Section 7

Kissing Number Problem

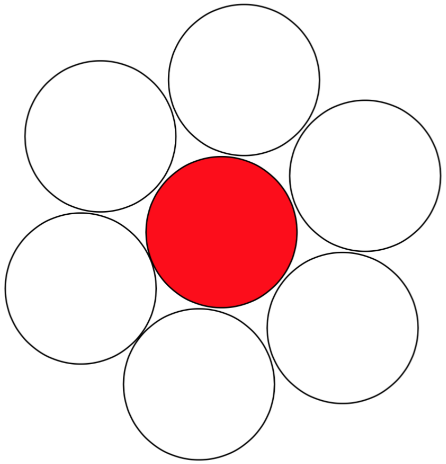
Definition

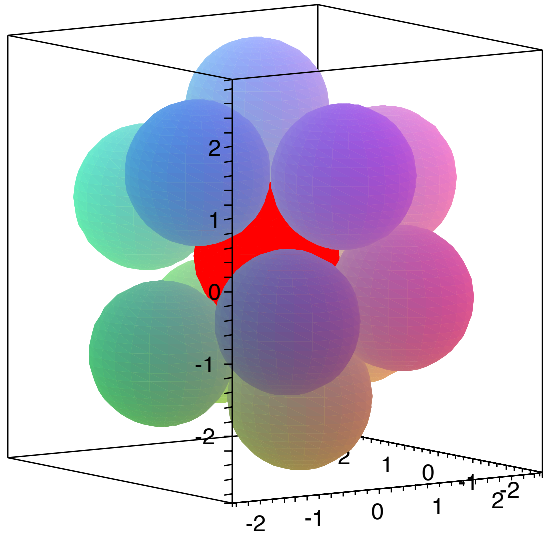
Given $n, K \in \mathbb{N}$, determine whether n unit spheres can be placed adjacent to a central unit sphere so that their interiors do not overlap

Funny story: ***Newton and Gregory went down the pub...***

Some examples

$n = 6, K = 2$ $\qquad$ $n = 12, K = 3$





more dimensions

n	τ (lattice)	τ (nonlattice)
0	0	
1	2	
2	6	
3	12	
4	24	
5	40	
6	72	
7	126	
8	240	
9	272	(306)*
10	336	(500)*
11	438	(582)*
12	756	(840)*
13	918	(1130)*
14	1422	(1582)*
15	2340	
16	4320	
17	5346	
18	7398	
19	10668	
20	17400	
21	27720	
22	49896	

Equivalent formulation

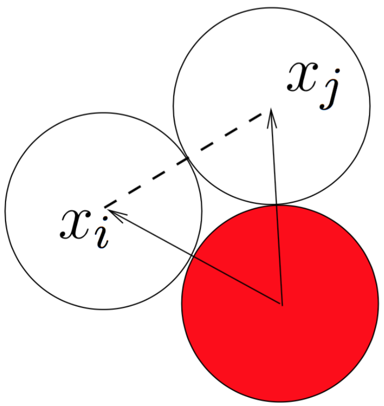
Given $n, K \in \mathbb{N}$**, determine whether there exist** n **vectors** $x_1, \ldots, x_n \in \mathbb{R}^K$ **such that:**

$$\begin{array}{rcl} \forall i \leq n \qquad \|x_i\|_2^2 & = & 1 \\ \forall i < j \leq n \qquad \|x_i - x_j\|_2^2 & \geq & 1. \end{array}$$

Spherical codes

- $\mathbb{S}^{K-1} \subset \mathbb{R}^K$ unit sphere centered at origin
- *K-dimensional spherical z-code*:
  - (finite) subset $\mathcal{C} \subset \mathbb{S}^{K-1}$
  - $\forall x \neq y \in \mathcal{C} \qquad x \cdot y \leq z$
- non-overlapping interiors:

$$\forall i < j \quad ||x_i - x_j|| \geq 2 \qquad \Longleftrightarrow x_i \cdot x_j \geq \cos(\frac{\pi}{3}) = \frac{1}{2}$$



...can use norm-1 projections on $\mathbb{S}^{K-1}$ instead

Lower bounds

- Construct spherical $\frac{1}{2}$-code $\mathcal{C}$ with $|\mathcal{C}|$ large
- Nonconvex NLP formulations
- SDP relaxations
- Combination of the two techniques

