J Comb Optim
DOI 10.1007/s10878-008-9197-2

New error-correcting pooling designs associated with finite vector spaces

Jizhu Nan · Jun Guo

Abstract Let $\mathbb{F}_q^n$ be a n-dimensional vector space over $\mathbb{F}_q$. In this paper we construct a new family of inclusion matrices associated with subspaces of $\mathbb{F}_q^n$, and exhibit their disjunct properties.

Keywords Pooling designs · s^e-disjunct matrix · Subspaces

1 Introduction

The basic problem of group testing is to identify the set of defective items in a large population of items. Suppose we have n items to be tested and that there are at most r defective items among them. Each *test* (or *pool*) is (or contains) a subset of items. We assume some testing mechanism exists which if applied to an arbitrary subset of the population gives a *negative outcome* if the subset contains no positive and *positive outcome* otherwise. Objectives of group testing vary from minimizing the number of tests, limiting number of pools, limiting pool sizes to tolerating a few errors. It is conceivable that these objectives are often contradicting, thus testing strategies are application dependent. A group testing algorithm is *non-adaptive* if all tests must be specified without knowing the outcomes of other tests. A non-adaptive testing algorithm is useful in many areas such as DNA library screening.

A group testing algorithm is *error tolerant* if it can detect some errors in test outcomes. A mathematical model of error-tolerance designs is an s^e*-disjunct* matrix.

J. Nan · J. Guo
Dept. of Applied Math., Dalian University of Technology, Dalian 116024, People's Republic of China

J. Guo (✉)
Math. and Inf. College, Langfang Teachers' College, Langfang 065000, People's Republic of China
e-mail: guojun_lf@163.com

Published online: 11 December 2008

A binary matrix M is said to be s^e-*disjunct* if given any $s+1$ columns of M with one designated, there are $e+1$ rows with a 1 in the designated column and 0 in each of the other s columns. An s^0-disjunct matrix is said to be s-*disjunct*.

The constructions of s^e-disjunct matrices were given by many authors (see Balding and Torney 1996; Du and Hwang 2006; Du et al. 2006; D'yachkov et al. 2005; Huang and Weng 2004; Macula 1996, 1997; Ngo 2008; Ngo and Du 1999, 2002). In this paper we construct a new family of inclusion matrices associated with subspaces, and exhibit their disjunct properties. (See Theorems 3.2 and 3.3).

2 The finite vector space

In this section we will first introduce the concepts of finite vector space, and then introduce some counting formulas in the vector space.

Let $\mathbb{F}_q$ be a finite field with q elements, where q is a prime power. Let $\mathbb{F}_q^n$ be the n-dimensional *row vector space* over the finite field $\mathbb{F}_q$. The set $GL_n(\mathbb{F}_q)$ of all $n \times n$ nonsingular matrices over $\mathbb{F}_q$ forms a group under matrix multiplication, called the *general linear group* of degree n over $\mathbb{F}_q$. Clearly, $GL_n(\mathbb{F}_q)$ is transitive on the set of all subspaces of the same dimension in $\mathbb{F}_q^n$ (Wan 2002).

Let m_1, m_2 be two integers. Then the *Gaussian coefficient*

$$\begin{bmatrix} m_2 \\ m_1 \end{bmatrix}_q = \frac{\prod_{t=m_2-m_1+1}^{m_2}(q^t-1)}{\prod_{t=1}^{m_1}(q^t-1)}.$$

By convenience $\left[\begin{smallmatrix} m_2 \\ 0 \end{smallmatrix}\right]_q = 1$ and $\left[\begin{smallmatrix} m_2 \\ m_1 \end{smallmatrix}\right]_q = 0$ whenever $m_1 < 0$ or $m_2 < m_1$.

Proposition 2.1 (Wan 2002, Theorem 1.7) *Let $0 \le m \le n$. Then the number of m-dimensional subspaces of $\mathbb{F}_q^n$ is $\left[\begin{smallmatrix} n \\ m \end{smallmatrix}\right]_q$.*

Proposition 2.2 (Wan et al. 1966, Chap. 1, Theorem 5) *The number of $m \times n$ matrices with rank i over $\mathbb{F}_q$ is*

$$N(i; m \times n) = q^{i(i-1)/2} \begin{bmatrix} m \\ i \end{bmatrix}_q \prod_{t=n-i+1}^{n} (q^t - 1).$$

