Invariant Semidefinite Programs

Christine Bachoc

Université Bordeaux I, IMB

Modern trends in Optimization and its Application, IPAM

Optimization Tutorials, september 14-17, 2010

Outline of Part I

Invariant semidefinite programs,
B., Dion C. Gijswijt (CWI Amsterdam), Alexander Schrijver (CWI Amsterdam) and Frank Vallentin (TU Delft), arxiv:1007.2905

- Invariant semidefinite programs
- C*-algebras
- Representation theory of compact groups
- Applications to coding theory

Semidefinite programs

- A semidefinite program (SDP) in standard form:

$$\max \left\{ \langle C, X \rangle : X \succeq 0, \langle A_1, X \rangle = b_1, \ldots, \langle A_m, X \rangle = b_m \right\},$$

  where X, C, A_i are real symmetric matrices and $b_i \in \mathbb{R}$.
- $X \succeq 0$ stands for: X is positive semidefinite, meaning that X is a real symmetric matrix with non negative eigenvalues.
- $\langle C, X \rangle = \text{trace}(CX)$ is the standard inner product.
- A matrix X satisfying the above conditions is called a feasible solution; $\langle C, X \rangle$ is the objective function. Its maximum over the feasible region is called the optimal value of the program.

Semidefinite programs

- The set of positive semidefinite matrices is a closed convex cone which is self dual which means that:

$$A \succeq 0 \text{ iff for all } B \succeq 0,\ \langle A, B\rangle \geq 0.$$

- To the initial sdp (primal program) is associated a dual program:

$$\min \left\{ \langle b, x\rangle : -C + x_1 A_1 + \cdots + x_m A_m \succeq 0 \right\},$$

  where $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$.

- Weak duality holds: the primal optimal value is upper bounded by the dual optimal value.

Semidefinite programs

- Proof of weak duality: let X be primal feasible and x dual feasible.

$$\begin{aligned}\langle C, X\rangle &= \langle C - (x_1 A_1 + \cdots + x_m A_m), X\rangle + \langle x_1 A_1 + \cdots + x_m A_m, X\rangle \\ &= \underbrace{-\langle -C + x_1 A_1 + \cdots + x_m A_m, X\rangle}_{\leq 0} + \underbrace{x_1\langle A_1, C\rangle + \cdots + x_m\langle A_m, X\rangle}_{= x_1 c_1 + \cdots + x_m c_m} \\ &\leq x_1 b_1 + \cdots + x_m b_m = \langle b, x\rangle.\end{aligned}$$

- Strong duality, i.e. equality of the primal and dual optimal values hold under mild conditions i.e. Slatter condition: there exists a primal strictly feasible.

Invariant semidefinite programs

- We shall consider complex semidefinite programs where X, A_i, C are complex hermitian matrices, i.e. $X \in \mathbb{C}^{n\times n}$ and $X = X^*$.
- Let $G \subset U_n(\mathbb{C})$ be a finite group. It acts on positive semidefinite hermitian matrices by: $g.X = gXg^*$.
- The SDP is said to be G-invariant if:
  - X is feasible iff gX is feasible
  - $\langle X, C\rangle = \langle g.X, C\rangle$ (e.g. $g.C = C$ for all $g \in G$)
- A G-invariant SDP has an optimal solution which is itself invariant by G:

$$X' := \frac{1}{|G|}\sum_{g\in G} g.X$$

Invariant semidefinite programs

Theorem

If the SDP

$$\max\left\{\langle C, X\rangle : X \succeq 0, \langle A_1, X\rangle = b_1, \ldots, \langle A_m, X\rangle = b_m\right\}$$

is invariant by G, then it has the same optimal value as:

$$\max\left\{\langle C', X\rangle : X \in (\mathbb{C}^{n\times n})^G, X \succeq 0, \langle A_1', X\rangle = b_1, \ldots, \langle A_m', X\rangle = b_m\right\},$$

where

$$(\mathbb{C}^{n\times n})^G = \{X \in \mathbb{C}^{n\times n} : g.X = X\}$$

and

$$A_i' := \frac{1}{|G|}\sum_{g\in G} g.A_i.$$

Matrix $*$-algebras

- A matrix $*$-algebra $\mathcal{A}$ is a linear subspace of $\mathbb{C}^{n\times n}$ which is closed under multiplication and under taking the conjugate transpose.
- $\mathcal{A} = (\mathbb{C}^{n\times n})^G$ is a matrix $*$-algebra.
- Structure of matrix $*$-algebras:

Theorem

There exists $m_1, \ldots, m_d$ integers and an isomorphism φ of matrix $*$-algebras such that:

$$\varphi : \mathcal{A} \to \bigoplus_{k=1}^{d} \mathbb{C}^{m_k\times m_k}.$$

Moreover φ preserves inner products and the property of being positive semidefinite.

