Systems of $\{$-1,0,1$\}$-vectors

Bashaeva Lina

Scientific advisor: Andrei M. Raigorodskii, PhD

Higher School of Econmics
Faculty of Mathematics

Contents

Introduction

Definition

A **graph** (sometimes called undirected graph for distinguishing from a directed graph, or simple graph for distinguishing from a multigraph) is a pair $G = (V, E)$, where V is a set whose elements are called vertices (singular: vertex), and E is a set of paired vertices, whose elements are called edges (sometimes links or lines).

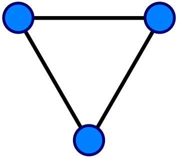

Figure: A graph with three vertices and three edges.

Introduction

Definition

$G = (V, E)$ is n-dimensional **distance graph** (or graph of distances), if:

$$V \subseteq \mathbb{R}_n,$$
$$E \subseteq \left\{\{\overline{x}, \overline{y}\} : \overline{x}, \overline{y} \in V, |\overline{x} - \overline{y}| = a\right\}, \quad a > 0,$$

Definition

An **independent set** of a graph G is a subset of its vertices , no two of which are adjacent:

$$G_0 = (V_0, E_0) : \left\{V_0 \subseteq V, E_0 = \varnothing\right\},$$

A power of the biggest independent set of a graph G is called the **independent number** and usually denoted by $\alpha(G)$.

Motivation

What is maximum power of a set of n-dimensional $\{$-1,0,1$\}$ - vectors, the pairwise scalar product of which is not equal to t?

1. **Extremal combinatorics:** the independent number of a distance graph, the vertices of which are n-dimensional $\{$-1,0,1$\}$ -vectors with k nonzero coordinates, and the edges are drawn between those vertices whose scalar product is equal to a given number t (that is, which are at the distance $\sqrt{(2(k-t))}$)
2. **Coding theory:** the problem of codes with forbidden distance. $\{$-1,0,1$\}$ -vectors can be considered words of an equilibrium ternary code of weight k. Then the condition for the absence of a scalar product equal to t is equivalent to the condition that there is no Hamming distance of the value $2(k-t)$.

Formulation of the problem

We consider the graphs $G_n = (V_n, E_n)$, for which $n = 2k$, $k \in \mathbb{N}$, $t = 0$:

$$V_n = \left\{ \mathbf{x} = (x_1, \ldots, x_n) : x_i \in \{-1, 0, 1\}, |\{i : x_i = \pm 1\}| = |\{i : x_i = 0\}| = k \right\}$$

$$\mathrm{E}_n = \left\{ \{\mathbf{x},\mathbf{y}\} : \mathbf{x},\mathbf{y} \in V_n, |\mathbf{x}\text{-}\mathbf{y}| = \sqrt{2k} \right\}.$$

1. an extremely important role in the problems of Borsuk and Nelson-Erdös-Hadwiger
2. the question of finding the independence number of G_n is still open

Our aim is to improve the existing lower bounds for $\alpha(G_n)$, given in the diploma work of V. Moskva in 2010 for even k.

Preliminaries: Moskva's construction

$$\underbrace{i_1, i_2, \ldots, i_{2m+1}}_{1}, \underbrace{j_1, \ldots, j_{2k-2m-1}}_{2}$$

Let $m \in \{1, \ldots, k-1\}$ (the parameter by which the maximum will be taken). We divide the set of coordinate positions into two parts with cardinalities $2m+1$ and $2k-2m-1$. For each $i \in \{1, \ldots, m+1\}$ consider all vectors for which in the first part there are exactly $m+i$ nonzero coordinates, and in the second, respectively, exactly $k-m-i$ nonzero coordinates.
We introduce a restriction on the number of negative coordinates in each part: if in the first part there are A and in the second - B then the inequality $i - 2A - B \geq 1$ must be satisfied.
Add to the construction all vectors obtained from those already constructed by multiplying by -1. It is clear that no orthogonal pairs will arise.

Thus, the cardinality of the optimal construction is

$$2\max_{m}\sum_{i=1}^{m+1}\sum_{A,B:i-2A-B\geq 1}\binom{2m+1}{m+i}\binom{m+i}{A}\binom{2k-2m-1}{k-m-i}\binom{k-m-i}{B}. \quad (1)$$

My construction of last year's term paper

Let $k_1 \leq k$ be a number of ones for each vector in construction, $k_2 = k - k_1$ – number of minus ones, $q \leq n, t, r : t + 2r \leq q$ are fixed parameters. Split a set of coordinate positions $N = \{1, \dots, n\}$ into two subsets: $Q = \{1, \dots, q\}$ and $N \setminus Q = \{q + 1, \dots, n\}$. Then split the subset Q into two parts: the first sub-part of power $t + 2r$ and the second of power $q - t - 2r$. Each vector from the construction has $t + i, 1 \leq i \leq r$ ones in the first sub-part and $k_1 - t - i$ ones in the second sub-part. Each vector has j negative coordinates in Q and $k_2 - j$ minus ones in $N \setminus Q$. A restriction is imposed on j: for any two vectors, the power of the intersection along the coordinates of opposite signs in Q must be less than the power of the intersection along the coordinates of the same sign in N. Then, their scalar product is positive.

