Applications of semidefinite programming in Algebraic Combinatorics

Hajime Tanaka

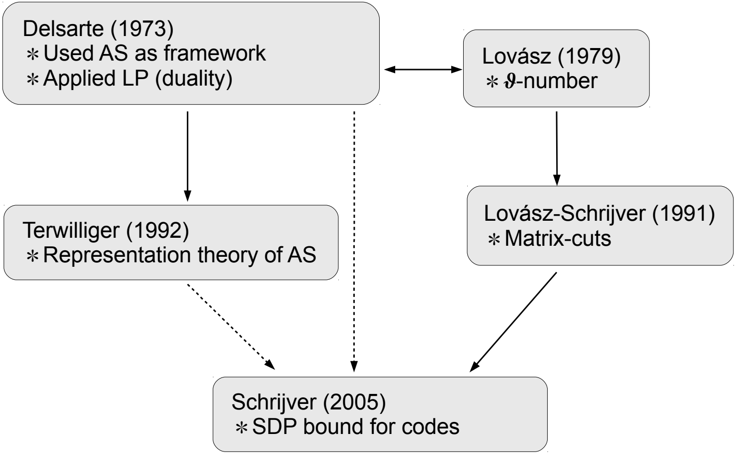
Tohoku University

The 23rd RAMP Symposium
October 24, 2011

We often want to $\cdots$

1. Bound the value of a numerical parameter of certain combinatorial configurations.
   - parameter $=$ size, index, $\cdots$
   - configurations $=$ codes, designs, spreads, ovoids, $\cdots$
2. Show that optimal (or nearly optimal) configurations have certain additional "regularity".
3. Classify the optimal (or nearly optimal) configurations.

A chart (“AS” stands for “association schemes”)



Codes

- $Q = \{0, 1, \ldots, q-1\}$: an alphabet of size $q \geqslant 2$
- $C \subseteq Q^n$: an (unrestricted) code of length n
- $\partial_H(\cdot, \cdot)$: the Hamming distance on Q^n :

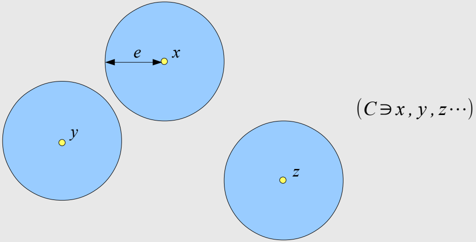
$$\partial_H(x, y) := |\{i = 1, \ldots, n : x_i \neq y_i\}|$$

  for $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in Q^n$.
- $d(C) := \min\{\partial_H(x, y) : x, y \in C, x \neq y\}$: the minimum distance of C

A classical problem

Remark

Set $e := \lfloor \frac{d(C)-1}{2} \rfloor$. Then



In other words, C is e-error-correcting.

- $A_q(n,d) := \max\{|C| : d(C) \geqslant d\}$

Problem

Determine $A_q(n,d)$. (Hard in general)

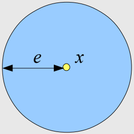
More modest problem

Problem*

Find a good upper bound on $A_q(n,d)$.

Example

- $B_e(x) := \{y \in Q^n : \partial_H(x,y) \leqslant e\}$: the ball with radius e and center x, where $e := \lfloor \frac{d-1}{2} \rfloor$:



- $A_q(n,d) \cdot |B_e(x)| = A_q(n,d) \cdot \sum_{i=0}^{e} \binom{n}{i} (q-1)^i \leqslant q^n$
- A code attaining equality in this sphere-packing bound is called a perfect code.

