COMBINATORICA 1 (3) (1981) 225—231

EVERY LARGE SET OF EQUIDISTANT (0, +1, −1)-VECTORS FORMS A SUNFLOWER

by

M. DEZA and P. FRANKL

C. N. R. S. Paris, 75007
15 Quai Anatole France

Received 16 July 1980

A theorem of Deza asserts that if $H_1, \ldots, H_m$ are s-sets any pair of which intersects in exactly d elements and if $m \geqq s^2 - s + 2$, then the H_i form a Δ-system, i.e. $\left|\bigcap_{i=1}^{m} H_i\right| = d$. In other words, every large equidistant (0, 1)-code of constant weight is trivial. We give a (0, +1, −1) analogue of this theorem.

1. Introduction

An equidistant (0, 1) code is a set $A = \{a_1, \ldots, a_m\}$ of (0, 1) vectors in $\mathbf{R}^n$ each having the same number s of non-zero entries such that the scalar products (a_i, a_j) have the same value d $(1 \leqq i < j \leqq m)$. Deza [2] proved

Theorem 1.1. *If $s = 2d$, $m > d^2 - d + 1$, then for an equidistant code A we can find d positions $1 \leqq i_1 < \ldots < i_d \leqq n$ such that all the vectors in A have 1 in these positions.*

A slight modification of the argument in [2] gives

Theorem 1.1′. *If $A = \{a_1, \ldots, a_m\}$ is an equidistant code with*

$$m > \max\{d+2, (s-d)^2 + (s-d) + 1\}$$

then we can find d positions $1 \leqq i_1 < \ldots < i_d \leqq n$ such that all the vectors in A have 1 in these positions.

For large n the bound is sharp if and only if $(d+2) \geqq (s-d)^2 + (s-d) + 1$ or there exists a projective plane of order $s-d$ (van Lint [6]). For some more results on this topic see the survey paper [3]. A. J. Hoffman [5] asked how this results extend to (0, +1, −1)-vectors, in view of applications to eigenvalues of directed graphs.

The main result of this paper is

Theorem 1.2. *Suppose $B = \{b_1, b_2, \ldots, b_m\}$ is a set of (0, +1, −1)-vectors in $\mathbf{R}^n$. Suppose that for some integers s, d with $s > d \geqq 1$ we have*

$$(b_i, b_i) = s, \ (b_i, b_j) = d \quad \textit{for every} \quad 1 \leqq i \neq j \leqq m.$$

AMS subject classification (1980): 05 C 65; 05 C 35, 05 B 25

If $m > \max\{(s-d)^2+(s-d)+1,\ (s-d)(d+2)\}$, *then we can find d positions* $1 \leqq i_1 < i_2 < \ldots < i_d \leqq n$ *such that for every* $1 \leqq j \leqq d$ *all the vectors have the same non-zero* i_j*'th entry.*

The case $d<0$ is much simpler. The answer is

Theorem 1.3 [1]. *Suppose* $B=\{b_1, b_2, \ldots, b_m\}$ *is as in Theorem* 1.2, $s \geqq 1$ *but* $d \leqq -1$. *Then* $m \leqq [1-s/d]$, *and this is best possible for large n.*

In fact Delsarte, Goethals, Seidel prove this theorem in a much more general setting, using Gegenbauer polynomials. (They consider the case of vectors in $\mathbf{R}^n$ of equal length with given set A of values of pairwise scalar products. The special case $A=\{\pm d\}$ is exactly the case of a set of equiangular lines, i.e. equidistant set of points in the elliptic space $\mathbf{E}^{n-1}$.) We give an elementary proof.

In the case $d=0$ obviously $m \leqq n$. The case of equality corresponds to the so-called weighing matrices, $W(n, s)$. In the case $s=n$ a weighing matrix is just a Hadamard matrix of order n. For more information about weighing matrices confer [4].

2. The proof of Theorem 1.3

Set $v=b_1+b_2+\ldots+b_m$. Then $(v, v) \geqq 0$, or, equivalently

(1) $$ms+m(m-1)d \geqq 0.$$

Dividing by $m(-d)$ and rearranging we obtain $m \leqq 1+\frac{s}{-d}$. As m is an integer we also have $m \leqq [1-(s/d)]$, proving (1).