Reducing invariant semidefinite programs

- Let $\varphi(X) = (X_1, \ldots, X_d)$, $\varphi(C') = (C_1, \ldots, C_d)$, $\varphi(A'_i) = (A_{i1}, \ldots, A_{id})$. The symmetrized SDP transforms to:

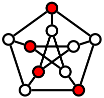
$$\max \Big\{ \sum_{k=1}^{d} \langle C_k, X_k \rangle : X_k \succeq 0,\ k = 1, \ldots, d$$

$$\sum_{k=1}^{d} \langle A_{ik}, X_k \rangle = b_i,\ i = 1, \ldots, m \Big\}$$

- The sizes of the matrix variables have changed from n to m_k.
- Need of an explicit isomorphism φ to compute $\varphi(C') = (C_1, \ldots, C_d)$, $\varphi(A'_i) = (A_{i1}, \ldots, A_{id})$.

Example: Lovász theta number of a graph

- Let $\Gamma = (V, E)$ a finite graph, $|V| = n$. An independent set S is a subset of V such that $S^2 \cap E = \emptyset$.



- The independence number of Γ:

$$\alpha(\Gamma) = \max_{S \text{ independent}} |S|$$

- Hard to compute. Lovász theta number provides an easy to compute approximation in the form of the optimal value of an SDP.

Example: Lovász theta number of a graph

- 1978, L. Lovász, *On the Shannon capacity of a graph*.

$$\vartheta(\Gamma) = \max\Big\{\langle J_n, X\rangle : \begin{array}{l} X = (X_{ij})_{1\le i,j\le n},\ X \succeq 0 \\ \langle I_n, X\rangle = 1, \\ X_{ij} = 0 \quad (i,j)\in E \end{array}\Big\}$$

- He proves the Sandwich Theorem:

Theorem

$$\alpha(\Gamma) \le \vartheta(\Gamma) \le \chi(\overline{\Gamma})$$

Proof of $\alpha(\Gamma) \le \vartheta(\Gamma)$: if S is an independent set, then B:

$$B_{ij} = \frac{1}{|S|}\mathbf{1}_S(i)\,\mathbf{1}_S(j)$$

is feasible. Moreover $\sum_{i,j} B_{ij} = |S|$, thus $|S| \le \vartheta(\Gamma)$.

Graphs with symmetries

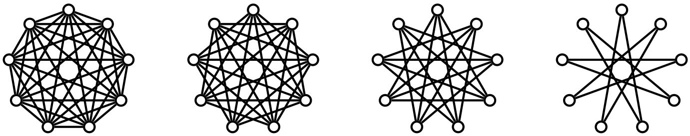
- Assume $G = \mathrm{Aut}(\Gamma)$ is the group of permutations $\sigma \in S_n$ that sends edges to edges.
- Then G acts on $X \in \mathbb{C}^{n\times n}$ by permutations:

$$\sigma.X = P(\sigma)XP(\sigma)^* = (X_{\sigma^{-1}(i)\sigma^{-1}(j)})_{i,j}$$

  and leaves ϑ invariant. Thus ϑ can be replaced by its symmetrization under G.
- $X_{ij} = \langle X, E_{ij}\rangle$. The matrix E'_{ij} is the characteristic function of the orbit under G of the pair (i,j). When $(i,j) \in [n]^2$ they form a basis of $(\mathbb{C}^{n\times n})^G$. We need to compute the image of this basis by the isomorphism φ.

An easy example: circular graphs

- Let p, q integers, $p \geq 2q$. Let $K_{p/q}$ the graph with vertex $V = [p]$ and edge set $E = \{(i,j) : q \leq |i-j| \leq p-q\}$.
- Examples: $K_{9/1} = K_9$, $K_{9/2} = \overline{C_9}$, $K_{9/3}$, $K_{9/4} = C_9$.



- The dihedral group D_p of order $2p$ acts on $K_{p/q}$.

An easy example: circular graphs

- With the discrete Fourier transform, we have $X \in (\mathbb{C}^{p\times p})^{D_p}$ iff
$$X_{ij} = \sum_{k=0}^{\lfloor p/2 \rfloor} x_k \cos(\frac{2k\pi}{p}(i-j)).$$
- The map $X \mapsto (x_0, \dots, x_{\lfloor p/2 \rfloor})$ is the wanted isomorphism
$$\varphi : (\mathbb{C}^{p\times p})^G \to \mathbb{C}^{1+\lfloor p/2 \rfloor}$$
- The sdp ϑ becomes the linear program:
$$\begin{aligned}\vartheta(K_{p/q}) \quad = \quad \max\Big\{ px_0 \ :\quad & x_k \geq 0, \ \sum_{k=0}^{\lfloor p/2 \rfloor} x_k = 1, \\ & \sum_{k=0}^{\lfloor p/2 \rfloor} x_k \cos(\frac{2jk\pi}{p}) = 0, \quad q \leq j \leq \lfloor p/2 \rfloor \ \Big\}\end{aligned}$$