Association schemes

- X : a finite set
- $\mathbb{C}^X$: the $|X|$-dimensional column vector space over $\mathbb{C}$
- $\mathbb{C}^{X\times X}$: the set of $|X| \times |X|$ matrices over $\mathbb{C}$

- $\mathcal{R} = \{R_0, \dots, R_n\}$: a set of non-empty subsets of $X \times X$
- $A_0, \dots, A_n \in \mathbb{C}^{X\times X}$: the adjacency matrices :

$$(A_i)_{xy} := \begin{cases} 1, & \text{if } (x,y) \in R_i, \\ 0, & \text{if } (x,y) \notin R_i. \end{cases}$$

$\mathcal{R} = \{R_0, \ldots, R_n\}$: a set of non-empty subsets of $X \times X$

Definition

The pair $(X, \mathcal{R})$ is a (symmetric) association scheme if

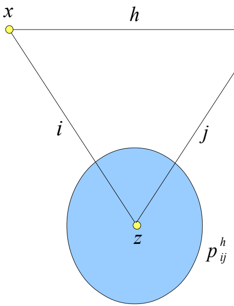
(AS1) $A_0 = I$ (the identity matrix),

(AS2) $A_0 + \cdots + A_n = J$ (the all 1's matrix),

(AS3) $A_i^{\mathsf{T}} = A_i \ (i = 0, \ldots, n)$,

(AS4) $A_i A_j \in \boldsymbol{A} := \mathrm{span}\{A_0, \ldots, A_n\} \ (i, j = 0, \ldots, n)$.

(AS4) $A_iA_j = \sum_{h=0}^{n} p_{ij}^h A_h \in \boldsymbol{A} := \mathrm{span}\{A_0, \ldots, A_n\}$



Remark

- $\boldsymbol{A}$ is a commutative matrix $*$-algebra, so has a basis of primitive idempotents, i.e., $E_iE_j = \delta_{ij}E_i, \ E_0 + \cdots + E_n = I$.
- $\boldsymbol{A}$: the Bose–Mesner algebra of $(X, \mathcal{R})$

Things to keep in mind

The Bose–Mesner algebra $\boldsymbol{A}$ of $(X, \mathcal{R})$ is commutative and has

- a basis of 0-1 (so nonnegative) matrices $A_0, \ldots, A_n$,
- a basis of idempotent (so positive semidefinite) matrices $E_0, \ldots, E_n$.

- Each of the bases is an orthogonal basis with respect to $\langle M, N \rangle = \text{trace}(M^* N)$ $(M, N \in \mathbb{C}^{X \times X})$.

The Hamming schemes

- $Q = \{0, 1, \ldots, q-1\}$
- $X = Q^n$
- $(x, y) \in R_i \overset{\text{def}}{\iff} \partial_H(x, y) = i \quad (i = 0, \ldots, n)$
- $\mathcal{R} = \{R_0, \ldots, R_n\}$
- $H(n, q) = (X, \mathcal{R})$: the Hamming scheme

Remark

- $H(n, q)$ admits $G := \mathfrak{S}_q \wr \mathfrak{S}_n$ as the group of automorphisms.
- A coincides with the commutant of G in $\mathbb{C}^{X \times X}$

The LP bound (Delsarte, 1973)

- For $x \in X$, set $\hat{x} = (0, \ldots, 0, 1, 0, \ldots, 0)^{\mathsf{T}} \in \mathbb{C}^X$ (a 1 in position x)
- $C \subseteq X$: a code with minimum distance $d(C) \geqslant d$
- $\chi_C = \sum_{x \in C} \hat{x}$: the characteristic vector of C
- $M := \frac{1}{|C|} \chi_C \chi_C^{\mathsf{T}} \in \mathbb{C}^{X \times X}$: nonnegative & positive semidefinite
- $\langle M, I \rangle = 1, \ \langle M, J \rangle = |C|$
- $\langle M, A_i \rangle = 0$ for $i = 1, \ldots, d-1$

The LP bound (Delsarte, 1973), continued

- Consider the following SDP problem:

$$\ell_{\mathsf{LP}} = \ell_{\mathsf{LP}}(n,q,d) = \max\langle M, J\rangle$$

  subject to

  1. $\langle M, I\rangle = 1$,
  2. $\langle M, A_i\rangle = 0 \ \ (i = 1,\ldots,d-1)$,
  3. M : nonnegative & positive semidefinite.