Proposition 2.3 *For $1 \le m, r \le n$ and $\max\{0, 2m-n\} \le i \le m$, let P and Q be two fixed m-dimensional subspaces of $\mathbb{F}_q^n$ with $\dim(P \cap Q) = i$. Then the number of r-dimensional subspaces S of $\mathbb{F}_q^n$ satisfying $\dim(P \cap S) = \dim(S \cap Q) = j$ is*

$$p_{j,j}^{i}(m, r; n) = \sum_{\beta+\gamma=j \le \alpha+\beta,\, \alpha+\beta+\rho+\gamma=r} q^{\omega} \prod_{l=j-\beta+1}^{\alpha} (q^l - 1)$$
$$\times \begin{bmatrix} m-i-\gamma \\ \alpha+\beta-j \end{bmatrix}_q \begin{bmatrix} m-i \\ \alpha \end{bmatrix}_q \begin{bmatrix} i \\ \beta \end{bmatrix}_q \begin{bmatrix} m-i \\ \gamma \end{bmatrix}_q \begin{bmatrix} n-2m+i \\ \rho \end{bmatrix}_q,$$

where $\omega = j(j+1)/2+\alpha(\alpha-1)/2+\beta(\beta-1)/2+2m\rho+i(\alpha+\gamma-\rho)-j(\alpha+\beta)-(\beta\gamma+\rho\alpha+\rho\beta+\rho\gamma)$. *In particular, for a given* m*-dimensional subspace* P *of* $\mathbb{F}_q^n$*, the number of* r*-dimensional subspaces intersecting* P *at* j*-dimensional subspaces of* $\mathbb{F}_q^n$ *is*

$$p_{j,j}^m(m,r;n) = q^{(r-j)(m-j)} \begin{bmatrix} n-m \\ r-j \end{bmatrix}_q \begin{bmatrix} m \\ j \end{bmatrix}_q.$$

Proof Denote by $P_{j,j}^i(m,r;n)$ the set of r-dimensional subspaces S of $\mathbb{F}_q^n$ satisfying both $\dim(P\cap S) = \dim(S\cap Q) = j$. By the transitivity of $GL_n(\mathbb{F}_q)$ on the set of subspaces with the same dimension, we may assume

$$P = (I^{(m)}\ 0), \qquad Q = \begin{pmatrix} 0^{(i,m-i)} & I^{(i)} & 0 & 0 \\ 0 & 0 & I^{(m-i)} & 0 \end{pmatrix}.$$

For any $S \in P_{j,j}^i(m,r;n)$, write S in block as

$$S = \begin{array}{c} \begin{matrix} m-i & i & m-i & n-2m+i \end{matrix} \\ \begin{pmatrix} S_1 & S_2 & S_3 & S_4 \end{pmatrix} \end{array}.$$

Suppose rank $S_4 = \rho$, rank $(S_1, S_4) = \rho+\alpha$, rank $(S_1, S_3, S_4) = \rho+\alpha+\gamma$ and $\beta = r-(\rho+\alpha+\gamma)$. By suitable row elementary transformations, we may pick S as

$$\begin{array}{c} \begin{matrix} m-i & i & m-i & n-2m+i \end{matrix} \\ \begin{pmatrix} S_{11} & S_{12} & S_{13} & 0 \\ 0 & S_{22} & 0 & 0 \\ 0 & S_{32} & S_{33} & 0 \\ S_{41} & S_{42} & S_{43} & S_{44} \end{pmatrix} \begin{matrix} \alpha \\ \beta \\ \gamma \\ \rho \end{matrix} \end{array},$$

where rank $S_{44} = \rho$, rank $S_{11} = \alpha$, rank $S_{33} = \gamma$ and rank $S_{22} = \beta$. Note that there are

$$\begin{bmatrix} m-i \\ \alpha \end{bmatrix}_q, \qquad \begin{bmatrix} i \\ \beta \end{bmatrix}_q, \qquad \begin{bmatrix} m-i \\ \gamma \end{bmatrix}_q \quad \text{and} \quad \begin{bmatrix} n-2m+i \\ \rho \end{bmatrix}_q$$

choices of subspaces S_{11}, S_{22}, S_{33} and S_{44}, respectively. By the transitivity of $GL_n(\mathbb{F}_q)$ on the set of subspaces with the same dimension, the number of S's does not depend on the particular choices of $S_{11}, S_{22}, S_{33}, S_{44}$. Without loss of generality we may assume

$$S_{11} = (I^{(\alpha)}0), \qquad S_{22} = (I^{(\beta)}0),$$
$$S_{33} = (I^{(\gamma)}0), \qquad S_{44} = (I^{(\rho)}0).$$