Group representations

- Let G be a compact group. Examples: $G = O_n(\mathbb{R})$, $U_n(\mathbb{C})$, a finite group.
- G is endowed with its Haar measure λ: a positive measure on G which is left and right invariant ($\lambda(gA) = \lambda(Ag) = \lambda(A)$).
- A finite dimensional representation of G is a finite dimensional complex vector space V on which G acts linearly and continuously.
- Such a representation is always a unitary representation: indeed, starting from an arbitrary inner product $\langle u, v \rangle$ on V one can construct a G-invariant inner product:

$$\langle u, v \rangle' = \int_G \langle gu, gv \rangle d\lambda(g).$$

Group representations

- V is said to be irreducible if it contains no non trivial subspace W such that $gW = W$ for all $g \in G$ (i.e. no G-subspace).
- If W is a G-subspace then $W^{\perp}$ is also a G-subspace, where orthogonality is with respect to a G-invariant inner product. Thus the space V splits into the direct sum of irreducibles (Maschke theorem).
- The G-homomorphisms are the homomorphisms of linear spaces that commute with the action of G, i.e. the $T : V_1 \to V_2$ such that $T(gv) = gT(v)$. If $V_1 = V_2 = V$ they form the algebra $\mathrm{End}^G(V)$ which is a C^*-algebra.

Group representations

- From Maschke theorem, V has an irreducible decomposition

$$V = W_0 \perp W_1 \perp \cdots \perp W_d$$

- Grouping the components which are pairwise G-isomorphic defines the isotypic subspaces of V.
- We fix a set $\mathcal{R} = \{R_k, k \geq 0\}$ of representatives of the isomorphism classes of irreducible representations of G.
- For $k \geq 0$, let $\mathcal{MI}_k$ denote the isotopic subspace of V related to R_k, i.e. the sum of the G-subspaces of V which are isomorphic to R_k. Then $\mathcal{MI}_k \simeq R_k^{m_k}$ and m_k is called the multiplicity of R_k in V.

Group representations

- Schur lemma : if V is irreducible, then

$$\mathrm{End}^G(V) = \{\lambda\,\mathrm{Id},\ \lambda \in \mathbb{C}\} \simeq \mathbb{C}.$$

  Proof: if $T \in \mathrm{End}^G(V)$, then T has an eigenvalue λ.
  $W := \ker(T - \lambda I)$ is a non zero G-subspace of V thus $W = V$.

- In general, if

$$V = \perp_{k \in I_V} \mathcal{MI}_k, \quad \mathcal{MI}_k \simeq R_k^{m_k}, \quad I_V := \{k : m_k \neq 0\}.$$

  then

$$\mathrm{End}^G(V) \simeq \bigoplus_{k \in I_V} \mathbb{C}^{m_k \times m_k}.$$

Group representations

- Let M be a compact set, on which G acts continuously. We assume M is given a G-invariant positive measure μ. Examples: $G = O_n(\mathbb{R})$ and $M = S^{n-1}$; $G = \mathrm{Aut}(\Gamma)$ and $M = V$.
- The space $\mathcal{C}(M)$ of complex valued continuous functions on M is a unitary representation of G, for the action:

$$(g.f)(x) := f(g^{-1}x)$$

  and the inner product:

$$\langle f_1, f_2 \rangle = \frac{1}{\mu(M)} \int_M f_1(x)\overline{f_2(x)} d\mu(x).$$

- $\mathcal{C}(M)$ is infinite dimensional (but we shall consider only finite dimensional G-subspaces $V \subset \mathcal{C}(M)$).

Group representations

- An explicit isomorphism $\text{End}^G(V) \simeq \oplus\mathbb{C}^{m_k \times m_k}$: let

$$\mathcal{MI}_k = \bigoplus_{i=1}^{m_k} H_{k,i}, \quad H_{k,i} \simeq R_k.$$

- Let $(e_{k,i,1}, \ldots, e_{k,i,d_k})$ an orthonormal basis of $H_{k,i}$, where $d_k = \dim(R_k)$, such that the complex numbers $\langle ge_{k,i,s}, e_{k,i,t} \rangle$ do not depend on i.
- We define $m_k \times m_k$ matrices $E_k(x, y)$ by:

$$E_{k,ij}(x, y) := \sum_{s=1}^{d_k} e_{k,i,s}(x)\overline{e_{k,j,s}(y)}.$$

Group representations

- $E_k(x,y)$ is G-invariant:

$$E_k(gx,gy) = E_k(x,y).$$

- A change in the decomposition of $\mathcal{MI}_k$ or in the choice of basis of $H_{k,i}$ changes $E_k(x,y)$ to $AE_k(x,y)A^*$ for some $A \in \mathrm{Gl}_{m_k}(\mathbb{C})$.
- To $(F_1,\dots,F_{|I_V|}) \in \oplus_{k \in I_V} \mathbb{C}^{m_k \times m_k}$ we associate

$$F(x,y) = \sum_{k \in I_V} \langle F_k, \overline{E_k(x,y)} \rangle$$

  which in turn defines the element $T_F \in \mathrm{End}^G(V)$:

$$(T_F(f))(x) := \int_M F(x,y) f(y) d\mu(y).$$

Example: the binary Hamming space

- Let $H_n := \{0,1\}^n$, with the Hamming distance $d_H(x,y)$:

$$d_H(x,y) := |\{i, 1 \le i \le n : x_i \ne y_i\}|.$$

- The group $G := S_2 \wr S_n$ acts on H_n and leaves d_H invariant.
- Moreover, G acts two-point homogeneously on H_n, meaning that the orbits of G on pairs $(x,y) \in H_n^2$ are characterized by the value of $d_H(x,y)$.
- Decomposition of $\mathbb{C}^{H_n}$ as a G-module: let $\chi_z(x) := (-1)^{x \cdot z}$ denote the characters of $(\{0,1\}^n, +)$.

$$\begin{aligned} \mathbb{C}^{H_n} &= \bigoplus_{z \in H_n} \mathbb{C}\chi_z \\ &= \bigoplus_{k=0}^{n} P_k, \quad P_k := \bigoplus_{wt(z)=k} \mathbb{C}\chi_z \end{aligned}$$

The binary Hamming space

- The subspaces P_k are invariant under G, irreducible and pairwise non isomorphic. They must be because remember

$$n+1 = \dim((\mathbb{C}^{H_n \times H_n})^G) = \dim(\mathrm{End}^G(\mathbb{C}^{H_n})) = \sum m_k^2.$$

- The multiplicities m_k are equal to 1.

$$\begin{aligned} E_k(x,y) &= \sum_{wt(z)=k} \chi_z(x)\chi_z(y) = \sum_{wt(z)=k} (-1)^{(x-y)\cdot z} \\ &= \sum_{j=0}^{k} (-1)^j \binom{t}{j}\binom{n-t}{k-j}, \quad t := d_H(x,y) \\ &= K_k^n(t) \quad \text{Krawtchouk polynomials.} \end{aligned}$$

The binary Hamming space

- A binary code with minimal distance d is a subset C of H_n such that

$$d_H(C) := \min\{d_H(x,y) : x \neq y, (x,y) \in C^2\} = d.$$

- In view of applications to error correction, combinatorial coding theory asks for

$$A(n,d) := \max\{|C| : C \subset H_n, d_H(C) \geq d\}.$$

- $A(n,d)$ is the independence number of the graph $\Gamma(n,d)$ with vertex set $V = H_n$ and edge set

$$E = \{(x,y) \in H_n^2 : 1 \leq d_H(x,y) \leq d-1\}.$$

An upper bound for $A(n,d)$

- We have
$$A(n,d) \leq \vartheta'(\Gamma(n,d)) = (\vartheta'(\Gamma(n,d)))^G$$
where in ϑ' we add the constraint: $X_{ij} \geq 0$.
- We have seen: $F \in (\mathbb{C}^{H_n \times H_n})^G$ iff
$$F(x,y) = \sum_{k=0}^{n} f_k K_k^n(d_H(x,y))$$
and: $F \succeq 0$ iff $f_k \geq 0$ for all $0 \leq k \leq n$.
- Thus the SDP defining $(\vartheta'(\Gamma(n,d)))^G$ becomes a linear program in the $n+1$ variables f_k with at most $n+1$ inequalities. In coding theory it is known under the name of Delsarte linear programming bound and prior to Lovász (Delsarte, 1973).