- Then $A_q(n,d) \leqslant \ell_{\mathsf{LP}}$.

Remark

ℓ_{LP} is the strengthening of Lovász's ϑ-number due to Schrijver (1979).

- By projecting M to $\boldsymbol{A}$, ℓ_{LP} turns to an LP:

$$(A_q(n,d) \leqslant)\ \ell_{\mathsf{LP}} = \ell_{\mathsf{LP}}(n,q,d) = \max\langle M,J\rangle$$

subject to

1. $\langle M,I\rangle = 1$,
2. $\langle M,A_i\rangle = 0\ \ (i=1,\dots,d-1)$,
3. $\sum_{i=0}^{n} \frac{\langle M,A_i\rangle}{\langle A_i,A_i\rangle} A_i = \sum_{i=0}^{n} \frac{\langle M,E_i\rangle}{\langle E_i,E_i\rangle} E_i \ \geqslant 0\ \&\ \succcurlyeq 0$, i.e.,
   $\langle M,A_i\rangle \geqslant 0\ \ (i=d,\dots,n),\ \ \langle M,E_i\rangle \geqslant 0\ \ (i=1,\dots,n)$.

Example

- $\ell_{\mathsf{LP}}(16,2,6) = 256$. In fact:
- $A_2(16,6) = 256$ (attained by the Nordstrom–Robinson code).

Delsarte exploited duality of LP $\cdots$

Remark

Many of the known universal bounds on $A_q(n,d)$, including the sphere packing bound, are obtained by constructing nice feasible solutions to the dual problem of ℓ_{LP}.

- $e := \lfloor \frac{d-1}{2} \rfloor$
- $\Psi_e(z) := \sum_{i=0}^{e}(-1)^i\binom{z-1}{i}\binom{n-z}{e-i}(q-1)^{e-i}$: Lloyd polynomial

Theorem (Lloyd)

If a perfect e-error-correcting code exists, then $\Psi_e(z)$ has e distinct zeros among the integers $1,2,\ldots,n$.

More generally, $\cdots$

- In fact, this reduction to LP works for any SDP problem

  $$\max \langle M, B_0 \rangle$$

  subject to

  1. $\langle M, B_i \rangle = b_i \quad (i = 1, \ldots, m)$,
  2. M is positive semidefinite,

  whenever $B_0, \ldots, B_m \in \boldsymbol{A}$ for an association scheme $(X, \mathcal{R})$.
- This was worked out in detail by Goemans and Rendl (1999) for the MAX-CUT problem.

Yet one more application

- Wilson (1984) used Delsarte's method to show the following Erdős–Ko–Rado theorem:

Theorem (Erdős–Ko–Rado, 1961)

Let $v > (t+1)(n-t+1)$ and let C be a collection of n-element subsets of $\Omega := \{1, \ldots, v\}$ with the property $|x \cap y| \geqslant t$ for all $x, y \in C$. Then

$$|C| \leqslant \binom{v-t}{n-t},$$

with equality if and only if

$$C = \{x \subseteq \Omega : |x| = n,\ w \subseteq x\}$$

for some t-element subset $w \subseteq \Omega$.

- This theorem has been extended to many other association schemes; cf. T. (2006, 2010).

The SDP bound (Schrijver, 2005)

- For simplicity, we only consider binary codes, i.e., codes in $H(n,2)$.
- $Q = \{0,1\}$, $X = Q^n$
- $A_0, A_1, \ldots, A_n$: the adjacency matrices

Remark

- The Bose–Mesner algebra $\boldsymbol{A} = \operatorname{span}\{A_0, \ldots, A_n\}$ coincides with the commutant of $G := \mathfrak{S}_2 \wr \mathfrak{S}_n$ in $\mathbb{C}^{X \times X}$.
- Below we shall consider the commutant of $\mathfrak{S}_n$ in $\mathbb{C}^{X \times X}$.