Then S has a matrix presentation of the form

$$\begin{pmatrix} I & 0^{(\alpha,m-i-\alpha)} & 0 & S_{122} & 0 & S_{132} & 0 & 0 \\ 0 & 0 & I & 0^{(\beta,i-\beta)} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & S_{322} & I & 0^{(\gamma,m-i-\gamma)} & 0 & 0 \\ 0 & S_{412} & 0 & S_{422} & 0 & S_{432} & I & 0^{(\rho,n-2m+i-\rho)} \end{pmatrix}. \quad (1)$$

Note that the matrix representation of S of the form (1) is unique. By $\dim(P \cap S) = j$ and $\dim(Q \cap S) = j$, we deduce that $\beta + \gamma = j \leq \alpha + \beta$ and rank $S_{132} = \alpha + \beta - j$. By Proposition 2.2, the desired result follows. □

3 The construction

In this section, we construct a family of inclusion matrices associated with subspaces of $\mathbb{F}_q^n$, and exhibit its disjunct property. Let $\mathcal{M}(m, n)$ be the set of all m-dimensional subspaces of $\mathbb{F}_q^n$.

Definition 3.1 Given integers $1 \leq r, m \leq n-1$ and $\max\{0, r+m-n\} \leq j \leq \min\{r, m\}$. Let $M(r, m; n)$ be the binary matrix whose rows (resp. columns) are indexed by $\mathcal{M}(r, n)$ (resp. $\mathcal{M}(m, n)$). We also order elements of these sets lexicographically. $M(r, m; n)$ has a 1 in row i and column l if and only if the i-th subspace of $\mathcal{M}(r, n)$ intersect the l-th subspace of $\mathcal{M}(m, n)$ at j-dimensional subspaces of $\mathbb{F}_q^n$.

By Propositions 2.1 and 2.3, $M(r, m; n)$ is a $\begin{bmatrix} n \\ r \end{bmatrix}_q \times \begin{bmatrix} n \\ m \end{bmatrix}_q$ matrix, whose constant row (resp. column) weight is $p^r_{j,j}(r, m; n) = q^{(r-j)(m-j)} \begin{bmatrix} n-r \\ m-j \end{bmatrix}_q \begin{bmatrix} r \\ j \end{bmatrix}_q$ (resp. $p^m_{j,j}(m, r; n) = q^{(r-j)(m-j)} \begin{bmatrix} n-m \\ r-j \end{bmatrix}_q \begin{bmatrix} m \\ j \end{bmatrix}_q$).

Theorem 3.2 *Let* $1 \leq r, m \leq n-1$ *and* $\max\{0, r+m-n\} \leq j \leq \min\{r, m\}$. *If* $1 \leq d \leq \lfloor p^m_{j,j}(m, r; n)/\alpha \rfloor + 1$, *then* $M(r, m; n)$ *is* d^e*-disjunct, where* $e = p^m_{j,j}(m, r; n) - d\alpha - 1$, $\alpha = \max\{p^l_{j,j}(m, r; n) \mid \max\{0, 2m-n\} \leq l \leq m-1\}$ *and* $p^i_{j,j}(m, r; n)$ *is given by Proposition* 2.3.

Proof Let $C, C_1, C_2, \ldots, C_d$ be $d+1$ distinct columns of $M(r, m; n)$. To obtain the maximum numbers of subspaces P of $\mathcal{M}(r, n)$ satisfying $\dim(P \cap C) = j$ and $\dim(P \cap C_i) = j$, by Proposition 2.3 we may assume that the number of subspaces P of $M(r, n)$ satisfying $\dim(P \cap C) = j$ and $\dim(P \cap C_i) = j$ is

$$\alpha = \max\{p^l_{j,j}(m, r; n) \mid \max\{0, 2m-n\} \leq l \leq m-1\}.$$

Hence the number of subspaces P of $\mathcal{M}(r, n)$ satisfying $\dim(P \cap C) = j$ and $\dim(P \cap C_1), \ldots, \dim(P \cap C_d) \neq j$ is at least

$$p^m_{j,j}(m, r; n) - d\alpha.$$

It follows that $e = p^m_{j,j}(m, r; n) - d\alpha - 1$. Since $e \geq 0$, we obtain

$$d \leq \left\lfloor \frac{p^m_{j,j}(m, r; n)}{\alpha} \right\rfloor + 1.$$

□

Remarks If $j = \min\{r, m\}$, then $M(r, m; n)$ is studied in D'yachkov et al. (2005), Ngo (2008), Ngo and Du (2002). If $j < \min\{r, m\}$, then $M(r, m; n)$ can not be obtained from pooling space (Huang and Weng 2004).

