Introduction to Semidefinite Programming II: Variations on the theta number

MFO seminar on Semidefinite Programming

Monique Laurent - CWI - Amsterdam & Tilburg University

May 30, 2010

Stable sets and graph coloring

$G = (V, E)$ graph $\rightsquigarrow$ $\bar{G} = (V, \bar{E})$: complementary graph.

- $\alpha(G)$= maximum size of a stable set $\rightsquigarrow$ stability number
- $\omega(G)$= maximum size of a clique in G $\rightsquigarrow$ clique number
- $\chi(G)$= (vertex) coloring number of G.
- $\chi_f(G)$= fractional coloring number of G.
- **Lovász' theta number:**

$$\vartheta(G) = \max\ J \cdot X \ \text{ s.t. }\ Tr(X) = 1,\ X_{ij} = 0\ (ij \in E),\ X \succeq 0$$

Sandwich theorem: $\alpha(G) \leq \vartheta(G) \leq \chi_f(\bar{G}) \leq \chi(\bar{G})$.

Approximating the Shannon capacity

- **The strong product** $G \cdot G'$ has vertex set $V \times V'$ and
$$(uu', vv') \in E(G \cdot G')$$
$$\Updownarrow$$
$$(u = v \text{ or } uv \in E(G)) \text{ and } (u' = v' \text{ or } u'v' \in E(G')).$$

- **Shannon capacity:** $\Theta(G) := \sup_k \sqrt[k]{\alpha(G^k)}$ [Shannon 1956]

- **Product property:** $\vartheta(G \cdot G') = \vartheta(G)\, \vartheta(G')$.

  Hence: $\Theta(G) \leq \vartheta(G)$. [Lovász 1979]

- This permits to show: $\Theta(C_5) = \sqrt{5}$.

  Proof: $\alpha(C_5^2) = 5$ and $\vartheta(C_5) = \sqrt{5}$.

- **Open:** $\Theta(C_7) =?$

Perfect graphs

Recall: $\omega(G) \leq \chi(G)$.

Berge [1962] calls G **perfect** if $\omega(G') = \chi(G')$ for *all* induced subgraphs G' of G.

Note: If G is perfect then no induced subgraph of G is an odd circuit of length ≥ 5 or its complement.

Perfect graph theorem [Lovász 1972]

G perfect $\Longleftrightarrow$ $\bar{G}$ perfect.

Strong perfect graph theorem

[Chudnovsky-Robertson-Seymour-Thomas 2002]

G perfect $\Longleftrightarrow$ no induced subgraph of G is an odd circuit of length ≥ 5 or its complement.

Polyhedral characterization of perfect graphs

- **Stable set polytope:** $\mathrm{STAB}(G)$= convex hull of incidence vectors of all stable sets in G.
- **Clique constrained polytope:**

$$Q\mathrm{STAB}(G) = \{x \in \mathbb{R}_+^V \mid x(C) \leq 1 \ \ (C \text{ clique})\}$$

Obviously, $\mathrm{STAB}(G) \subseteq Q\mathrm{STAB}(G)$.

- **Theorem** [Fulkerson-Chvatal 1972/75]

$$\mathrm{STAB}(G) = Q\mathrm{STAB}(G) \iff G \text{ is perfect.}$$

But this does not help (yet) for optimization!

$$\alpha(G) = \max_{x \in \mathrm{STAB}(G)} e^T x \ \leq \ \chi_f(\bar{G}) = \max_{x \in Q\mathrm{STAB}(G)} e^T x.$$

Finding a maximum stable set in a perfect graph is in P

- For G perfect, computing $\alpha(G)$ and $\chi(G)$ is in P.

  Proof: $\alpha(G) = \vartheta(G)$ and $\chi(G) = \vartheta(\bar{G})$.

- **For G perfect, one can also find a maximum stable set in polynomial time.** [Grötschel-Lovász-Schrijver 1981]

  • Order the vertices $v_1, \ldots, v_n$.

  • Construct graphs $G_0 := G \supseteq G_1 \supseteq \ldots \supseteq G_n$.

  - If $\alpha(G_0 \setminus v_1) = \alpha(G_0)$, set $G_1 = G_0 \setminus v_1$
  - otherwise, set $G_1 = G_0$.

  - Iterate. Then G_n is a maximum stable set.

- **Use SDP!**

  Open: Find a combinatorial algorithm?