The Terwilliger algebra of $H(n, 2)$

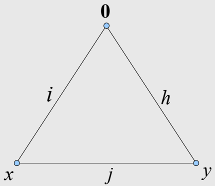
- $\mathbf{0} := (0, \ldots, 0) \in X$
- $E_0^\vee, \ldots, E_n^\vee \in \mathbb{C}^{X \times X}$: the dual idempotents :

$$(E_i^\vee)_{xy} := \begin{cases} 1, & \text{if } x = y,\ (\mathbf{0}, x) \in R_i, \\ 0, & \text{otherwise.} \end{cases}$$

- $\boldsymbol{T} := \mathbb{C}[A_0, \ldots, A_n, E_0^\vee, \ldots, E_n^\vee]$: the Terwilliger algebra
- $\boldsymbol{T} = \mathrm{span}\{E_i^\vee A_j E_h^\vee : i, j, h = 0, \ldots, n\}$

Remark

- $(E_i^\vee A_j E_h^\vee)_{xy} = \begin{cases} 1, & \text{if } (\mathbf{0}, x) \in R_i,\ (x, y) \in R_j,\ (\mathbf{0}, y) \in R_h, \\ 0, & \text{otherwise.} \end{cases}$



- $E_i^\vee A_j E_h^\vee = 0$ unless i, j, h satisfy the triangle inequality.

Two matrices from a code

- $C \subseteq X$: an (unrestricted) code

Lemma (Schrijver, 2005)

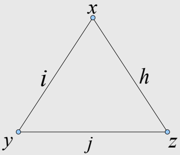
The matrices

$$M^1_{SDP} = \sum_{i,j,h} \lambda_{ijh} E_i^\vee A_j E_h^\vee, \quad M^2_{SDP} = \sum_{i,j,h} (\lambda_{0jj} - \lambda_{ijh}) E_i^\vee A_j E_h^\vee$$

with

$$\lambda_{ijh} := \frac{|X|}{|C|} \cdot \frac{|\{(x,y,z) \in C^3 : (x,y,z) \text{ satisfies } (\triangle)\}|}{|\{(x,y,z) \in X^3 : (x,y,z) \text{ satisfies } (\triangle)\}|}$$

are nonnegative & positive semidefinite, where



$(\triangle)$

- It follows that $\lambda_{000} = 1$ and $\sum_{i=0}^{n} \binom{n}{i} \lambda_{0ii} = |C|$.

- Consider the following SDP problem:

$$\ell_{\mathrm{SDP}} = \ell_{\mathrm{SDP}}(n, 2, d) = \max \sum_{i=0}^{n} \binom{n}{i} \lambda_{0ii}$$

subject to

1. $\lambda_{000} = 1$,
2. $\lambda_{ijh} = \lambda_{i'j'h'}$ if (i', j', h') is a permutation of (i, j, h),
3. $\sum_{i,j,h} \lambda_{ijh} E_i^{\vee} A_j E_h^{\vee}$: nonnegative & positive semidefinite,
4. $\sum_{i,j,h} (\lambda_{0jj} - \lambda_{ijh}) E_i^{\vee} A_j E_h^{\vee}$: nonnegative & positive semidefinite,
5. $\lambda_{ijh} = 0$ if $\{i, j, h\} \cap \{1, 2, \ldots, d-1\} \neq \emptyset$.

- Then $A_2(n, d) \leqslant \ell_{\mathrm{SDP}}$.