Review on Part I

- Semidefinite programs having symmetries can be reduced to smaller size, through an isomorphism

$$\varphi : (\mathbb{C}^{n\times n})^G \to \bigoplus_{k=1}^{d} \mathbb{C}^{m_k\times m_k}.$$

- An example: Lovász theta number of a graph Γ with automorphism group G.
- Applications to the binary Hamming space H_n. Here

$$\varphi : (\mathbb{C}^{2^n\times 2^n})^G \to \bigoplus_{k=0}^{n} \mathbb{C}$$

$$F \mapsto (f_0,\dots,f_n), \quad F(x,y) = \sum_{k=0}^{n} f_k K_k^n(d(x,y)).$$

Outline of Part II

- Stronger SDP upper bounds for $A(n, d)$
- Other spaces in coding theory
- Extremal problems on the sphere

Stronger upper bounds for $A(n, d)$

- Idea: exploit constraints on k-subsets of binary words.
- A. Schrijver, 2005, *New code upper bounds from the Terwilliger algebra and semidefinite programming*. Uses triples.
- D.C. Gijswijt, H.D. Mittelmann, A. Schrijver, *Semidefinite code bounds based on quadruple distances*. They give a general framework for k-tuples.
- Let $\mathcal{P}_k$ the set of subsets of H_n of size at most k. Symmetric matrices X indexed by H_n can be viewed as functions:

  $$X : \mathcal{P}_2 \to \mathbb{C}$$

  We want to introduce functions:

  $$X : \mathcal{P}_k \to \mathbb{C}$$

Stronger upper bounds for $A(n,d)$

- Let $X : \mathcal{P}_k \to \mathbb{C}$ and let $T \in \mathcal{P}_k$. Let $M_T(X)$ be indexed by:

$$I_T := \{S \in \mathcal{P}_{(k+|T|)/2} : T \subset S\}$$

and defined by:

$$\big(M_T(X)\big)_{S,S' \in I_T} := X(S \cup S').$$

- Let the semidefinite program:

$$\vartheta_k(n,d) := \max\Big\{ \sum_{v \in H_n} X(\{v\}) : \quad \begin{array}{ll} X(\emptyset) = 1 & \\ X(S) = 0 & d_H(S) \leq d-1 \\ M_T(X) \succeq 0 & T \in \mathcal{P}_k \Big\} \end{array}$$

Stronger upper bounds for $A(n, d)$

- Then we have

$$A(n,d) \leq \vartheta_k(n,d).$$

  Proof: if C is a binary code with minimal distance d, then X defined by

$$X(S) = \prod_{x \in S} \mathbf{1}_C(x) = \begin{cases} 1 \text{ if } S \subset C \\ 0 \text{ otherwise} \end{cases}$$

  is a feasible solution, and $\sum_{v \in H_n} X(\{v\}) = |C|$.
- For $k = 2$ we recover Lovász $\vartheta'(\Gamma(n,d))$.

Stronger upper bounds for $A(n, d)$

- The group $G = \text{Aut}(H_n)$ acts on $\mathcal{P}_k$ and leaves $\vartheta_k(n, d)$ invariant, thus one can restrict to X being G-invariant:

$$X(gS) = X(S) \quad \text{for all } g \in G,\ S \in \mathcal{P}_k.$$

- The number of orbits of G on $\mathcal{P}_k$ is of the order of $n^{2^{k-1}-1}$. Thus the resulting program has polynomial size (for fixed k).
- Then,

$$M_T(X) \in (\mathbb{C}^{I_T \times I_T})^{\text{Stab}(T,G)}.$$

Stronger upper bounds for $A(n,d)$

- The case $k=3$: we can assume $T=\{0^n\}$. Then, $\text{Stab}(T,G)=S_n$. We need to understand

$$(\mathbb{C}^{H_n\times H_n})^{S_n}.$$

- The orbit of $(x,y)\in H_n\times H_n$ under S_n is given by the triple: $(wt(x), wt(y), d_H(x,y))$.
- A. Schrijver, 2005: block diagonalization of $(\mathbb{C}^{H_n\times H_n})^{S_n}$.
- F. Vallentin, 2007: using the framework of group representations and work of Dunkl, gives an expression of the $E_k(x,y)$ with Hahn polynomials.

Stronger upper bounds for $A(n, d)$

- In the case $k = 4$, there are two cases:
  - $|T| = 2$, $(\mathbb{C}^{H_n \times H_n})^{S_w \times S_{n-w}}$
  - $T = \emptyset$, $(\mathbb{C}^{H_n^2 \times H_n^2})^G$
- $T = 2$: easy.

$$(\mathbb{C}^{H_n \times H_n})^{S_w \times S_{n-w}} = (\mathbb{C}^{H_w \times H_w})^{S_w} \otimes (\mathbb{C}^{H_{n-w} \times H_{n-w}})^{S_{n-w}}.$$

- $T = \emptyset$: not so easy. Amounts to have an alphabet of size 4.

$$(\mathbb{C}^{H_n^2 \times H_n^2})^G = \left(((\mathbb{C}^{4\times 4})^{S_2})^{\otimes n}\right)^{S_n} = \mathrm{Sym}^n((\mathbb{C}^{4\times 4})^{S_2}),$$

- D.C. Gijswijt (2010): a general method to decompose $\mathrm{Sym}^n(\mathcal{A})$ from a decomposition of $\mathcal{A}$.