MINLP formulation

Maculan, Michelon, Smith 1995

Parameters:

- K: space dimension
- n: upper bound to $\mathsf{kn}(K)$

Variables:

- $x_i \in \mathbb{R}^K$: center of i-th vector
- $\alpha_i = 1$ iff vector i in configuration

$$\left.\begin{array}{rrcl} \max & \sum_{i=1}^{n} \alpha_i & & \\ \forall i \le n & ||x_i||^2 & = & \alpha_i \\ \forall i < j \le n & ||x_i - x_j||^2 & \ge & \alpha_i \alpha_j \\ \forall i \le n & x_i & \in & [-1,1]^K \\ \forall i \le n & \alpha_i & \in & \{0,1\} \end{array}\right\}$$

Reformulating the binary products

- **Additional variables:** $\beta_{ij} = 1$ **iff vectors** i, j **in configuration**
- **Linearize** $\alpha_i \alpha_j$ **by** β_{ij}
- **Add constraints:**

$$\begin{array}{lrcl} \forall i < j \le n & \beta_{ij} & \le & \alpha_i \\ \forall i < j \le n & \beta_{ij} & \le & \alpha_j \\ \forall i < j \le n & \beta_{ij} & \ge & \alpha_i + \alpha_j - 1 \end{array}$$

AMPL and Baron

- **Certifying YES**
  - $n = 6, K = 2$: **OK, 0.60s**
  - $n = 12, K = 3$: **OK, 0.07s**
  - $n = 24, K = 4$: **FAIL, CPU time limit (100s)**
- **Certifying NO**
  - $n = 7, K = 2$: **FAIL, CPU time limit (100s)**
  - $n = 13, K = 3$: **FAIL, CPU time limit (100s)**
  - $n = 25, K = 4$: **FAIL, CPU time limit (100s)**

Almost useless

Modelling the decision problem

$$\left.\begin{array}{rrcl}\max\limits_{x,\alpha} & \alpha & & \\ \forall i\le n & \|x_i\|^2 & = & 1 \\ \forall i<j\le n & \|x_i - x_j\|^2 & \ge & \alpha \\ \forall i\le n & x_i & \in & [-1,1]^K \\ & \alpha & \ge & 0\end{array}\right\}$$

- **Feasible solution** (x^*, α^*)
- *KNP instance is YES iff* $\alpha^* \ge 1$

[Kucherenko, Belotti, Liberti, Maculan, *Discr. Appl. Math.* 2007]

AMPL and Baron

- Certifying YES
  - $n = 6, K = 2$: FAIL, CPU time limit (100s)
  - $n = 12, K = 3$: FAIL, CPU time limit (100s)
  - $n = 24, K = 4$: FAIL, CPU time limit (100s)
- Certifying NO
  - $n = 7, K = 2$: FAIL, CPU time limit (100s)
  - $n = 13, K = 3$: FAIL, CPU time limit (100s)
  - $n = 25, K = 4$: FAIL, CPU time limit (100s)

Apparently even more useless

But more informative (arccos $\alpha =$ min. angular sep)

Certifying YES by $\alpha \geq 1$

- $n = 6, K = 2$: OK, 0.06s
- $n = 12, K = 3$: OK, 0.05s
- $n = 24, K = 4$: OK, 1.48s
- $n = 40, K = 5$: FAIL, CPU time limit (100s)

What about polar coordinates?

$$
\begin{aligned}
y = (y_1, \ldots, y_K) \quad &\rightarrow \quad (\rho, \vartheta_1, \ldots, \vartheta_{K-1}) \\
\rho \quad &= \quad ||y|| \\
\forall k \le K \qquad y_k \quad &= \quad \rho \sin \vartheta_{k-1} \prod_{h=k}^{K-1} \cos \vartheta_h
\end{aligned}
$$

- **Only need to decide** $s_k = \sin \vartheta_k$ **and** $c_k = \cos \vartheta_k$
- **Get polynomial program in** s, c
- **Numerically more challenging to solve**
- *But maybe useful for bounds?*