To show that (1) is best possible we have to construct $m=[1-(s/d)]$ $(0, +1, -1)$-vectors $b_1, b_2, \ldots, b_m$ such that $(b_i, b_i)=s$ and $(b_i, b_j)=d$ for $1 \leqq i, j \leqq m$, $i \neq j$. We represent these vectors as the rows of a matrix M.

Every column of M will have at most 2 non-zero entries, more exactly for every (i, j), $1 \leqq i < j \leqq m$ M contains $-d$ columns with 1 in the i'th, -1 in the j'th positions, and for every $1 \leqq i \leqq m$ it contains $s+d[s/-d]$ columns with $+1$ in the i'th position and zeros elsewhere, and it contains no other columns. Thus M is defined up to the order of columns, but this does not affect the scalar products of the rows. Every row contains $(m-1)d+s+d[s/-d]=s$ non-zero entries, and for any two different rows there are $-d$ columns where none of them has zero. Moreover these two non-zero entries have opposite signs, giving scalar product d. ∎

Remark 2.1. The proof of Theorem 1.3 shows that if $m=1-(s/d)$ then necessarily $b_1+b_2+\ldots+b_m$ is the all-zero vector.

3. Some lemmas

Definition 3.1. For $1 \leqq i \leqq n$, and a collection of $(0, 1, -1)$-vectors $B=\{b_1, b_2, \ldots, b_m\}$ let $q_i^0(B)$, $q_i^+(B)$, $q_i^-(B)$ denote the number of vectors which have 0, $+1$, -1, respectively, in the i'th coordinate. If it causes no confusion we simply write q_i^0, q_i^+, q_i^- or q^0, q^+, q^- if the current value of i is clear from the context.

Lemma 3.2. *Suppose* $B=\{b_1, b_2, \dots, b_m\}$ *is as in Theorem* 2, *then*

$$(q^0+q^-)m(s-d) \geqq (2q^-+q^0)^2q^+. \tag{2}$$

Corollary 3.3. *With the same notation as in Lemma* 3.2 *we have*

$$(q^0+q^+)m(s-d) \geqq (2q^++q^0)^2q^-. \tag{3}$$

Proofs of the lemma and the corollary. Let us define δ_j for $1\leqq j\leqq m$ such that $\delta_j = q^0+q^-$ if b_j has $+1$ in the i'th position and $\delta_j=-q^+$ otherwise. Let b_j' be the vector which agrees with b_j in every position except possibly in the i'th position, where it has 0. Set $v=\sum_{1\leqq j\leqq m}\delta_j b_j'$. Let us expand the inequality $(v, v)\geqq 0$. We obtain

$$\begin{aligned}(q^0+q^-)^2q^+(s-1)+(q^0+q^-)^2q^+(q^+-1)(d-1)+(q^+)^2q^-(s-1)+\\ +(q^+)^2q^0s+(q^+)^2q^-(q^--1)(d-1)+(q^+)^22q^-q^0d+\\ +(q^+)^2q^0(q^0-1)d-2(q^+)^2(q^-+q^0)q^0d-\\ -2(q^+)^2(q^-+q^0)q^-(d+1)\geqq 0.\end{aligned}$$

After rearranging and dividing by q^+(if $q^+=0$ the lemma holds trivially) we obtain

$$(q^0+q^-)(q^++q^0+q^-)(s-d)-(2q^-+q^0)^2q^+\geqq 0.$$

As $m=q^++q^0+q^-$ the statement of the lemma is proved. Now the corollary follows applying the lemma to $B^-=\{-b_1, -b_2, \dots, -b_m\}$. ▮

Remark 3.4. Actually Lemma 3.2 is a direct consequence of the following statement: Let $V=\{v_1, v_2, \dots, v_n\}$ be a set of real vectors in $\mathbf{R}^u$ such that all the Euclidean distances $\|v_i-v_j\|$ are the same, say f. Suppose there are q_x vectors which have x in the i'th position $(1\leqq i\leqq n)$. Then we have the following inequality for the harmonical mean

$$\operatorname{Harm}(q_x, q_y) = \frac{2}{\dfrac{1}{q_x}+\dfrac{1}{q_y}} \leqq \left(\frac{f}{x-y}\right)^2$$

for arbitrary real numbers x, y. (The proof uses a negative type inequality (see [7]) for the square of Euclidean distance.) In our case $\|v_i\|=s$ and so $f^2=2s-2d$. Thus $(q_x+q_y)/q_xq_y\geqq(x-y)^2/(s-d)$. In the special case of (0, 1) vectors it becomes $q_{+1}(m-q_{+1})\leqq(s-d)m$, which is the inequality of [2, Lemma 3.1] crucial for the proof of Theorem 1.1.

