Pooling designs with d-disjunct property and block weight $d + 1$

Chih-wen Weng (翁志文)

(with Yu-pei Huang and Wu, Hsin-Jung)

Department of Applied Mathematics, National Chiao Tung University, Taiwan

July 12, 2010

Definition

An incidence structure $(P, \mathcal{B})$ is called *d-disjunct* if any block in $\mathcal{B}$ is not covered by the union of d other blocks.

Definition

An incidence structure $(P, \mathcal{B})$ is called *d-disjunct* if any block in $\mathcal{B}$ is not covered by the union of d other blocks.

Assume $P = \{1, 2, \ldots, v\}$, $\mathcal{B} = \{B_1, B_2, \ldots, B_b\}$ and M is be the incidence matrix of $(P, \mathcal{B})$, i.e.

$$M_{ij} = \begin{cases} 1, & i \in B_j; \\ 0, & i \notin B_j \end{cases}$$

for $1 \leq i \leq v$ and $0 \leq j \leq b$.

Definition

An incidence structure $(P, \mathcal{B})$ is called *d-disjunct* if any block in $\mathcal{B}$ is not covered by the union of d other blocks.

Assume $P = \{1, 2, \dots, v\}$, $\mathcal{B} = \{B_1, B_2, \dots, B_b\}$ and M is be the incidence matrix of $(P, \mathcal{B})$, i.e.

$$M_{ij} = \begin{cases} 1, & i \in B_j; \\ 0, & i \notin B_j \end{cases}$$

for $1 \leq i \leq v$ and $0 \leq j \leq b$.

The incidence matrix M of a d-disjunct incidence structure can be used in non-adaptive group testing programming, in which $v << b$ is preferred.

1. Let M be a $v \times b$ incidence matrix of an incidence structure and set $F_2 = \{0, 1\}$. Define the output function $o_M : F_2^b \to F_2^v$ by

$$o_M(P) := M \star P = \bigcup_{P_i = 1} M_i,$$

where $\star$ is the matrix product by using Boolean sum to replace addition.

1. Let M be a $v \times b$ incidence matrix of an incidence structure and set $F_2 = \{0, 1\}$. Define the output function $o_M : F_2^b \to F_2^v$ by

$$o_M(P) := M \star P = \bigcup_{P_i = 1} M_i,$$

where $\star$ is the matrix product by using Boolean sum to replace addition.

2. If the incidence structure is d-disjunct, then $o_M \restriction F_2^b(\leq d)$ is known to be injective, where $F_2^b(\leq d)$ is the set of binary vectors of length b and Hamming weight at most d.

1. Let M be a $v \times b$ incidence matrix of an incidence structure and set $F_2 = \{0, 1\}$. Define the output function $o_M : F_2^b \to F_2^v$ by

$$o_M(P) := M \star P = \bigcup_{P_i = 1} M_i,$$

where $\star$ is the matrix product by using Boolean sum to replace addition.

2. If the incidence structure is d-disjunct, then $o_M \restriction F_2^b(\leq d)$ is known to be injective, where $F_2^b(\leq d)$ is the set of binary vectors of length b and Hamming weight at most d.

3. This means that for each element u in the image of o_M on $F_2^b(\leq d)$, we know which $P \in F_2^b$ to have $o_M(P) = u$.

1. Let M be a $v \times b$ incidence matrix of an incidence structure and set $F_2 = \{0, 1\}$. Define the output function $o_M : F_2^b \to F_2^v$ by

$$o_M(P) := M \star P = \bigcup_{P_i=1} M_i,$$

where $\star$ is the matrix product by using Boolean sum to replace addition.

2. If the incidence structure is d-disjunct, then $o_M \restriction F_2^b(\leq d)$ is known to be injective, where $F_2^b(\leq d)$ is the set of binary vectors of length b and Hamming weight at most d.
3. This means that for each element u in the image of o_M on $F_2^b(\leq d)$, we know which $P \in F_2^b$ to have $o_M(P) = u$.
4. In application, P is interpreted as the unknown infected subset $\{j \mid P_j = 1\}$ of a given set of b items, and u is interpreted as the sequence of test results. Then the injective property of o_M implies that the infected subset can be determined from the sequence of test results if the number of infected items is known in advance to be at most d.

