge to f^* and satisfy $D_k^* \leq P_k^* \leq f^*$. How fast?

Theorem (Schmüdgen, Putinar, ...) If f > 0 on S, then f = s + gt for sums of squares s, t in $\mathbb{R}[X_1, \ldots, X_n]$.

Corollary (Lasserre). $(D_k^*)_{k \in \mathcal{N}}$ and $(P_k^*)_{k \in \mathcal{N}}$ are increasing sequences that converge to f^* and satisfy $D_k^* \leq P_k^* \leq f^*$. How fast?

Theorem. There exists $C \in \mathbb{N}$ depending on f and g and $c \in \mathbb{N}$ depending on g such that

$$f^* - D_k^* \le \frac{C}{\sqrt[c]{k}}$$
 for big k .

On the complexity of Schmüdgen's Positivstellensatz Journal of Complexity **20**, No. 4, 529—543 (2004)

Optimization of polynomials on compact semialgebraic sets SIAM Journal on Optimization 15, No. 3, 805-825 (2005)

• Feasible solutions of (D_k) are certified lower bounds of f^* .

- Feasible solutions of (D_k) are certified lower bounds of f^* .
- Method converges from below to f^* .

- Feasible solutions of (D_k) are certified lower bounds of f^* .
- Method converges from below to f^* .
- Method converges to unique minimizers.

Optimization of polynomials on compact semialgebraic sets SIAM Journal on Optimization 15, No. 3, 805-825 (2005)

- Feasible solutions of (D_k) are certified lower bounds of f^* .
- Method converges from below to f^* .
- Method converges to unique minimizers. Disadvantage: Possibly from outside the set S.

Optimization of polynomials on compact semialgebraic sets SIAM Journal on Optimization 15, No. 3, 805-825 (2005)

- Feasible solutions of (D_k) are certified lower bounds of f^* .
- Method converges from below to f^* .
- Method converges to unique minimizers. Disadvantage: Possibly from outside the set S.
- If there is a unique minimizer and it lies in the interior of S,

Optimization of polynomials on compact semialgebraic sets SIAM Journal on Optimization 15, No. 3, 805-825 (2005)

- Feasible solutions of (D_k) are certified lower bounds of f^* .
- Method converges from below to f^* .
- Method converges to unique minimizers. Disadvantage: Possibly from outside the set S.
- If there is a unique minimizer and it lies in the interior of S, then the method produces a sequence of intervals containing f^* whose endpoints converge to f^* .

Optimization of polynomials on compact semialgebraic sets SIAM Journal on Optimization 15, No. 3, 805-825 (2005)

Implementations

- Henrion and Lasserre: GloptiPoly http://www.laas.fr/~henrion/software/gloptipoly/
- Prajna, Papachristodoulou, Parrilo: SOSTOOLS http://control.ee.ethz.ch/~parrilo/sostools/
- Both use the free SeDuMi solver by Jos Sturm
- But they need MATLAB and the MATLAB Symbolic Toolbox

Example: The maximum cut problem

Given a graph, i.e., an $n \in \mathbb{N}$ (number of nodes) and a set $E \subseteq \{(i,j) \in \{1,\ldots,n\}^2 \mid i < j\}$

(of edges),

Example: The maximum cut problem

Given a graph, i.e., an $n \in \mathbb{N}$ (number of nodes) and a set

$$E \subseteq \{(i, j) \in \{1, \dots, n\}^2 \mid i < j\}$$

(of edges), find the maximum cut value, i.e., the maximal possible number of edges that connect nodes with different signs when each node is assigned a sign + or -.