Lemma 3.5. *Suppose* $B=\{b_1, b_2, \dots, b_m\}$ *is as in Theorem* 2 *and that*

$$m > \max\{(s-d)^2+(s-d)+1,\ (s-d)(d+2)\},$$

then for every $1\leqq i\leqq n$ *either*

$$\max(q^-, q^+)\geqq m-(s-d+1) \tag{4}$$

or

$$\max(q^-, q^+)\leqq (s-d+1), \tag{5}$$

or $d=1$, $s\leqq 3$.

Proof. By symmetry reasons we may assume $q^- \leqq q^+$. Suppose the statement of the Lemma is not true i.e. $s-d+2 \leqq q^+ \leqq m-s+d-2$. Using lemma 3.2 we deduce

$$m(s-d) \geqq q^+(2q^- + q^0)^2/(q^- + q^0) \geqq q^+(q^0 + q^-) \geqq (s-d+2)(m-s+d-2).$$

Rearranging and using the fact that m is an integer we deduce

$$m \leqq \left[\frac{1}{2}(s-d+2)^2\right].$$

Now for $s-d>2$

$$(s-d+1)(s-d)+1 > \frac{1}{2}(s-d+2)^2,$$

while, for $s-d \leqq 2$, $(d+2)(s-d) \geqq \left[\frac{1}{2}(s-d+2)^2\right]$, except for $d=1$, $s \leqq 3$. ▮

Definition 3.6. We say i is *light (heavy)* if $1 \leqq i \leqq n$ and (5) ((4), respectively) holds in Lemma 3.5.

Definition 3.7. We say b_j *contributes* to the heavy position i if its i'th entry is non-zero and of the same sign as the majority of non-zero i'th entries.

Lemma 3.8. *Suppose $B=\{b_1, \ldots, b_m\}$ is as in Theorem 1.2 and that*

$$m > \max\{(s-d)^2+(s-d)+1,\ (d+2)(s-d)\},$$

then every b_j contributes to at least d heavy positions, or $d=1$, $s \leqq 3$.

Proof. Arguing indirectly we may assume that, for example, b_1 contributes to only $d-t$ heavy columns with $1 \leqq t \leqq d$. As $(b_i, b_1)=d$ for $2 \leqq i \leqq m$, b_i agrees with b_1 in at least t of the remaining $s-d+t$ non-zero positions of b_1. Thus, using Lemma 3.5 $(s-d+t)(s-d) \geqq (m-1)t$, yielding $m \leqq (s-d)^2+(s-d)+1$, a contradiction. ▮

Lemma 3.9. *Suppose $B=\{b_1, b_2, \ldots, b_m\}$ is as in Theorem 2 and that*

$$m > \max\{(s-d)^2+(s-d)+1,\ (d+2)(s-d)\},$$

then there are at most d heavy positions unless $d=1$, $s \leqq 3$.

Proof. Suppose the contrary and let $i_1, i_2, \ldots, i_{d+1}$ be heavy positions. Changing the signs of all the entries in some of these columns we may assume $q_{i_t}^+ \geqq m-s+d$ for $1 \leqq t \leqq d+1$.

For $1 \leqq t \leqq d+1$ let A_t be the set of integers j, $1 \leqq j \leqq m$, for which b_j has not $+1$ in the i_t'th position. Let us set $A_0 = \{1, 2, \ldots, m\} - \bigcup_{t=1}^{d+1} A_t$.

We break the proof into three propositions.

Proposition 3.10. $|A_0| \leqq s-d$.