The results for $A(n, d)$

- (GMS 2010) The computation of $\vartheta_k(n, d)$ for $k = 3, 4$ has lead to improved upper bounds of $A(n, d)$ for values of n in the range $18 \leq n \leq 28$. In particular, $A(20, 8) = 256$ is proved.
- Using Delsarte LP method, very good upper bounds in the form of explicit functions of the parameters (n, d) where given from explicit dual feasible solution (MRRW (1978); Levenshtein).
- Using Delsarte LP method, the best known asymptotic bound for the rate

$$\frac{1}{n} \log(A(n, d)$$

was obtained (MRRW (1978)).
- Open question: is it possible to improve it with $\vartheta_k(n, d)$?

Comments

- The SDP program defining $\vartheta_k(n, d)$ can be viewed as a SDP relaxation of the independence number of a hypergraph.
- It has further applications to extremal problems in coding theory relative to constraints on k points.
- It can also be understood in terms of hierachies of SDP for 0/1 programs (Lovász-Schrijver, Lasserre).

Other spaces

- Let (M, d_M) be a metric space. We introduce

$$A(M,d) := \max\{|C| : C \subset M, d_M(C) \geq d\}.$$

- Many metric spaces are of interest in coding theory, due to the growing number of applications.
- It is a general fact that these spaces are usually huge spaces, affording huge groups of symmetries.
- One can follow the same line as for the Hamming space: $A(M,d)$ is the independence number of a graph $\Gamma(M,d)$ thus is upper bounded by $\vartheta'(\Gamma(M,d))$ on which the group $G = \text{Aut}(M, d_M)$ acts.

Examples

Space	Group	Polynomial
Hamming space $\mathbf{q}^n$	$S_q \wr S_n$	Krawtchouk
Johnson space	S_n	Hahn
q-Johnson space	$\mathrm{Gl}_n(\mathbb{F}_q)$	q-Hahn
A^n, A is H-sym	$H \wr S_n$	multivariate Krawtchouk
Projective space	$\mathrm{Gl}_n(\mathbb{F}_q)$	matrix q-Hahn
Permutations	$S_n \times S_n$	characters
Sphere S^{n-1}	$O_n(\mathbb{R})$	Gegenbauer
Projective spaces	$O_n(\mathbb{R})$, $U_n(\mathbb{C})$	Jacobi
Grassmann spaces	$O_n(\mathbb{R})$, $U_n(\mathbb{C})$	multivariate Jacobi

The sphere

- Euclidean space $\mathbb{R}^n$, inner product $x \cdot y = \sum_{i=1}^n x_i y_i$.

$$S^{n-1} := \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n : x \cdot x = 1\}.$$

- The orthogonal group $O_n(\mathbb{R})$ acts homogeneously on S^{n-1}.
- The angular distance $d_\theta(x, y)$ is $O_n(\mathbb{R})$-invariant:

$$d_\theta(x, y) = \arccos(x \cdot y)$$

- Moreover, $O_n(\mathbb{R})$ acts two-point homogeneously on S^{n-1}.

Spherical codes

- For a spherical code $C \subset S^{n-1}$, let

$$d_\theta(C) := \min\{d_\theta(x, y) : (x, y) \in C^2, x \neq y\}.$$

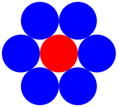
- Problem: to determine

$$A(S^{n-1}, \theta_{\min}) := \max\{|C| : C \subset S^{n-1}, d_\theta(C) \geq \theta_{\min}\}.$$

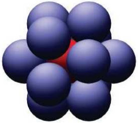
- Case $\theta_{\min} = \pi/3$: $A(S^{n-1}, \pi/3) = \tau_n$ is the kissing number of dimension n, the maximal number of spheres that can touch simultaneously a central sphere, without overlapping, all spheres having same radius.

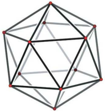
Kissing number in dimensions 2 and 3

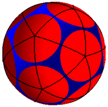
- Dimension 2: $\tau_2 = 6$, unique configuration.



- Dimension 3: Regular icosahedron, 12 points. The minimal angle is $\simeq 63.4^\circ$.







History

- 1694: Isaac Newton and David Gregory:

$$\tau_3 = 13\ ?$$

- 1953: Schütte and Van der Waerden prove $\tau_3 = 12$.
- 1956: other proof by Leech.