Example

The following 4×6 binary matrix is used to detect the infected item in $\{1, 2, 3, 4, 5, 6\}$, if the infected item is known to be at most one in advance (but do not know which one):

Tests/Items	1	2	**3**	4	5	6		$o_M((0,0,1,0,0,0)^T)$
one	1	1	1	0	0	0	$\rightarrow$	1
Two	1	0	0	1	1	0	$\rightarrow$	0
Three	0	1	0	1	0	1	$\rightarrow$	0
Four	0	0	1	0	1	1	$\rightarrow$	1

Example

The following 4×6 binary matrix is used to detect the infected item in $\{1, 2, 3, 4, 5, 6\}$, if the infected item is known to be at most one in advance (but do not know which one):

Tests/Items	1	2	**3**	4	5	6		$o_M((0,0,1,0,0,0)^T)$
one	1	1	1	0	0	0	→	1
Two	1	0	0	1	1	0	→	0
Three	0	1	0	1	0	1	→	0
Four	0	0	1	0	1	1	→	1

If there are two infected items, the above 4×6 matrix does not work for detecting them. For example, both the infected sets $\{3, 4\}$ and $\{1, 6\}$ have the same output $(1, 1, 1, 1)^T$. So it is impossible to recover the infected set from the output $(1, 1, 1, 1)^T$.

Relation to t-design

Relation to t-design

Definition

An incidence structure $(P, \mathcal{B})$ is called a t-(v, k, λ) design if

1. $|P| = v$,
2. $|B| = k$ for and $B \in \mathcal{B}$, and
3. any t-subset of P is contained in exactly λ blocks in $\mathcal{B}$.

Remark

1. A 2-$(v, k, 1)$ design is $(k-1)$-disjunct since a block has k points and it intersects another block in at most one point, so $k-1$ other blocks can cover at most $k-1$ points of a block, leaving at least one point uncovered.
2. If any point is incidence in at least two blocks, then any block in a d-disjunct matrix has size at least $d+1$.
3. A d-disjunct incidence structure is called a pooling design.

First result

First result

Theorem

Let $(P, \mathcal{B})$ *be a d-disjunct pooling design with constant block size* $d+1$, *and define* $v = |P|$ *and* $b = |\mathcal{B}|$. *Then* $b \leq \max\{v(v-1)/d(d+1), v-d\}$. *Moreover if* $v - d \leq v(v-1)/d(d+1)$, *then the above upper bound of b is reached if and only if* $(P, \mathcal{B})$ *is a* 2-$(v, d+1, 1)$ *design.*

First result

Theorem

Let $(P, \mathcal{B})$ be a d-disjunct pooling design with constant block size $d+1$, and define $v = |P|$ and $b = |\mathcal{B}|$. Then $b \leq \max\{v(v-1)/d(d+1), v-d\}$. Moreover if $v - d \leq v(v-1)/d(d+1)$, then the above upper bound of b is reached if and only if $(P, \mathcal{B})$ is a 2-$(v, d+1, 1)$ design.

The $v \times b$ incidence matrix

$$M = \begin{pmatrix} I_b \\ J_d \end{pmatrix}$$

satisfies the equality $b = v - d$, where I_b is the $b \times b$ identity matrix and J_d is the $d \times d$ all 1's matrix.

The following example gives the equality in previous theorem for $d = q - 1$.

Example

$(2-(q^2, q, 1)$ design$)$ Let q be a prime power. The affine plane F_q^2 over F_q has q^2 points and $q^2 + q$ lines. Of course any line has q points and any two lines intersect at at most 1 point. Hence the points-lines incidence matrix is $v \times b$ d-disjunct with with constant weight w, where $v = q^2$, $b = q^2 + q$ and $w = q = d + 1$ satisfy

$$b = q^2 + q = v(v-1)/d(d+1).$$

The following example gives the equality in previous theorem for $d = q - 1$.