Proof. Suppose the contrary and let $j_1, j_2, \ldots, j_{s-d+1} \in A_0$. For $1 \leqq t \leqq s-d+1$ let a_t be the vector which agrees with b_{j_t} except that it has 0 in the i_l'th position for $l=1, \ldots, d+1$. Then $(a_t, a_t)=s-d-1$ and $(a_t, a_{t'})=-1$ for $1 \leqq t \neq t' \leqq s-d+1$, contradicting Theorem 1.3. ▮

From the above, Lemma 3.9 and Theorem 1.2 easily follow for

$$m > \max\{(s-d)^2+(s-d)+1,\ (d+2)(s-d+1)\}.$$

In order to get the sharp result, we need more technical propositions.

Proposition 3.11. If $|A_0|=s-d$ then the A_s's are pairwise disjoint.

Proof. By Remark 2.1, the fact $|A_0|=s-d$ implies $\sum_{1\leqq l\leqq s-d} a_l=0$. Let us choose $1\leqq j\leqq m$ such that it is contained in two different A_t's. Then $(b_j, b_{i_l})=d$ implies $(b_j, a_l)\geqq 1$ for $1\leqq l\leqq s-d$. But this leads to $0=(b_j, \sum a_l)\geqq s-d$, a contradiction. Now Proposition 3.10 yields $\max_{1\leqq t\leqq d+1} |A_t|>s-d$, thus by Lemma 3.5 we have $\max_{1\leqq t\leqq d+1} |A_t|=s-d+1$. By symmetry we assume $|A_1|=s-d+1$. ∎

Proposition 3.12. For $2\leqq t\leqq d+1$ we have

$$|A_t-A_1| \leqq s-d;$$

moreover, if we have equality then $A_t\cap A_1\neq\emptyset$.

Proof. The contrary would mean that for some $2\leqq t\leqq d+1$ we have $s-d\leqq|A_t|\leqq \leqq s-d+1$, and $A_t\cap A_1=\emptyset$ holds. For $j\in(A_1\cup A_t)$ let a_j denote the vector we obtain from b_j by putting zero into the i_1 and i_t'th position.

Let us set $w=|A_t| \sum_{j\in A_1} a_j-|A_1| \sum_{j\in A_t} a_j$. Expanding the inequality $(w, w)\geqq 0$, and using $(a_j, a_{j'})\geqq d$ for $j\in A_1$, $j'\in A_t$, $(a_j, a_{j'})\leqq d-1$ for $j\neq j'\in A_1$ or $j\neq j'\in A_t$, and $(a_j, a_j)\leqq s-1$ for $j\in(A_1\cup A_t)$ we obtain a contradiction. ∎

Now from the first part of Proposition 3.12 we deduce

$$\Big|\bigcup_{1\leqq j\leqq d+1} A_j\Big| \leqq |A_1|+\sum_{2\leqq j\leqq d+1} |A_j-A_1| \leqq (s-d)(d+1)+1.$$

Thus $|A_0|\geqq s-d$. Using Propositions 3.10 and 3.11 we deduce $|A_0|=s-d$ and the sets A_t are pairwise disjoint. Hence by the second part of Proposition 3.12 we have $|A_t-A_1|\leqq s-d-1$, yielding $|A_0|\geqq s-d+1$, in contradiction with Proposition 3.10. This proves Lemma 3.9. ∎

4. The proof of Theorem 1.2

Suppose first $(s, d)\neq(2, 1)$ or $(3, 1)$. Then by Lemma 3.9 there are at most d heavy positions and by Lemma 3.8 every vector contributes to at least d, i.e. to each of them. Hence there are d positions in which all the vectors have the same non-zero entry. This concludes the proof for this case.

If $s=2$, $d=1$ then by symmetry we assume $b_1=(1, 1, 0, 0, \ldots, 0)$. As $(b_i, b_1)=1$ for $i=2, 3$, again by symmetry we may assume $b_2=(1, 0, 1, 0, \ldots, 0)$, $b_3= =(1, 0, 0, 1, 0, \ldots, 0)$. But now $(b_i, b_j)=1$ for $1\leqq i\leqq 3$, $4\leqq j\leqq m$ yields that all the b_j's have 1 in the first position.