Finding a minimum vertex coloring in G perfect is in P

- It suffices to find a stable set S meeting *all* max. size cliques.

  Indeed, then *color* $G \setminus S$ with $\omega(G \setminus S) = \omega(G) - 1$ colors, and S with one more color $\rightsquigarrow \omega(G)$ coloring of G

- **Strategy:** Grow a list of (affinely independent) maximum size cliques $Q_1, \ldots, Q_t$, and a stable set S meeting each Q_i.

  To find such S, *compute* a maximum weight stable set S for the weight function $w := \sum_{i=1}^{t} \chi^{Q_i}$.

  As G is perfect, $w(S) = t$, thus S meets each Q_i.

- If $\omega(G \setminus S) = \omega(G)$, find a maximum size clique Q_{t+1} in $G \setminus S$ and iterate.

- Else, S meets all maximum size cliques $\rightsquigarrow$ we are done.

Geometric reformulation of the theta number

- **The theta body** $\mathrm{TH}(G)$ is defined as

$$\left\{x \in \mathbb{R}^V \mid \exists X \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0,\ \mathrm{diag}(X) = x,\ X_{ij} = 0\ (ij \in E)\right\}$$

- **Obviously:** $\mathrm{STAB}(G) \subseteq \mathrm{TH}(G) \subseteq Q\mathrm{STAB}(G)$

- **Theorem:** G perfect $\Longleftrightarrow \mathrm{TH}(G) = \mathrm{STAB}(G)$
$\Longleftrightarrow \mathrm{TH}(G) = Q\mathrm{STAB}(G)$
$\Longleftrightarrow \mathrm{TH}(G)$ is a polytope.

- **Geometric reformulation of the theta number:**

$$\begin{aligned} \vartheta(G) &= \max_{x \in \mathrm{TH}(G)} e^T x \\ &= \max_{Y \in \mathcal{S}^+_{n+1}} \sum_{i \in V} Y_{ii} \ \text{ s.t. } \ Y_{00} = 1,\ Y_{0i} = Y_{ii}\ (i \in V), \\ &\qquad\qquad\qquad\qquad\qquad Y_{ij} = 0\ (ij \in E). \end{aligned}$$

Dual formulations of the theta number

$$\vartheta(G) = \max\ J\cdot X \ \text{ s.t. }\ I\cdot X = 1,\ X_{ij} = 0\ (ij \in E),\ X \succeq 0$$

$$= \min\ t \ \text{ s.t. }\ \underbrace{tI + \underbrace{\overbrace{\sum_{ij\in E} E_{ij}}^{A} - J}_{-Y}}_{Z \ \rightsquigarrow\ \frac{1}{t-1}Z=(u_i\cdot u_j)} \succeq 0$$

$$= \min\ t \ \text{ s.t. }\ tI + A - J \succeq 0,\ A_{ij} = 0\ (i = j \text{ or } ij \in \bar{E})$$

$$= \min\ \lambda_{\max}(Y) \ \text{ s.t. }\ Y_{ij} = 1\ (i = j \ \text{ or } ij \in \bar{E})$$

$$= \min\ t \ \text{ s.t. }\ u_i\cdot u_j = -\frac{1}{t-1}\ (ij \in \bar{E}),\ u_j \text{ unit vectors.}$$

$\rightsquigarrow$ *strict vector coloring*

Deriving the geometric formulation of $\vartheta(G)$ via $\mathrm{TH}(G)$

$$\vartheta(G) = \min\ t \ \text{ s.t. } \underbrace{tI + A - J}_{B} \succeq 0,\ A_{ij} = 0\ (i = j \text{ or } ij \in \bar{E}).$$

Schur complement: $B \succeq 0 \Longleftrightarrow C := \begin{pmatrix} t & e^T \\ e & I + \frac{1}{t}A \end{pmatrix} \succeq 0$

$$\begin{aligned} \vartheta(G) &= \min_{C \succeq 0}\ C_{00} \ \text{ s.t. }\ C_{0i} = C_{ii} = 1\ (i \in V), C_{ij} = 0\ (ij \in \bar{E}) \\ &= \max_{D \succeq 0}\ -\sum_{i \in V}(2D_{0i} + D_{ii}) \ \text{ s.t. }\ D_{00} = 1,\ D_{ij} = 0\ (ij \in E) \end{aligned}$$

Lemma: If D is optimum then $D_{0i} + D_{ii} = 0 \ \ \forall i \in V$.