Theorem 3.3 *Let* $1 \le r, m \le n-1$ *and* $\max\{0, r+m-n\} \le j \le \min\{r, m\}$. *Then the following* (i)–(iv) *hold*:

(i) *If* $d \le 1+q+\cdots+q^{m-1}$ *and* $m-1 > r$, *then there exists a* j *such that* $M(r, m; n)$ *is* d^e*-disjunct, where* $e \ge \begin{bmatrix} m \\ r \end{bmatrix}_q - d\begin{bmatrix} m-1 \\ r \end{bmatrix}_q + (d-1)\begin{bmatrix} m-2 \\ r \end{bmatrix}_q$.

(ii) *If* $1+q+\cdots+q^{m-1} \ge d \ge q+2$ *and* $m > r \ge 2$, *then there exists a* j *such that* $M(r, m; n)$ *is* d^e*-disjunct, where* $e \ge \begin{bmatrix} m \\ r \end{bmatrix}_q - d\begin{bmatrix} m-1 \\ r \end{bmatrix}_q + (2d-3)\begin{bmatrix} m-2 \\ r \end{bmatrix}_q - (d-2)\begin{bmatrix} m-3 \\ r \end{bmatrix}_q$.

(iii) *If* $1+q+\cdots+q^{m-1} \ge d \ge q+3$ *and* $m > r \ge 2$, *then there exists a* j *such that* $M(r, m; n)$ *is* d^e*-disjunct, where* $e \ge \begin{bmatrix} m \\ r \end{bmatrix}_q - d\begin{bmatrix} m-1 \\ r \end{bmatrix}_q + (3d-7)\begin{bmatrix} m-2 \\ r \end{bmatrix}_q - (2d-6)\begin{bmatrix} m-3 \\ r \end{bmatrix}_q$.

(iv) *If* $1+q+\cdots+q^{m-1} \ge d \ge q+4$ *and* $m > r \ge 2$, *then there exists a* j *such that* $M(r, m; n)$ *is* d^e*-disjunct, where* $e \ge \begin{bmatrix} m \\ r \end{bmatrix}_q - d\begin{bmatrix} m-1 \\ r \end{bmatrix}_q + (4d-13)\begin{bmatrix} m-2 \\ r \end{bmatrix}_q - (3d-12)\begin{bmatrix} m-3 \\ r \end{bmatrix}_q$.

Proof (i) Let $j = r$, by [Ngo 2008, Theorem 2.1], the desired result follows.

(ii)–(iv) Let $j = r$, by [Ngo 2008, Corollaries 3.3–3.5], the desired results follow. □

Acknowledgements This paper is supported by NSF of China (No. 10771023), Natural Science Foundation of Hebei Province, China (No. A2008000128), educational committee of Hebei Province, China (No. 2008142), and Langfang Teacher's College (LSZZ200803).

References

Balding DJ, Torney DC (1996) Optimal pooling designs with error detection. J Comb Theory Ser A 74:131–140

Du D, Hwang F (2006) Pooling designs and non-adaptive group testing: important tools for DNA sequencing. World Scientific, Singapore

Du D, Hwang F, Wu W, Znati T (2006) New construction for transversal design. J Comput Biol 13:990–995

D'yachkov AG, Hwang FK, Macula AJ, Vilenkin PA, Weng C (2005) A construction of pooling designs with some happy surprises. J Comput Biol 12:1129–1136

Huang T, Weng C (2004) Pooling spaces and non-adaptive pooling designs. Discrete Math 282:163–169

Macula AJ (1996) A simple construction of d-disjunct matrices with certain constant weights. Discrete Math 162:311–312

Macula AJ (1997) Error-correcting non-adaptive group testing with d^e-disjunct matrices. Discrete Appl Math 80:217–222

Ngo HQ (2008) On a huperplane arrangement problem and tighter analysis of an error-tolerant pooling design. J Comb Optim 15:61–76

Ngo H, Du D (1999) A survey on combinatorial group testing algorithms with applications to DNA library screening. In: Discrete mathematical problems with medical applications, New Brunswick, NJ, 1999. DIMACS series in discrete mathematics and theoretical computer science, vol 55. Am Math Soc, Providence, pp 171–182

Ngo H, Du D (2002) New constructions of non-adaptive and error-tolerance pooling designs. Discrete Math 243:161–170

Wan Z (2002) Geometry of classical groups over finite fields, 2nd edn. Science, Beijing

Wan Z, Dai Z, Feng X, Yang B (1966) Studies in finite geometry and the construction of incomplete block designs. Science, Beijing (in Chinese)