SDP relaxation of Euclidean distances

- **Linearization of scalar products**

$$\forall i,j \le n \qquad x_i \cdot x_j \longrightarrow X_{ij}$$

where X is an $n \times n$ symmetric matrix

- $\|x_i\|_2^2 = x_i \cdot x_i = X_{ii}$
- $\|x_i - x_j\|_2^2 = \|x_i\|_2^2 + \|x_j\|_2^2 - 2x_i \cdot x_j = X_{ii} + X_{jj} - 2X_{ij}$
- $X = xx^\top \Rightarrow X - xx^\top = 0$ **makes linearization exact**
- **Relaxation:**

$$X - xx^\top \succeq 0 \Rightarrow \mathsf{Schur}(X,x) = \begin{pmatrix} I_K & x^\top \\ x & X \end{pmatrix} \succeq 0$$

SDP relaxation of binary constraints

- $\forall i \le n \qquad \alpha_i \in \{0,1\} \Leftrightarrow \alpha_i^2 = \alpha_i$
- **Let A be an $n \times n$ symmetric matrix**
- **Linearize $\alpha_i\alpha_j$ by A_{ij} (hence α_i^2 by A_{ii})**
- $A = \alpha\alpha^\top$ **makes linearization exact**
- **Relaxation:** $\mathsf{Schur}(A, \alpha) \succeq 0$

SDP relaxation of [MMS95]

$$
\left.\begin{array}{rrcl}
\max & \sum_{i=1}^{n} \alpha_i & & \\
\forall i \le n & X_{ii} & = & \alpha_i \\
\forall i < j \le n & X_{ii} + X_{jj} - 2X_{ij} & \ge & A_{ij} \\
\forall i \le n & A_{ii} & = & \alpha_i \\
\forall i < j \le n & A_{ij} & \le & \alpha_j \\
\forall i < j \le n & A_{ij} & \le & \alpha_i \\
\forall i < j \le n & A_{ij} & \ge & \alpha_i + \alpha_j - 1 \\
 & \mathsf{Schur}(X, x) & \succeq & 0 \\
 & \mathsf{Schur}(A, \alpha) & \succeq & 0 \\
\forall i \le n & x_i & \in & [-1, 1]^K \\
 & \alpha & \in & [0, 1]^n \\
 & X & \in & [-1, 1]^{n^2} \\
 & A & \in & [0, 1]^{n^2}
\end{array}\right\}
$$

Python, PICOS and Mosek

- bound always equal to n
- prominent failure :-(
- **Why?**
  - *can combine inequalities to remove A from SDP*
  - *integrality of α completely lost*

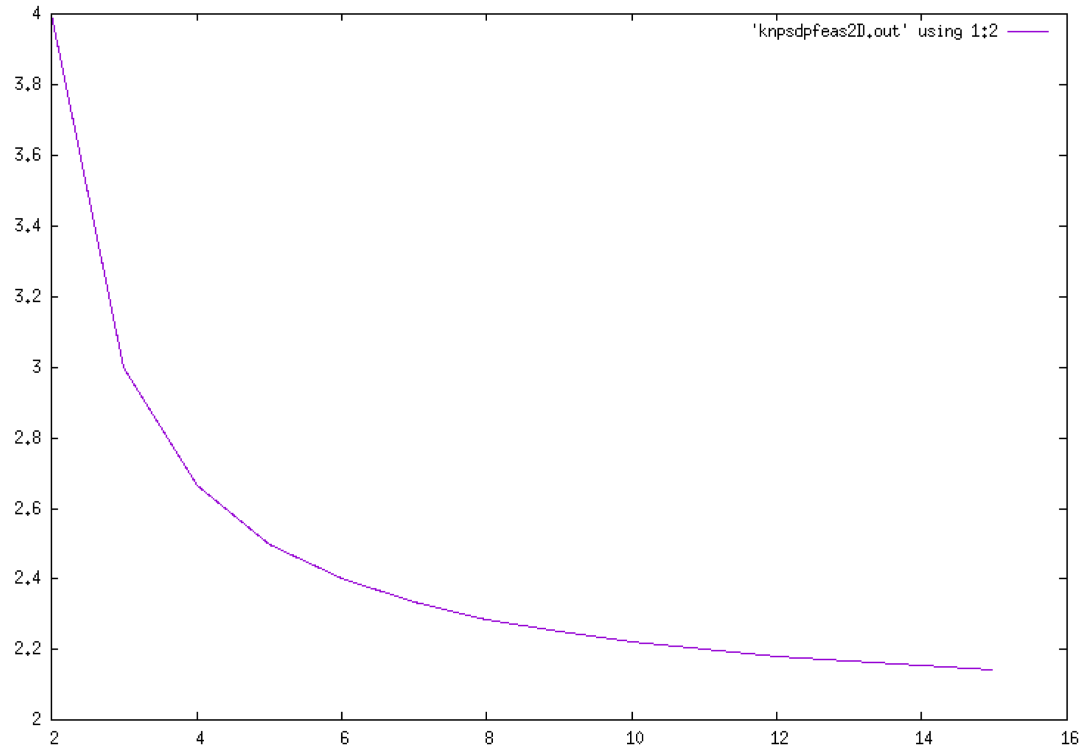
SDP relaxation of [KBLM07]