Else multiply the ith row/column of D by $-\frac{D_{0i}}{D_{ii}} \rightsquigarrow$ better objective.

Changing signs at positions $0i$:

$$\begin{aligned} \vartheta(G) &= \max_{Y \succeq 0} \sum_{i \in V} Y_{ii} \ \text{ s.t. }\ Y_{00} = 1,\ Y_{0i} = Y_{ii}\ (i \in V),\ Y_{ij} = 0\ (ij \in E) \\ &= \qquad \max\ e^T x \ \text{ s.t. }\ x \in \mathrm{TH}(G). \end{aligned}$$

Reformulation of $\vartheta(G)$ via orthonormal representations

Definition: An *orthonormal representation* (O.R.) of G is a set of unit vectors u_i $(i \in V)$ satisfying $u_i^T u_j = 0$ for nonedges ij.

Theorem:

1. *Linear inequality description:* $\mathrm{TH}(G)$ is equal to

$$\{x \in \mathbb{R}^V \mid \textstyle\sum_{i\in V}(c^T u_i)^2 x_i \leq 1 \quad \forall c, u_i \text{ unit vectors with } u_i \text{ O.R. of } G\}$$

2. *Extreme point description:* $\mathrm{TH}(\bar{G})$ is equal to

$$\text{conv}\{\big((c^T u_1)^2, \ldots, (c^T u_n)^2\big) \mid \quad c, u_i \text{ unit vectors with } u_i \text{ O.R. of } G\}.$$

Thus: $\mathrm{TH}(G) = \{x \in \mathbb{R}^V \mid y^T x \leq 1 \quad \forall y \in \mathrm{TH}(\bar{G})\}$.

Reciprocity property

Theorem: $\alpha(G)\chi_f(G) \geq n$, with equality if G is vertex transitive.

Theorem: $\vartheta(G)\vartheta(\bar{G}) \geq n$, with equality if G is vertex transitive.

Corollary: $\vartheta(C_5) = \sqrt{5}$.

- Let $a = \vartheta(G)$, $aI + A - J \succeq 0$, $A_{ij} = 0$ if $i = j$ or $ij \in \bar{E}$.
  Let $b = \vartheta(\bar{G})$, $bI + B - J \succeq 0$, $B_{ij} = 0$ if $i = j$ or $ij \in E$.
  Then, $(aI + A - J) \circ (bI + B - J) \succeq 0$,
  $(aI + A - J) \circ J \succeq 0$, $J \circ (bI + B - J) \succeq 0$.
  Summing up: $abI - J \succeq 0 \Longrightarrow ab \geq n$.
- Let $x \in \mathrm{TH}(G)$ maximizing $\vartheta(G)$, and let $y \in \mathrm{TH}(\bar{G})$.
  As G is vertex transitive, we may assume that $x = ke$.
  Thus, $\vartheta(G) = kn$.
  Then, $x^T y \leq 1 \Longrightarrow k\ e^T y \leq 1 \Longrightarrow \frac{\vartheta(G)}{n}\vartheta(\bar{G}) \leq 1$.

Strengthening the theta number

$$\alpha(G) \leq \vartheta(G) \leq \bar{\chi}_f(G) \leq \bar{\chi}(G)$$

$$\vartheta(G) = \max\ J \cdot X \ \text{ s.t. } X \succeq 0,\ Tr(X) = 1,\ X_{ij} = 0\ (ij \in E)$$

1. **Improve toward** $\alpha(G)$ [McEliece et al. 78] [Schrijver 79]
   Add nonnegativity conditions: $X \geq 0$ $\rightsquigarrow \vartheta'(G)$
2. **Improve toward** $\bar{\chi}(G)$ [Szegedy 94]
   Relax the edge conditions: $X_{ij} \leq 0\ \ (ij \in E)$ $\rightsquigarrow \vartheta^+(G)$

$$\alpha(G) \leq \vartheta'(G) \leq \vartheta(G) \leq \vartheta^+(G) \leq \bar{\chi}(G).$$

There are SDP hierarchies converging to $\alpha(G)$, $\chi_f(G)$, and $\chi(G)$.