My construction of last year's term paper

Thus, the cardinality of this construction is

$$\sum_{i=r}^{\min\{2r,k-t\}} \binom{t+2r}{t+i}\binom{q-t-2r}{k-t-i} \sum_{j=min\{0,k_2-(n-q)\}}^{l} \binom{q-k_1}{j} \cdot \binom{n-q}{k_2-j},$$

where

$$l = \frac{[3t-n+q-1]}{2} + k_2 - k_1.$$

A power of the optimal construction is a maximum by parameters

$0 \leqslant r \leqslant k_1 - t, 0 \leqslant t \leqslant q,$

$0 \leqslant q \leqslant n, 0 \leqslant k_1 \leqslant k.$

Main results

- probably lower estimates are good for fixed k_1 and k_2, but for unfixed k_1,k_2 and $n = 2k$ (we checked $k = 1, \ldots, 14$ using Python math package) lower estimates are much worse than known
- after some observations we came up with modified construction

New construction

Let $n = 2k$, as before.

We split the set of coordinate positions into three parts: the first $t_1 + 2r_1$ coordinate positions are going into the first part, the next $t_2 + 2r_2$ positions – into the second part and the last $n - (t_1 + t_2 + 2(r_1 + r_2))$ – into the third part. Each vector has at least $t_1 + i, i = r_1, \ldots, 2r_1$ nonzero coordinates in the first part and $t_2 + j, j = r_2, \ldots, 2r_2$ nonzero coordinates in the second part. Let's call such vectors as vectors of "i, j" $-$level.

New construction

We limit the number of negative coordinates so that the scalar product of any two vectors does not equal zero, like in Moskva's construction: if a vector of "i, j" $-$level has A_1 minus ones in the first part, A_2 – in the second part and B – in the third part, then the following inequality holds:

$$2(A_1 + A_2) + B \leq i + j + t_1 + t_2 - 1 = s_{ij}.$$

New construction

So, a power of the new construction is

$$\sum_{i=r_1}^{2r_1} \binom{t_1+2r_1}{t_1+i} \sum_{j=r_2}^{2r_2} \binom{t_2+2r_2}{t_2+j} \binom{2k-(t_1+t_2+2r_1+2r_2)}{k-(i+j+t_1+t_2)} \sum_{J=0}^{s_{ij}} \binom{k-s_{ij}+1}{J} \sum_{i_1=0}^{[\frac{s_{ij}-J}{2}]} \binom{j}{i_1} \tag{2}$$

A power of the optimal construction is a maximum by parameters r_1, r_2, t_1, t_2.

Observations and preliminary results

- Moskva's construction turned out to be a special case of this construction for $t_1 = 1, r_1 = m*, t_2 = 0, r_2 = 0$
- we also used Python to find the optimal values for r_1 and r_2, putting $t_1 = 1$ and $t_2 = 0$ to get the greatest value of (2) .
- for $n = 1, \dots, 28$ we got the same values as in (1), in all cases it turned out that the parameters $r_1 = m*, r_2 = 0$ are optimal.

Future research

- construction (1) can be considered a 2-*part* construction, and (2) – a 3-*part* for $t_2 + r_2 \neq 0$
- one can consider $d - part$ constructions for any $d = 1, ..., n$.

Current question: is it true that for any n a two-part construction gives a better lower bound than d-*part* for $d > 2$, or are there n such that d-*part* for $d > 2$ give a better lower bound, than a 2-*part* construction?

Appendices

Table: Lower and upper bounds for the independence numbers G_{2k}

n	$\alpha \geq$	$\alpha \leq$
2	2	2
4	6	6
6	32	32
8	80	80
10	512	930
12	1428	5000
14	8192	30000
16	29700	312426
18	134420	1807920
20	635206	–
22	2983344	61258430
24	13929188	–

*A dash means no reasonable estimate.

Appendices

Figure: A code for finding optimal parameters r_1, r_2 with $t_1 = 1, t_2 = 0$

[5]

```
for m in range(12):
    max_number = 0
    for r1 in range(2*m+1):
        for r2 in range(2*m+1):
            number=2*sum(
                binom(1 + 2 * r1, r1 + 1 + i) * sum(
                    binom(2 * r2, r2 + j) *
                    binom(2 * (m - (1 + r1 + r2))+1, m -
                                        (1 + r1 + r2 + i + j)) *
                    sum(
                        binom(m - (1 + r1 + r2 + i + j), h) * sum(
                            binom(r2+j, i1) * sum(
                                binom(r1+1+i, j1)
                                for j1 in range((i + j - h) // 2 - i1 + 1)
                            ) for i1 in range((i + j - h) // 2 + 1)
                        ) for h in range(i + j + 1)
                    ) for j in range(min(r2+1,m-(r1+1+i)-(r2+1)+1))
                ) for i in range(min(r1 + 1,m+1)))
            if number > max_number:
                print('r1=', r1,'r2=',r2)
                max_number = number

    print(max_number)
```

Appendices

Figure: Some given estimates by the code

```
n= 0 Optimal estimate= 0
n= 2 Optimal estimate= 0
r1= 0 r2= 0
n= 4 Optimal estimate= 6.0
r1= 0 r2= 0
n= 6 Optimal estimate= 20.0
r1= 0 r2= 0
r1= 1 r2= 0
n= 8 Optimal estimate= 80.0
r1= 0 r2= 0
r1= 0 r2= 1
r1= 1 r2= 0
n= 10 Optimal estimate= 336.0
r1= 0 r2= 0
r1= 0 r2= 1
r1= 1 r2= 0
n= 12 Optimal estimate= 1428.0
r1= 0 r2= 0
r1= 0 r2= 1
r1= 1 r2= 0
r1= 2 r2= 0
n= 14 Optimal estimate= 6528.0
r1= 0 r2= 0
r1= 0 r2= 1
r1= 0 r2= 2
r1= 1 r2= 0
r1= 1 r2= 1
r1= 2 r2= 0
n= 16 Optimal estimate= 29700.0
```

*Jumping to a line with new parameters means that the construction with parameters on a next line gives a larger value than with parameters on the previous line.

Thanks for attention:)