$$\left.\begin{array}{rrcl}
\max & \alpha & & \\
\forall i \le n & X_{ii} & = & 1 \\
\forall i < j \le n & X_{ii} + X_{jj} - 2X_{ij} & \ge & \alpha \\
& X & \in & [-1,1]^{n^2} \\
& X & \succeq & 0 \\
& \alpha & \ge & 0
\end{array}\right\}$$

Python, PICOS and Mosek

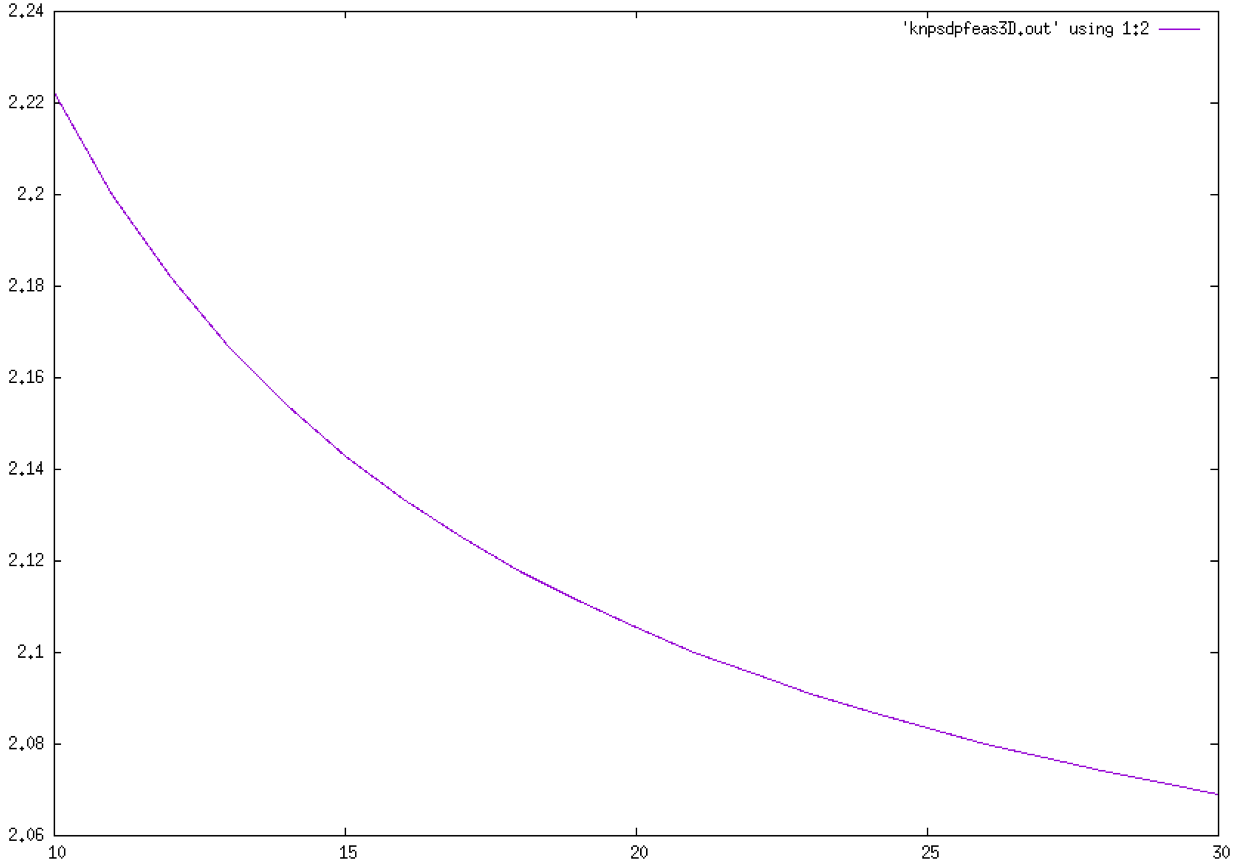
With $K = 2$

n	α^*
2	4.00
3	3.00
4	2.66
5	2.50
6	2.40
7	2.33
8	2.28
9	2.25
10	2.22
11	2.20
12	2.18
13	2.16
14	2.15
15	2.14



Python, PICOS and Mosek

With $K = 3$



Enforces some separation between "relaxed vectors"

An SDP-based heuristic

1. $X^* \in \mathbb{R}^{n^2}$: SDP relaxation solution of [KBLM07]
2. Perform *Principal Component Analysis* (PCA), get $\bar{x} \in \mathbb{R}^{nK}$
   - *concatenate K eigenvectors $\in \mathbb{R}^n$ corresponding to K largest eigenvalues*