Computational results (Schrijver, 2005)

Bounds on $A_2(n,d)$

n	d	best lower bound known	ℓ_{SDP}	best upper bound previously known	ℓ_{LP}
19	6	1024	1280	1288	1289
23	6	8192	13766	13774	13775
25	6	16384	47998	48148	48148
19	8	128	142	144	145
20	8	256	274	279	290
25	8	4096	5477	5557	6474
27	8	8192	17768	17804	18189
28	8	16384	32151	32204	32206
22	10	64	87	88	95
25	10	192	503	549	551
26	10	384	886	989	1040

2 $\lambda_{ijh} = \lambda_{i'j'h'}$ if (i',j',h') is a permutation of (i,j,h).

Remark

- If we omit 2, then ℓ_{SDP} essentially coincides with the application of "matrix cuts" (Lovász–Schrijver, 1991) to ℓ_{LP}, followed by projecting to $\boldsymbol{T}$.
- Gijswijt (2005) observed that 2 makes a huge difference in the resulting bound.
- Currently, several hierarchies of SDP bounds on $A_2(n,d)$ of the form

$$\ell_{\text{LP}} \geqslant \ell_{\text{SDP}}^{(1)} \geqslant \ell_{\text{SDP}}^{(2)} \geqslant \cdots \geqslant \ell_{\text{SDP}}^{(k)} \geqslant \ldots (\geqslant A_2(n,d))$$

have been proposed, and some numerical computations have also been carried out (e.g., Lasserre (2001), Laurent 2007), Gvozdenović–Laurent–Vallentin (2009), Gijswijt–Mittelmann–Schrijver (2010)).

ℓ_{SDP} is huge $\cdots$

- $E_i^{\vee} A_j E_h^{\vee} \in \boldsymbol{T} \subseteq \mathbb{C}^{X \times X}$, $|X| = 2^n$.
- We reduce the size of ℓ_{SDP} by describing the Wedderburn decomposition (or block-diagonalization) of the matrix $*$-algebra $\boldsymbol{T}$ (in a form convenient for the computation).
- $\boldsymbol{T}$ is the commutant of $\mathfrak{S}_n$ in $\mathbb{C}^{X \times X}$, and its Wedderburn decomposition was also found by Dunkl (1976) in the study of Krawtchouk polynomials.

Structure of irreducible T-modules

- $T = \text{span}\{E_i^\vee A_j E_h^\vee : i, j, h = 0, \ldots, n\}$
- $\sum_{i=0}^n E_i^\vee = I, \ E_i^\vee E_j^\vee = \delta_{ij} E_i^\vee$
- $W \subseteq \mathbb{C}^X$: an irreducible T-module
- $r := \min\{i = 0, \ldots, n : E_i^\vee W \neq 0\}$: the endpoint of W

Theorem (Go, 2002)

We have

$$W = E_r^\vee W \perp E_{r+1}^\vee W \perp \cdots \perp E_{n-r}^\vee W.$$

More precisely,

$$\dim E_i^\vee W = \begin{cases} 1, & \textit{if } i = r, r+1, \ldots, n-r, \\ 0, & \textit{otherwise.} \end{cases}$$

The isomorphism class of W is determined by the endpoint r.

- With respect to the decomposition

  $$W = E_r^\vee W \perp E_{r+1}^\vee W \perp \cdots \perp E_{n-r}^\vee W,$$

  the matrix representing $E_i^\vee$ on W is

  $$E_i^\vee|_W = \left[\begin{array}{ccccc} & & & & \\ & & & & \\ & & 1 & & \\ & & & & \\ & & & & \end{array}\right] \begin{array}{l} r \\ \vdots \\ i \\ \vdots \\ n-r \end{array}$$

- $(E_i^\vee A_j E_h^\vee)|_W$ is a 'sparse' matrix. (In fact, its (i,h)-entry is written by a dual Hahn polynomial.)

Remark

$\ell_{\mathrm{SDP}}(n,q,d)$ $(q \geqslant 3)$ was studied by Gijswijt–Schrijver–T. (2006).