Example: The maximum cut problem

Given a graph, i.e., an $n \in \mathbb{N}$ (number of nodes) and a set

$$E \subseteq \{(i, j) \in \{1, \dots, n\}^2 \mid i < j\}$$

(of edges), find the maximum cut value, i.e., the maximal possible number of edges that connect nodes with different signs when each node is assigned a sign + or -.

maximize
$$\sum_{(i,j)\in E} \frac{1}{2}(1-x_ix_j)$$

subject to
$$x_i^2 = 1 \text{ for all } i \in \{1,\ldots,n\}$$

maximize
$$\sum_{(i,j)\in E} \frac{1}{2}(1-x_ix_j)$$

maximize
$$\sum_{(i,j)\in E} \frac{1}{2}(1-x_i x_j)$$

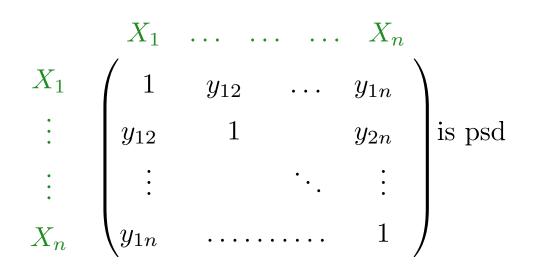
Note that
$$\begin{pmatrix} 1 & x_1 x_2 & \dots & x_1 x_n \\ x_2 x_1 & 1 & & x_2 x_n \\ \vdots & & \ddots & \vdots \\ x_n x_1 & \dots & & 1 \end{pmatrix}$$
 is psd

maximize
$$\sum_{(i,j)\in E} \frac{1}{2}(1-x_i x_j)$$

First MAXCUT relaxation

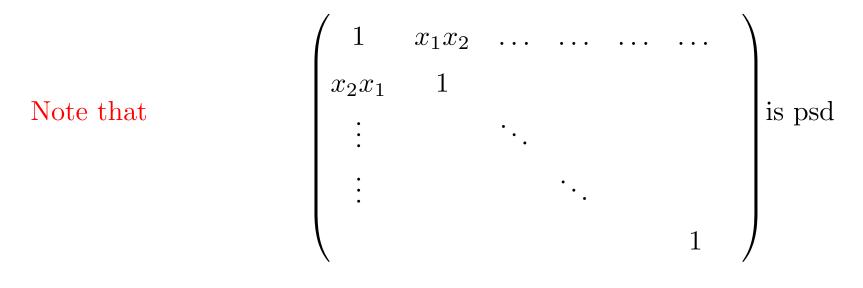
(P₁) maximize
$$\sum_{(i,j)\in E} \frac{1}{2}(1-y_{ij})$$

subject to $y_{ij} \in \mathbb{R}$ $(1 \le i < j \le n)$



maximize
$$\sum_{(i,j)\in E} \frac{1}{2}(1-x_i x_j)$$

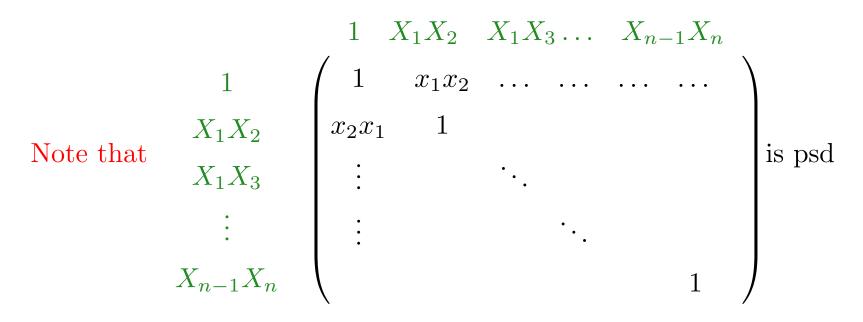
maximize
$$\sum_{(i,j)\in E} \frac{1}{2}(1-x_i x_j)$$