Example

$(2-(q^2, q, 1)$ design$)$ Let q be a prime power. The affine plane F_q^2 over F_q has q^2 points and $q^2 + q$ lines. Of course any line has q points and any two lines intersect at at most 1 point. Hence the points-lines incidence matrix is $v \times b$ d-disjunct with with constant weight w, where $v = q^2$, $b = q^2 + q$ and $w = q = d + 1$ satisfy

$$b = q^2 + q = v(v-1)/d(d+1).$$

The first q which is not a prime power is when $q = 6 = d + 1$. In this case the equality does not hold by the Bruck-Ryser-Chowla Theorem. Then there is no 5-disjunct pooling design with 36 points, 42 blocks and constant bock size 6. We will construct a 5-disjunct pooling design with 36 points, 37 blocks and constant block size 6.

Forward difference property

1. Let q be a prime power and $m \geq q$ be an integer.

Forward difference property

1. Let q be a prime power and $m \geq q$ be an integer.
2. Let $F_q := \{0, a^0, a^1, \ldots, a^{q-2}\}$ denote the finite field of q elements, where a is a generator of the cyclic multiplication group $F_q^* := F_q - \{0\}$.

Forward difference property

1. Let q be a prime power and $m \geq q$ be an integer.
2. Let $F_q := \{0, a^0, a^1, \ldots, a^{q-2}\}$ denote the finite field of q elements, where a is a generator of the cyclic multiplication group $F_q^* := F_q - \{0\}$.
3. Let $m \geq q$ be an integer. Let $\mathbb{Z}_m := \{0, 1, \ldots, m-1\}$ be the addition group of integers modulo m. We use the order of integers to order the elements in $\mathbb{Z}_m$, e.g. $0 < 1$.

Forward difference property

1. Let q be a prime power and $m \geq q$ be an integer.
2. Let $F_q := \{0, a^0, a^1, \ldots, a^{q-2}\}$ denote the finite field of q elements, where a is a generator of the cyclic multiplication group $F_q^* := F_q - \{0\}$.
3. Let $m \geq q$ be an integer. Let $\mathbb{Z}_m := \{0, 1, \ldots, m-1\}$ be the addition group of integers modulo m. We use the order of integers to order the elements in $\mathbb{Z}_m$, e.g. $0 < 1$.
4. A subset $T \subseteq \mathbb{Z}_m \times F_q$ is said to have the forward difference distinct property in $\mathbb{Z}_m \times F_q$ if the forward difference set

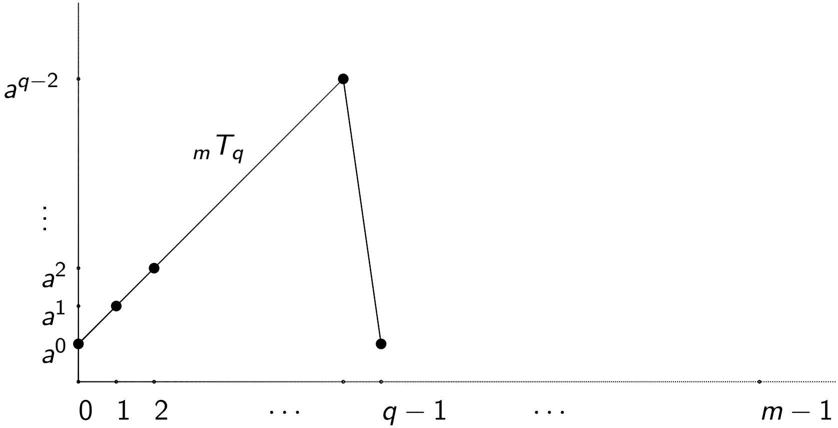
$$FD_T := \{(j, y) - (i, x) \mid (i, x), (j, y) \in T \text{ with } i < j\}$$

consists of $\frac{|T|(|T|-1)}{2}$ elements.