The known values of τ_n:

- 1979: $\tau_8 = 240$ (E_8) and $\tau_{24} = 196560$ (min. vectors of the Leech lattice) Levenshtein; indep. Odlysko et Sloane
- 2003: $\tau_4 = 24$ (D_4), Oleg Musin

Extremal problems on sphere

- $C \subset S^{n-1}$ avoids $\Omega \subset (S^{n-1})^2$ if, for all $(x, y) \in C^2$, $(x, y) \notin \Omega$.
- Let λ a measure on S^{n-1}, let

$$A(S^{n-1}, \Omega, \lambda) = \sup \{\lambda(C) : C \subset S^{n-1} \text{ measurable}, \ C \text{ avoids } \Omega\}.$$

- $\Omega = \{(x, y) : d_\theta(x, y) \in]0, \theta_{\min}[\}$ and λ is the counting measure denoted μ_c.
  Ω-avoiding sets are spherical codes with minimal distance $\theta_{\min}$.
- $\Omega = \{(x, y) : d_\theta(x, y) = \theta\}$ for some value $\theta \neq 0$, and $\lambda = \mu$.

Extremal problems on sphere

- Computing $A(S^{n-1}, \Omega, \lambda)$ is difficult.
- We aim at a SDP relaxation, of "theta type".
- Problem: the analog on S^{n-1} of the cone of psd matrices.

Definition

We say that $F \in \mathcal{C}((S^{n-1})^2)$ is positive definite, denoted $F \succeq 0$, if $F(x, y) = \overline{F(y, x)}$ and, for all k, for all $(x_1, \ldots, x_k) \in (S^{n-1})^k$,

$$\big(F(x_i, x_j)\big)_{1 \leq i,j \leq k} \succeq 0.$$

Primal and dual theta numbers

- Recall the theta (prime) number of the graph $\Gamma = (V, E)$, $V = [n]$:

$$\vartheta'(\Gamma) = \max\Big\{\langle X, J_n\rangle : \quad X \succeq 0,\ X \geq 0, \\ \langle X, I_n\rangle = 1,\ X_{ij} = 0 \text{ for all } (i,j) \in E\Big\}.$$

- The dual expression:

$$\vartheta'(\Gamma) = \min\Big\{t : X \succeq 0, \quad X_{ii} \leq t - 1, \\ X_{ij} \leq -1 \text{ for all } (i,j) \notin E\Big\}.$$

Theta numbers for the sphere

- Replace $\mathbb{C}^{n\times n}$ with $\mathcal{C}(S^{n-1} \times S^{n-1})$.
- Let $\Omega^c = \{(x, y) : (x, y) \notin \Omega \text{ and } x \neq y\}$

$$\begin{aligned}\vartheta_2(S^{n-1}, \Omega) = \sup \{\ \langle F, 1\rangle :\quad & F \succeq 0,\ F \geq 0, \\ & \langle F, \mathbf{1}_\Delta\rangle = 1, \\ & F(x, y) = 0 \text{ for all } (x, y) \in \Omega\ \}.\end{aligned}$$

$$\begin{aligned}\vartheta_1(S^{n-1}, \Omega) = \inf \{\ t : F \succeq 0,\quad & F(x, x) \leq t - 1, \\ & F(x, y) \leq -1 \text{ for all } (x, y) \in \Omega^c\ \},\end{aligned}$$

Theta numbers for the sphere

- These cone linear programs are not pairwise dual because the topological dual of $\mathcal{C}(S^{n-1})$ is the space $\mathcal{M}(S^{n-1})$ of Borel regular measures on S^{n-1}.
- The appropriate version depends on the nature of Ω:

$$\begin{aligned} \Omega =]0, \theta_{min}[\quad & A(S^{n-1}, \Omega, \lambda) \leq \vartheta_1(S^{n-1}, \Omega) \\ \Omega = \{\theta\} \quad & A(S^{n-1}, \Omega, \lambda) \leq \vartheta_2(S^{n-1}, \Omega) \end{aligned}$$

- These programs are invariant under $O_n(\mathbb{R})$. Thus we can assume $F \in \mathcal{C}((S^{n-1})^2)^{O_n(\mathbb{R})}$.

Harmonic analysis on S^{n-1}

- Harmonic polynomials:

$$\mathrm{Harm}_k^n := \{P \in \mathbb{R}[x_1, .., x_n], P \text{ hom. }, \deg(P) = k, \Delta P = 0\}$$

  where Δ is the Laplace operator:

$$\Delta = \sum_{k=1}^{n} \frac{\partial^2}{\partial x_k^2}.$$

  Harm_k^n is an irreducible representation of $O_n(\mathbb{R})$.
- Let H_k^n the functions on S^{n-1} obtained from Harm_k^n. Then

$$\mathcal{C}(S^{n-1}) = \oplus_{k \geq 0} H_k^n$$

Harmonic analysis on S^{n-1}

- $m_k = 1$ and $E_k(x, y) = P_k^n(x \cdot y)$ where $P_k^n(u)$ are the Gegenbauer polynomials with parameter $n/2 - 1$:

$$\begin{cases} P_k^n \in \mathbb{R}[t], \ \deg(P_k^n) = k, \ P_k^n(1) = 1 \\ \int_{-1}^{1} P_k^n(t) P_l^n(t)(1 - t^2)^{(n-3)/2} dt = 0, \quad k \neq l. \end{cases}$$