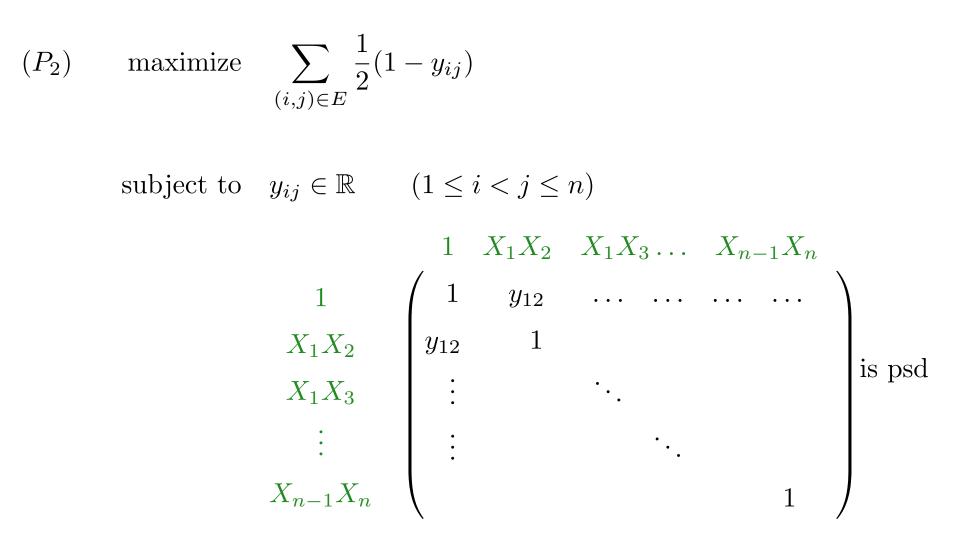
MAXCUT

maximize
$$\sum_{(i,j)\in E} \frac{1}{2}(1-x_i x_j)$$

subject to $x \in \{-1, 1\}^n$



Second MAXCUT relaxation



• The maximum cut problem is NP-complete

- The maximum cut problem is *NP*-complete
- The first relaxation gives a polynomial time algorithm which overestimates the maximum cut value at most by a factor of ≈ 1.1382 .

- The maximum cut problem is *NP*-complete
- The first relaxation gives a polynomial time algorithm which overestimates the maximum cut value at most by a factor of ≈ 1.1382 .
- The first relaxation is the famous algorithm of Goemans and Williamson.

- The maximum cut problem is *NP*-complete
- The first relaxation gives a polynomial time algorithm which overestimates the maximum cut value at most by a factor of ≈ 1.1382 .
- The first relaxation is the famous algorithm of Goemans and Williamson. From no polynomial algorithm it is known that it has a better approximation ratio.

- The maximum cut problem is NP-complete
- The first relaxation gives a polynomial time algorithm which overestimates the maximum cut value at most by a factor of ≈ 1.1382 .
- The first relaxation is the famous algorithm of Goemans and Williamson. From no polynomial algorithm it is known that it has a better approximation ratio. Existence of such an algorithm with ratio < 1.0625 implies P = NP (Hastad).

- The maximum cut problem is *NP*-complete
- The first relaxation gives a polynomial time algorithm which overestimates the maximum cut value at most by a factor of ≈ 1.1382 .
- The first relaxation is the famous algorithm of Goemans and Williamson. From no polynomial algorithm it is known that it has a better approximation ratio. Existence of such an algorithm with ratio < 1.0625 implies P = NP (Hastad).
- Solving the second relaxation is a polynomial time algorithm which yields the exact value for all planar graphs (consequence of results of Seymour, Barahona, Mahjoub),

- The maximum cut problem is *NP*-complete
- The first relaxation gives a polynomial time algorithm which overestimates the maximum cut value at most by a factor of ≈ 1.1382 .
- The first relaxation is the famous algorithm of Goemans and Williamson. From no polynomial algorithm it is known that it has a better approximation ratio. Existence of such an algorithm with ratio < 1.0625 implies P = NP (Hastad).
- Solving the second relaxation is a polynomial time algorithm which yields the exact value for all planar graphs (consequence of results of Seymour, Barahona, Mahjoub), and is conjectured to improve over the GW-algorithm.