The Set ${}_mT_q$

Let ${}_mT_q \subseteq \mathbb{Z}_m \times F_q$ be defined by

$$ {}_mT_q = \{(i, a^i) \mid i \in \mathbb{Z}_m, 0 \leq i \leq q-1\}. $$



The Set ${}_5T_5$

For $q = 5$, $a = 2$,

$$ {}_5T_5 = \{(0,1),(1,2),(2,4),(3,3),(4,1)\} $$

and

$$ \begin{aligned} FD_{{}_5T_5} = \{ \quad & (1,1),(1,2),(1,4),(1,3) \\ & (2,3),(2,1),(2,2) \\ & (3,2),(3,4) \\ & (4,0) \quad \}. \end{aligned} $$

The Set ${}_5T_5$

For $q = 5$, $a = 2$,

$${}_5T_5 = \{(0,1),(1,2),(2,4),(3,3),(4,1)\}$$

and

$$\begin{aligned} FD_{{}_5T_5} = \{ \quad & (1,1),(1,2),(1,4),(1,3) \\ & (2,3),(2,1),(2,2) \\ & (3,2),(3,4) \\ & (4,0) \quad \}. \end{aligned}$$

Since $|FD_{{}_5T_5}| = 10$, the set ${}_5T_5$ has the forward difference distinct property in $\mathbb{Z}_5 \times F_5$.

${}_mT_q$ has the forward difference distinct property

Lemma

The set ${}_mT_q$ has the forward difference distinct property in $\mathbb{Z}_m \times T_q$.

Proof.

Given any pair $(c, d) \in \mathbb{Z}_m \times F_q$, solve the equations

$$(c, d) = (j, a^j) - (i, a^i)$$

for $0 \leq i < j \leq q - 1$. Note that $1 \leq c \leq q - 1$ to have a solution. If $c = q - 1$ then $j = q - 1$ and $i = 0$. If $c \neq q - 1$ then $a^i = d/(a^{j-i} - 1) = d/(a^c - 1)$ and $j = c + i$. In each case the pair $(i, a^i), (j, a^j)$ is unique determined by the element $(c, d) \in \mathbb{Z}_m \times F_q$. □

Difference Property

A subset $T \subseteq \mathbb{Z}_m \times F_q$ is said to have the difference distinct property in $\mathbb{Z}_m \times F_q$ if the difference set $D_T := -FD_T \cup FD_T$ consists of $|T|(|T|-1)$ elements.

Difference Property

A subset $T \subseteq \mathbb{Z}_m \times F_q$ is said to have the difference distinct property in $\mathbb{Z}_m \times F_q$ if the difference set $D_T := -FD_T \cup FD_T$ consists of $|T|(|T|-1)$ elements.

Since ${}_mT_q$ intersects a vertical line in at most one point, we find $(0,x) \notin D_{{}_mT_q}$ for any $x \in F_q$.

Non-example ($m = q = 5$)

We have seen

$$FD_{{}_5T_5} = \{ \begin{array}{l} (1,1),(1,2),(1,4),(1,3) \\ (2,3),(2,1),(2,2) \\ (3,2),(3,4) \\ (4,0) \quad \}. \end{array}$$

Hence

$$-FD_{{}_5T_5} = \{ \begin{array}{l} (4,4),(4,3),(4,1),(4,2) \\ (3,2),(3,4),(3,3) \\ (2,3),(2,1) \\ (1,0) \quad \}. \end{array}$$

Since $|D_{{}_5T_5}| = 16 \neq 20$, the set ${}_5T_5$ does not have the difference distinct property in $\mathbb{Z}_5 \times F_5$.

Example $(m-1=q=5)$

$$
\begin{aligned}
FD_{{}_6T_5} = \{ \quad & (1,1),(1,2),(1,4),(1,3) \\
& (2,3),(2,1),(2,2) \\
& (3,2),(3,4) \\
& (4,0) \quad \}.
\end{aligned}
$$

Hence considering as the negative in $\mathbb{Z}_6 \times F_5$, we have

$$
\begin{aligned}
-FD_{{}_6T_5} = \{ \quad & (5,4),(5,3),(5,1),(5,2) \\
& (4,2),(4,4),(4,3) \\
& (3,3),(3,1) \\
& (2,0) \quad \}.
\end{aligned}
$$

Since $|D_{{}_6T_5}| = 20$ now, the set ${}_6T_5$ has the difference distinct property in $\mathbb{Z}_6 \times F_5$.