3. Use $\bar{x}$ as starting point for local NLP solver on [KBLM07]

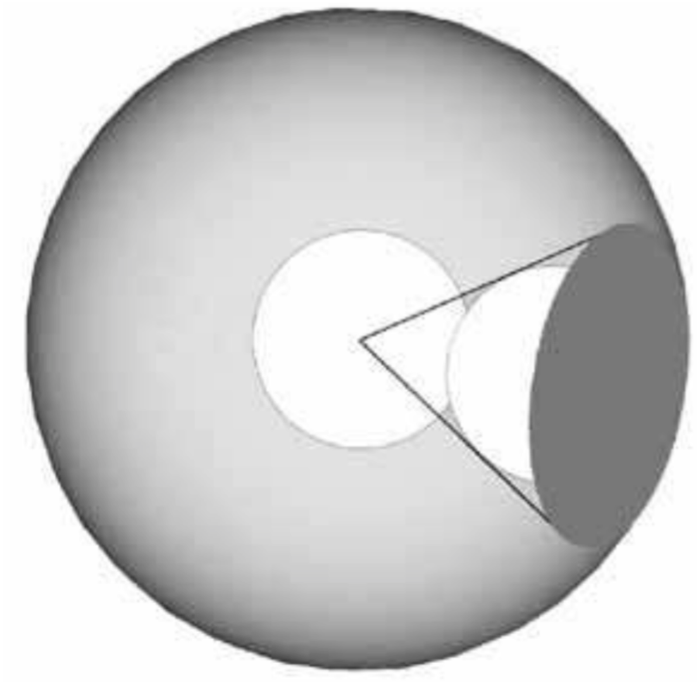
Python, PICOS, Mosek + AMPL, IPOPT

- $n = 6, K = 2$: **OK, 0.02s**
- $n = 12, K = 3$: **OK, 0.02s**
- $n = 24, K = 4$: **4% error, 0.32s**
- $n = 40, K = 5$: **5% error, 1.57s**
- $n = 72, K = 6$: **7% error, 12.26s**

Surface upper bound

Szpiro 2003, Gregory 1694

Consider a $\mathsf{kn}(3)$ configuration inscribed into a super-sphere of radius 3. Imagine a lamp at the centre of the central sphere that casts shadows of the surrounding balls onto the inside surface of the super-sphere. Each shadow has a surface area of 7.6; the total surface of the super-ball is 113.1. So $\frac{113.1}{7.6} = 14.9$ is an upper bound to $\mathsf{kn}(3)$.



At end of XVII century, yielded Newton/Gregory dispute

Another upper bound

Thm.