How to get stronger bounds?

$$\vartheta(G) = \max \sum_{i\in V} X_{ii} \text{ s.t. } X \text{ is indexed by } V \cup \{0\},$$
$$X \succeq 0, \quad X_{00} = 1, \ X_{0i} = X_{ii} \ (i \in V), \ X_{ij} = 0 \ (ij \in E).$$

Generalization: For $t \in \mathbb{N}$, index X by $\mathcal{P}_t := \{I \subseteq V \mid |I| \leq t\}$.
Denote the empty set by 0.

$$\mathbf{las^{(t)}(G)} := \max \sum_{i\in V} X_{ii} \text{ s.t. } X \text{ is indexed by } \mathcal{P}_t,$$

$$X \succeq 0, \ X_{00} = 1, \ X_{I,J} = X_{I',J'} \text{ if } I \cup J = I' \cup J',$$

$$X_{I,J} = 0 \text{ if } I \cup J \text{ contains an edge.}$$

Then: $\alpha(G) \leq \mathrm{las}^{(t)}(G)$, with equality if $t \geq \alpha(G)$.

Proof: S stable $\rightsquigarrow X_{I,J} = 1$ if $I \cup J \subseteq S$, and $X_{I,J} = 0$ otherwise.

Exploiting symmetry to compute the theta number

$$\vartheta'(G) = \max\ J\cdot X \text{ s.t. } X \succeq 0,\ Tr(X) = 1,\ X_{ij} = 0\ (ij \in E),\ X \geq 0$$

- $\mathcal{G} := \mathrm{Aut}(G)$: permutations of V preserving the edges of G.
- **The SDP defining $\vartheta(G)$ is invariant under action of $\mathcal{G}$:**

$$X \text{ feasible} \implies \forall g \in \mathcal{G}\ \ g(X) := (X_{g(i),g(j)}) \text{ feasible}$$
$$\implies \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} g(X) \text{ feasible}$$

with the same objective value.

$\rightsquigarrow$ We may assume that X **is invariant under action of $\mathcal{G}$:**

$$X_{i,j} = X_{i',j'} \text{ if } i' = g(i),\ j' = g(j) \text{ for some } g \in \mathcal{G}.$$

- $X = \sum_{t=1}^{N} x_t A_t$, where A_t are the 0/1 matrices corresponding to the orbits of $V \times V$ under action of $\mathcal{G}$.

  $\rightsquigarrow X \in \mathcal{A}_{\mathcal{G}}$: **algebra of invariant matrices**

$\rightsquigarrow$ SDP with N (# orbits) variables

A first explicit symmetry reduction

$\rightsquigarrow$ **One can write an explicit equivalent SDP with N variables and $N \times N$ matrices.** [de Klerk-Pasechnik-Schrijver 07]

- **Rescale the matrix A_t:** $B_t := \frac{A_t}{\sqrt{A_t \cdot A_t}}$

  $\rightsquigarrow$ orthormal basis of the algebra $\mathcal{A}_\mathcal{G}$ of invariant matrices

- **Multiplication parameters:** $B_r B_s = \sum_{t=1}^N \gamma_{r,s}^t B_t$

- **New $N \times N$ matrices:** $L_t = (\gamma_{t,s}^r)_{r,s=1}^N \quad (t = 1, \ldots, N)$

Theorem: For $x_1, \ldots, x_N \in \mathbb{R}$,

$$\sum_{t=1}^N x_t A_t \succeq 0 \Longleftrightarrow \sum_{t=1}^N x_t L_t \succeq 0.$$

Further symmetry reduction: block-diagonalization

$\rightsquigarrow$ **One can find an equivalent block-diagonal SDP with N variables and several smaller blocks.**

$\mathcal{A}_\mathcal{G}$: algebra of $V \times V$ matrices invariant under action of $\mathcal{G}$.

Wedderburn theorem: There exists a unitary matrix $U \in \mathbb{C}^{V\times V}$ such that

$$U\mathcal{A}_\mathcal{G}U^* = \bigoplus_{r=1}^{s} \underbrace{\mathbb{C}^{p_r\times p_r} \otimes I_{q_r}}_{*} \quad \text{for some} \quad p_1, q_1, \ldots, p_s, q_s \in \mathbb{N}.$$

$$* = \left\{ \begin{pmatrix} B & 0 & \ldots & 0 \\ 0 & B & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & B \end{pmatrix} \mid B \in \mathbb{C}^{p_r\times p_r} \quad \text{repeated } q_r \text{ times} \right\}$$

$\sum_{r=1}^{s} p_r^2 = N$: # of orbits of $V \times V$ under action of $\mathcal{G}$.