${}_{2q-1}T_q$ has the difference distinct property

Lemma

For $m \geq 2q-1$, the set ${}_mT_q$ has the difference distinct property in $\mathbb{Z}_m \times T_q$.

Proof.

We have $|FD_{{}_mT_q}| = |-FD_{{}_mT_q}| = q(q-1)/2$. The first coordinate of an element in $FD_{{}_{2q-1}T_q}$ runs from 1 to $q-1$, and the first coordinate of an element in $-FD_{{}_{2q-1}T_q}$ from $m+1-q$ to $m-1$. The assumption $m \geq 2q-1$ implies $-FD_{{}_{2q-1}T_q} \cap FD_{{}_{2q-1}T_q} = \emptyset$. □

${}_{2q-3}T_q$ has the difference distinct property

Lemma

The set ${}_{m}T_q$ has the difference distinct property for $m = 2q-3$.

Proof.

We have $|FD_{T_{m,q}}| = |-FD_{T_{m,q}}| = q(q-1)/2$. Let $(c,d) \in FD_{T_{m,q}}$. If $m = 2q-3$, then $1 \leq c \leq q-1$ and $q-2 \leq -c \leq 2q-4$. Thus the repetition of differences occurs only when $c = q-2$ or $c = q-1$. Note that $d = 0$ iff $c = q-1$, and $-d = 0$ iff $-c = q-2$. For $c = q-2$, suppose $(c', d') \in -FD_{{}_mT_q}$ and $(c', d') = (c, d)$. Then we have $c' = q-2$ and $d' = 0$. Hence $d = 0$, a contradiction. Similarly for $c = q-1$, we have $d = 0$ but $(q-1, 0) \notin -FD_{T_{m,q}}$.

□

${}_{2q-4}T_q$ has the difference distinct property

Lemma

The set ${}_mT_q$ has the difference distinct property for $m = 2q - 4$.

Proof.

Let $(c, d) \in FD_{T_{m,q}}$. Since $m = 2q - 4$, we have $1 \leq c \leq q - 1$ and $q - 3 \leq -c \leq 2q - 5$. Thus the repetition of differences occurs only when $c = q - 3$, $q - 2$ or $q - 1$. Note that $d = 0$ iff $c = q - 1$, and $-d = 0$ iff $-c = q - 3$. For $c = q - 1$ or $c = q - 3$, similar process as the above $m = 2q - 3$ case can be applied to get contradictions. For $c = q - 2$, $-c = q - 2$. Thus a repetition implies that there are $(q - 2, d_1), (q - 2, d_2) \in FD_{T_{m,q}}$ such that $d_1 = -d_2$. Note that the only two elements of $FD_{T_{m,q}}$ with the first coordinate $q - 2$ are $(q - 2, a^{q-2} - 1)$ and $(q - 2, a^{q-1} - a)$, where a is the generator chosen for F_q^*. So we have $a^{q-2} - 1 = -(a^{q-1} - a)$ and this implies $a = -1$, also a contradiction. □

Lines with any two intersecting in at most a point

Proposition

Suppose that ${}_mT_q \subseteq \mathbb{Z}_m \times F_q$ *has the difference distinct property in* $\mathbb{Z}_m \times F_q$. *Set* $\mathcal{B} = \{u +_m T_q \mid u \in \mathbb{Z}_m \times F_q\}$. *Then* $|L \cap L'| \leq 1$ *for any distinct* $L, L' \in \mathcal{B}$.

Proof.

Routine. □

Lines with any two intersecting in at most a point

Proposition

Suppose that ${}_mT_q \subseteq \mathbb{Z}_m \times F_q$ *has the difference distinct property in* $\mathbb{Z}_m \times F_q$. *Set* $\mathcal{B} = \{u +_m T_q \mid u \in \mathbb{Z}_m \times F_q\}$. *Then* $|L \cap L'| \leq 1$ *for any distinct* $L, L' \in \mathcal{B}$.

Proof.