The quadratic assignment problem (QAP)

- X : a finite set
- $\mathfrak{S}(X)$: the symmetric group on X
- $\pi(g) \in \mathbb{R}^{X \times X}$: the permutation matrix of $g \in \mathfrak{S}(X)$:

$$(\pi(g))_{xy} = \delta_{x,gy} \quad (x, y \in X).$$

- $W, A \in \mathbb{R}^{X \times X}$: the distance and flow matrices
- Consider the following QAP (without linear term):

$$\min_{g \in \mathfrak{S}(X)} \frac{1}{2} \langle W, \pi(g)^{\mathsf{T}} A \pi(g) \rangle.$$

$\min_{g\in\mathfrak{S}(X)} \frac{1}{2}\langle W, \pi(g)^{\mathsf{T}} A\pi(g)\rangle$

- Suppose $A = A_1 \in \boldsymbol{A}$ for some association scheme $(X, \mathcal{R})$.
- Write $E_i = \frac{1}{|X|}\sum_{j=0}^{n} Q_{ji}A_j \quad (i = 0, \ldots, n)$.
- Then $(A_0, \ldots, A_n) \in (\mathbb{C}^{X\times X})^{n+1}$ satisfies

  1. $A_0 = I$, and $A_1, \ldots, A_n$ are nonnegative,
  2. $A_0 + \cdots + A_n = J$,
  3. $\sum_{j=0}^{n} Q_{ji}A_j$ is positive semidefinite $\quad (i = 0, \ldots, n)$.

- Conditions 1, 2, 3 hold for

$$(\pi(g)^{\mathsf{T}} A_0\pi(g), \ldots, \pi(g)^{\mathsf{T}} A_n\pi(g))$$

for any $g \in \mathfrak{S}(X)$, as well as their convex combinations.

SDP relaxation of QAP (de Klerk et al., to appear)

- Thus, the SDP problem

$$\min \frac{1}{2}\langle W, M_1\rangle$$

subject to

1. $M_0 = I$, and $M_1, \ldots, M_n$ are nonnegative,
2. $M_0 + \cdots + M_n = J$,
3. $\sum_{j=0}^{n} Q_{ji} M_j$ is positive semidefinite $(i = 0, \ldots, n)$,

gives a lower bound on QAP.

Minimum bisection problem

- $|X| = 2m$
- W : nonnegative
- $A = \left(\begin{array}{c|c} 0_m & J_m \\ \hline J_m & 0_m \end{array}\right)$
- $A_1 := A, \quad A_2 := \left(\begin{array}{c|c} J_m - I_m & 0_m \\ \hline 0_m & J_m - I_m \end{array}\right)$
- $\boldsymbol{A} = \text{span}\{I, A_1, A_2\}$

Traveling salesman problem (de Klerk et al., 2008)

- W : nonnegative
- $A = \begin{pmatrix} 0 & 1 & & & & 1 \\ 1 & 0 & 1 & & & \\ & 1 & \ddots & \ddots & & \\ & & \ddots & \ddots & 1 & \\ & & & 1 & 0 & 1 \\ 1 & & & & 1 & 0 \end{pmatrix}$
- $\mathcal{A} = \text{span}\{I, A, A^2, \ldots, A^{\lfloor |X|/2 \rfloor}\}$

Recent progress

Remark

- When A is the commutant of a finite group G in $\mathbb{C}^{X\times X}$, then this SDP relaxation of QAP can be strengthened further (de Klerk–Sotirov, to appear).
- For $H(n,2)$ (so $G = \mathfrak{S}_2 \wr \mathfrak{S}_n$), the Terwilliger algebra $\boldsymbol{T}$ plays a role again in this case.

Related topics

1. Spherical codes and spherical designs
   - LP bound
   - Kissing number problem ($k(4) = 24$ (Musin, 2008))
   - SDP bound for spherical codes (Bachoc–Vallentin, 2008)
2. Designs
   - Dual concept to codes
   - LP bound