Let: $\mathcal{C}_z = \{x_i \in \mathbb{S}^{K-1} \mid i \leq n \wedge \forall j \neq i \ (x_i \cdot x_j \leq z)\}$; $c_0 > 0$; $f : [-1, 1] \to \mathbb{R}$ **s.t.:**
(i) $\sum_{i,j \leq n} f(x_i \cdot x_j) \geq 0$ **(ii)** $f(t) + c_0 \leq 0$ **for** $t \in [-1, z]$ **(iii)** $f(1) + c_0 \leq 1$
Then $n \leq \frac{1}{c_0}$

(**[Delsarte 1977]; [Pfender 2006]**)
Let $g(t) = f(t) + c_0$

$$\begin{aligned} n^2 c_0 &\leq n^2 c_0 + \sum_{i,j \leq n} f(x_i \cdot x_j) && \text{by (i)} \\ &= \sum_{i,j \leq n} (f(x_i \cdot x_j) + c_0) = \sum_{i,j \leq n} g(x_i \cdot x_j) \\ &\leq \sum_{i \leq n} g(x_i \cdot x_i) && \text{since } g(t) \leq 0 \text{ for } t \leq z \text{ and } x_i \in \mathcal{C}_z \text{ for } i \leq n \\ &= n g(1) && \text{since } \|x_i\|_2 = 1 \text{ for } i \leq n \\ &\leq n && \text{since } g(1) \leq 1. \end{aligned}$$

The *Linear Programming Bound*

- Condition (i) of Theorem valid for conic combinations of suitable functions $\mathcal{F} = \{f_1, \ldots, f_H\}$:

$$f(t) = \sum_{h \le H} c_h f_h(t) \quad \text{for some } c_h \ge 0$$

- Let $T = \{t_i \mid i \le s \wedge t_1 = -1 \wedge t_s = z \wedge \forall i < j \ (t_i < t_j)\}$, get LP:

$$\left.\begin{array}{rrcll}
\max\limits_{c \in \mathbb{R}^{K+1}} & c_0 & & & n = 1/c_0 \text{ smallest} \\
\forall\, t \in T & \sum\limits_{1 \le h \le H} c_h f_h(t) + c_0 & \le & 0 & \text{(ii)} \\
 & \sum\limits_{1 \le h \le H} c_h f_h(1) + c_0 & \le & 1 & \text{(iii)} \\
\forall\, 1 \le h \le H & c_h & \ge & 0 & \text{(conic comb.)}
\end{array}\right\}$$

- E.g. $\mathcal{F} =$ Gegenbauer polynomials [Delsarte 1977]
- $T \subseteq [-1, z]$, *don't know how to solve infinite LPs so we discretize it*

Some results

- ***Gegenbauer polynomials*** G_h^γ **(recursive definition):**

$$\begin{aligned} G_0^\gamma(t) &= 1, \quad G_1^\gamma(t) = 2\gamma t, \\ \forall h > 1 \; hG_h^\gamma(t) &= 2t(h+\gamma-1)G_{h-1}^\gamma(t) - (h-2\gamma-2)G_{h-2}^\gamma(t) \end{aligned}$$

(all normalized so $G_h^\gamma(1) = 1$)

- Special case $G_h^\gamma = P_h^{\gamma,\gamma}$ of *Jacobi polynomials*:

$$P_h^{\alpha,\beta} = \frac{1}{2^h} \sum_{i=0}^{h} \binom{h+\alpha}{i} \binom{h+\beta}{h-1} (t+1)^i (t-1)^{h-i}$$

- [Delsarte 1977, Odlyzko & Sloane 1998]

  $\mathsf{kn}(3) \leq 12$, $\mathsf{kn}(4) \leq 25$, $\mathsf{kn}(5) \leq 46$, $\mathsf{kn}(8) \leq 240$, $\mathsf{kn}(24) \leq 196560$

- *Used to prove the "Twelve spheres theorem" ($\mathsf{kn}(3) = 12$)*

- **My test: works for $K > 4$, couldn't make it work for $K = 3$**

Where does K appear in the LP bound?

- $\mathcal{F}$ **containing Gegenbauer polynomials**
- **In** $G_h^\gamma(t)$**,** $\gamma = \frac{K-3}{2}$
- K **determined by** ***lowest*** γ **appearing in** $\mathcal{F}$
- **E.g.** $\mathcal{F} = \{G_h^1(t), G_h^{1.5}(t) \mid h \leq 10\}$ **yields bound** $25.5581 \geq \mathsf{kn}(4) = 24$