Routine. □

1. Note that there are mq lines and mq points in $\mathbb{Z}_m \times F_q$, and a line has $q = |T|$ points with q different first coordinates.

Lines with any two intersecting in at most a point

Proposition

Suppose that ${}_mT_q \subseteq \mathbb{Z}_m \times F_q$ *has the difference distinct property in* $\mathbb{Z}_m \times F_q$. *Set* $\mathcal{B} = \{u +_m T_q \mid u \in \mathbb{Z}_m \times F_q\}$. *Then* $|L \cap L'| \leq 1$ *for any distinct* $L, L' \in \mathcal{B}$.

Proof.

Routine. □

1. Note that there are mq lines and mq points in $\mathbb{Z}_m \times F_q$, and a line has $q = |T|$ points with q different first coordinates.
2. Apparently more lines can be added to $\mathcal{B}$ still having the conclusion of the above proposition, for example, adding vertical lines to $\mathcal{B}$.

Lines with any two intersecting in at most a point

Proposition

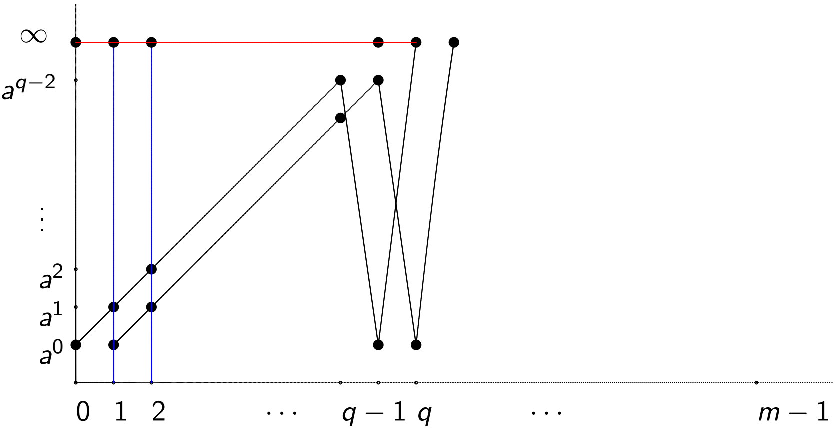
Suppose that ${}_mT_q \subseteq \mathbb{Z}_m \times F_q$ *has the difference distinct property in* $\mathbb{Z}_m \times F_q$. *Set* $\mathcal{B} = \{u +_m T_q \mid u \in \mathbb{Z}_m \times F_q\}$. *Then* $|L \cap L'| \leq 1$ *for any distinct* $L, L' \in \mathcal{B}$.

Proof.

Routine. □

1. Note that there are mq lines and mq points in $\mathbb{Z}_m \times F_q$, and a line has $q = |T|$ points with q different first coordinates.
2. Apparently more lines can be added to $\mathcal{B}$ still having the conclusion of the above proposition, for example, adding vertical lines to $\mathcal{B}$.
3. We will add m more points to P, add $m+1$ lines to $\mathcal{B}$, and add one more point to each original line in $\mathcal{B}$.

A picture for the finial result



Lines in $Z_m \times (F_q \cup \{\infty\})$

Second and final result

Theorem

There exists a q-disjunct pooling design $(P, \mathcal{B})$ *with* $|P| = m(q+1)$, $|\mathcal{B}| = m(q+1)+1$ *and constant block weight* $q+1$*, where* q *is a prime power, and* m *is an integer at least three satisfying* $m = 2q-4$, $m = 2q-3$ *or* $m \geq 2q-1$.

Second and final result

Theorem

There exists a q-disjunct pooling design $(P, \mathcal{B})$ *with* $|P| = m(q+1)$, $|\mathcal{B}| = m(q+1)+1$ *and constant block weight* $q+1$, *where* q *is a prime power, and* m *is an integer at least three satisfying* $m = 2q-4$, $m = 2q-3$ *or* $m \geq 2q-1$.

By choosing $q = 5$ and $m = 2q - 4 = 6$, there exists a 5-disjunct pooling design with 36 points, 37 blocks and constant block size 6.

The end

Thank you for your attention.