- The maximum cut problem is *NP*-complete
- The first relaxation gives a polynomial time algorithm which overestimates the maximum cut value at most by a factor of ≈ 1.1382 .
- The first relaxation is the famous algorithm of Goemans and Williamson. From no polynomial algorithm it is known that it has a better approximation ratio. Existence of such an algorithm with ratio < 1.0625 implies P = NP (Hastad).
- Solving the second relaxation is a polynomial time algorithm which yields the exact value for all planar graphs (consequence of results of Seymour, Barahona, Mahjoub), and is conjectured to improve over the GW-algorithm.
- The *n*-th relaxation yields the exact maximum cut value.

Proposition. Suppose $p \in \mathbb{R}[X_1, \ldots, X_n]$ such that

 $p \ge 0$ on $\{-1, 1\}^n$.

Then f is a square modulo the ideal

$$I := (X_1^2 - 1, \dots, X_n^2 - 1) \subseteq \mathbb{R}[X_1, \dots, X_n].$$

Proposition. Suppose $p \in \mathbb{R}[X_1, \ldots, X_n]$ such that

 $p \ge 0$ on $\{-1, 1\}^n$.

Then f is a square modulo the ideal

$$I := (X_1^2 - 1, \dots, X_n^2 - 1) \subseteq \mathbb{R}[X_1, \dots, X_n].$$

Proof by algebra. By chinese remainder theorem

$$\mathbb{R}[X_1,\ldots,X_n]/I \cong \mathbb{R}^{\{-1,1\}^n} \cong \mathbb{R}^{2^n}.$$

Proposition. Suppose $p \in \mathbb{R}[X_1, \ldots, X_n]$ such that

 $p \ge 0$ on $\{-1, 1\}^n$.

Then f is a square modulo the ideal

$$I := (X_1^2 - 1, \dots, X_n^2 - 1) \subseteq \mathbb{R}[X_1, \dots, X_n].$$

Proof by algebra. By chinese remainder theorem

$$\mathbb{R}[X_1,\ldots,X_n]/I \cong \mathbb{R}^{\{-1,1\}^n} \cong \mathbb{R}^{2^n}.$$

Proof by algebraic geometry. I is a zero-dimensional radical ideal.

Proposition. Suppose $p \in \mathbb{R}[X_1, \ldots, X_n]$ such that

 $p \ge 0$ on $\{-1, 1\}^n$.

Then f is a square modulo the ideal

$$I := (X_1^2 - 1, \dots, X_n^2 - 1) \subseteq \mathbb{R}[X_1, \dots, X_n].$$

Proof by algebra. By chinese remainder theorem

$$\mathbb{R}[X_1,\ldots,X_n]/I \cong \mathbb{R}^{\{-1,1\}^n} \cong \mathbb{R}^{2^n}.$$

Proof by algebraic geometry. I is a zero-dimensional radical ideal.

Corollary. $D_n^* = P_n^* = f^*$

The story goes on...

Theorem (Lasserre). For every $p \in \mathbb{R}[X_1, \ldots, X_n]$, the following are equivalent:

(i) $p \ge 0$ on \mathbb{R}^n

The story goes on...

Theorem (Lasserre). For every $p \in \mathbb{R}[X_1, \ldots, X_n]$, the following are equivalent:

- (i) $p \ge 0$ on \mathbb{R}^n
- (ii) For every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$p + \varepsilon \sum_{i=1}^{n} \sum_{k=0}^{N} \frac{X_i^{2k}}{k!}$$
 is sos.

Jean Lasserre: A sum of squares approximation of nonnegative polynomials http://front.math.ucdavis.edu/math.AG/0412398

The story goes on...

Theorem (Nie, Demmel, Sturmfels). If p > 0 on \mathbb{R}^n , then p is sos modulo its own gradient ideal

$$I := \left(\frac{\partial f}{\partial X_1}, \dots, \frac{\partial f}{\partial X_n}\right)$$

Nie, Demmel, Sturmfels: Minimizing Polynomials via Sum of Squares over the Gradient Ideal http://front.math.ucdavis.edu/math.OC/0411342