For $s=3$, $d=1$ suppose first that there are two rows which have in some position $+1$ and -1, respectively. By symmetry we may assume it is the first position and the vectors are b_1, b_2. As $(b_1, b_2)=1$, $b_1=(1, 1, 1, 0, \ldots, 0)$, $b_2= =(-1, 1, 1, 0, \ldots, 0)$. For $i \geqq 3$ $(b_i, b_1)=1$, $(b_i, b_2)=1$ yields that b_i has $(0, 1, 0)$ or $(0, 0, 1)$ in the first 3 positions. As $m \geqq 8$ one of these possibilities occurs at least 3 times: say the first one and for b_3, b_4, b_5. If not all the vectors have 1 in the second position then we may assume $b_6=(0, 0, 1, 1, 1, 0, \ldots, 0)$. As $(b_j, b_6)=1$ for $j=3, 4, 5$ we may as assume b_3 and b_4 have both 1 in the fourth position, and consequently $b_3=(0, 1, 0, 1, 0, 1, 0, \ldots, 0)$, $b_4=(0, 1, 0, 1, 0, -1, 0, \ldots, 0)$. But then b_5 cannot have 1 in the fourth position, we may assume $b_5=(0, 1, 0, 0, 1, 0, 1, 0, \ldots, 0)$. Now the only possible choice for b_7 is $(0, 1, 0, 0, 1, 0, -1, 0, \ldots, 0)$, thus $m \leqq 7$, a contradiction.

If in every position all the non-zero entries are equal then, possibly changing their signs simultaneously, we may assume that all the b_i's are $(0, 1)$-vectors. We may suppose $b_1=(1, 1, 1, 0, \ldots, 0)$. As $(b_i, b_1)=1$ for $2 \leqq i \leqq 8$, there are at least 3 vectors which have 1 in the same of the first three positions. By symmetry we may assume

$$b_2 = (1, 0, 0, 1, 1, 0, \ldots, 0),\ b_3 = (1, 0, 0, 0, 0, 1, 1, 0, \ldots, 0),$$

$$b_4 = (1, 0, 0, 0, 0, 0, 0, 1, 1, 0, \ldots, 0).$$

Now $(b_i, b_j)=1$ for $i=1, 2, 3, 4$ and $4 \leqq j \leqq m$ gives that all the b_j's have 1 in the first position, which concludes the proof of Theorem 1.2. ∎

Remark 4.1. For large n the bound given by Theorem 1.2 is best possible. For $d<s/2$ we can take the incidence vectors of a projective space of order $s-d$ (if it exists) plus $d-1$ columns of entirely ones. For $d \geqq s/2$ we give the vectors as the rows of the following matrix N.

Let J_r be the r by r all ones matrix and I_r the r by r identity matrix. We define the first $d+2$ columns of N by $(s-d)$-fold repetition of each row of $J_{d+2}-I_{d+2}$. In order to construct the remaining columns of N, we use the matrix M constructed in the proof of Theorem 3 with the following parameters: the rows have weight $s-d-1$ and the scalar product of each pair of rows is equal to -1. We define the right half of N to be the Kronecker product of I_{s-d} by M, i.e. the block-diagonal matrix with M in each of the $s-d$ diagonal blocks.

Now N has $(s-d)(d+2)$ rows, every row contains $(d+1)+(s-d-1)=s$ non-zero entries and all the scalar products of different rows are equal to d. For $s=4, d=2$

$$N = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & -1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

References

[1] P. DELSARTE, J-M. GOETHALS and J. J. SEIDEL, Spherical codes and designs, *Geometriae Dedicata* **6** (1977) 363—368.

[2] M. DEZA, Une propriété extrémale des plans projectifs finis dans une classe de codes equidistants, *Discrete Math.* **6** (1973) 343—352.

[3] M. DEZA, R. C. MULLIN and S. A. VANSTONE, Recent results on (r, λ)-designs and related configurations, *to appear*.

[4] A. V. GERAMITA and J. SEBERRY, *Orthogonal designs*, Marcel Dekker Inc., New York and Basel, 1979.

[5] A. J. HOFFMAN, private communication.

[6] J. H. VAN LINT, A theorem on equidistant codes, *Discrete Math.* **6** (1973) 353—358.

[7] I. J. SCHOENBERG, Metric spaces and positive definite functions, *Trans. Amer. Math. Soc.* **44** (1938) 522—536.