Application to the coding problem

- **Question:** What is the maximum cardinality $A(n,d)$ of a code $C \subseteq \{0,1\}^n$ with minimum Hamming distance d ?

  That is, $d_H(i,j) = |i \oplus j| \geq d$ for distinct $i,j \in C$.

- **Hamming graph $H(n,d)$:** vertex set $V = \{0,1\}^n$, with edges the pairs (i,j) with $d_H(i,j) \in [1, d-1]$.

  $\rightsquigarrow$ Compute $A(n,d) = \alpha(H(n,d))$.

- The Hamming graph has a rich automorphism group:
  - Permute the n coordinates.
  - Flip any set of coordinates: $i \in V \mapsto i_0 \oplus i$.

$\rightsquigarrow$ **Algebra of invariant matrices:** $\mathcal{A}_{\mathcal{G}} = \{\sum_{t=0}^{n} x_t A_t \mid x_t \in \mathbb{R}\}$, where $(A_t)_{i,j} = 1$ if $d_H(i,j) = t$ and $(A_t)_{i,j} = 0$ otherwise.

Link of the theta number to the Delsarte bound

Recall: $V = \{0,1\}^n$, and A_t is the $V \times V$ 0/1 matrix with entry 1 at positions with Hamming distance t.

$A_0 = I$, $\sum_{t=0}^n A_t = J$.

Fact: $\mathcal{A}_\mathcal{G}$ is a **commutative algebra** of dimension $n+1$ (known as the *Bose-Mesner algebra*).

Hence: All matrices in $\mathcal{A}_\mathcal{G}$ have a common basis of eigenvectors.

Thus: One can reduce the computation of $\vartheta(H(n,d))$ from an *SDP with* $2^n \times 2^n$ *matrices* to an *LP with* $\leq n+1$ *variables and constraints.*

Theorem: [Mc Eliece et al. 1978] [Schrijver 1979]
$\vartheta'(H(n,d))$ equals the LP bound introduced by Delsarte [1973].

Strengthening the Delsarte bound via the Lasserre bounds

Hamming graph: $G = H(n,d)$ with vertex set $V = \{0,1\}^n$ and with edges the pairs (i,j) with $d_H(i,j) \in [1, d-1]$.

- The SDP defining $\mathrm{las}^{(t)}(H(n,d))$ involves matrices of order $O(2^{nt})$.
- The number of orbits of $\mathcal{P}_t$ under action of $\mathrm{Aut}(G)$ is $O(n^{2^{2t-1}-1})$.

$\rightsquigarrow$ One can compute $\mathrm{las}^{(t)}(H(n,d))$ (to any precision) in time polynomial in n for any *fixed* t.

Practically:

• $t = 1$: This is the theta number (= LP Delsarte bound).

• $t = 2$: Gijswijt-Schrijver-Mittelmann [2010] give the *explicit block-diagonalization of the algebra of invariant matrices*, and compute (a strengthening of) the SDP bound $\mathrm{las}^{(2)}(H(n,d))$ for n up to 28 (for some values of d).

... An earlier SDP bound via the Terwilliger algebra

Historically, for the coding problem:

- X is indexed by $\emptyset$, and all singletons.
  $\rightsquigarrow$ 2-point bound (LP)
- X is indexed by $\emptyset$, all singletons, and all pairs.
  $\rightsquigarrow$ 4-point bound (SDP)
- **In-between:** X is indexed by $\emptyset$, all singletons, and all pairs containing a given element i_0. $\rightsquigarrow$ 3-point bound (SDP)

  For the Hamming graph, the algebra of invariant matrices is the Terwilliger algebra, of dimension $O(n^3)$, whose explicit block-diagonalization was given by Schrijver [2005].
- The block-diagonalization technique has since been applied to other problems (crossing number, quadratic assignment, etc.) [Bachoc, de Klerk, Pasechnik , Rendl, Sotirov, Vallentin, etc.]

Some numerical values for the coding problem

n	d	Delsarte (2-point)	Schrijver (3-point)	Gij-Mit-Sch (4-point)	lower bound
19	6	1289	1280	1237	1024
23	6	13775	13766	13 674	8192
19	8	145	142	135	128
20	8	290	274	**256**	**256**
25	8	6474	5477	5421	4096
26	8		9672	9275	4096
22	10	95	87	84	64
25	10	551	503	466	192
26	10	1040	886	836	384