- The measure $\mathbf{1}_{[-1,1]}(1 - t^2)^{(n-3)/2} dt$ is the measure induced by the Lebegue measure of S^{n-1} on inner products.
- $F \succeq 0$ and $O_n(\mathbb{R})$-invariant iff

$$F(x, y) = \sum_{k \geq 0} f_k P_k^n(x \cdot y) \text{ with } f_k \geq 0.$$

and the sum converges uniformly (Schöenberg 1942).

Spherical codes

- We obtain the linear program, where $s := \cos\theta_{\min}$:

$$\vartheta_1 \quad = \inf \Big\{ 1 + \sum_{k\geq 0} f_k : \quad f_k \geq 0 \\ \sum_{k\geq 0} f_k P_k^n(u) \leq -1 \quad -1 \leq u \leq s \Big\}$$

  We recover Delsarte LP bound. Moreover ϑ_1 is the limit of a decreasing sequence of finite dimensional SDP.

- A. Odlysko, NJA. Sloane (1978): computed upper bounds for the kissing number problem corresponding to $\theta_{\min} = \pi/3$.
  Cases where the LP bound is optimal: $n = 8, 24$.
  G. Kabatiansky, V. Levenshtein (1978): asymptotic upper bound.

Sets avoiding one angle

- Case $\Omega = \{\theta\}$. (B., G. Nebe, F. de Oliveira Filho, F. Vallentin 2009).

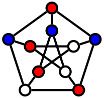
$$\begin{aligned} \vartheta_2 \quad = \sup \Big\{ f_0 \ : \quad & f_k \geq 0 \text{ for all } k \geq 0 \\ & \sum_{k \geq 0} f_k = 1 \\ & \sum_{k \geq 0} f_k P_k^n(s) = 0 \} \end{aligned}$$

- Let $m(s)$ be the minimum of $P_k^n(s)$ for $k = 0, 1, 2, \ldots$. Then

$$\vartheta_2 = \frac{m(s)}{m(s) - 1}.$$

Chromatic numbers

- Let $\Gamma = (V, E)$ be a finite graph. The chromatic number $\chi(\Gamma)$ of Γ is the smallest number of colors needed to color V s.t. connected vertices receive different colors.



- The color classes are independent sets of Γ that partition V. Hence

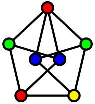
$$\chi(\Gamma) \geq \frac{|V|}{\alpha(\Gamma)} \geq \frac{|V|}{\vartheta(\Gamma)}.$$

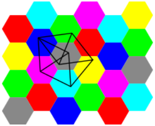
- Similarly, for $\Gamma = \Gamma(n, s)$ the graph with $V = S^{n-1}$ and $E = \{(x, y) : x \cdot y = s\}$, ($\chi_m$: measurable color classes)

$$\chi_m(\Gamma(n, s)) \geq \frac{1}{\vartheta_2} = \frac{m(s) - 1}{m(s)}.$$

Chromatic numbers

- The chromatic number of Euclidean space $\chi(\mathbb{R}^n)$: points at distance 1 receive different colors.
- $\chi(\mathbb{R}) = 2, \quad 4 \leq \chi(\mathbb{R}^2) \leq 7$:





- We have $\chi_m(\mathbb{R}^n) \geq \chi_m(\Gamma(n, s))$ for all s. Taking the limit when $s \to 1$,

$$\chi_m(\mathbb{R}^n) \quad \geq 1 + \frac{(j_{\alpha+1})^\alpha}{2^\alpha \Gamma(\alpha+1)|J_\alpha(j_{\alpha+1})|} \approx_{+\infty} (1.165)^n$$

where $\alpha = (n-3)/2$, J_α denotes the Bessel function of the first kind and j_α denotes its first positive zero.

Some results

- Chromatic numbers: the inequality $\chi_m(\mathbb{R}^n) \geq \chi_m(\Gamma(n,s))$ gave improvements on the known lower bounds for $n \geq 10$.
- F. Vallentin, F.M. de Oliveira Filho: better lower bounds obtained from a linear program involving functions of $\mathbb{R}^n$ (instead of S^{n-1}).
- Kissing numbers: with triple constraints, SDP improvements of upper bounds (B., F. Vallentin 2008; F. Vallentin, H. Mittelmann 2009).