Discrete Mathematics 338 (2015) 857–862

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

New error-correcting pooling designs with vector spaces over finite field

Xuemei Liu *, Xing Gao

College of Science, Civil Aviation University of China, Tianjin, 300300, PR China

ARTICLE INFO

Article history:
Received 12 April 2014
Received in revised form 22 November 2014
Accepted 7 January 2015
Available online 3 February 2015

Keywords:
Pooling designs
d^z-disjunct matrix
Vector space
Parameters
The test efficiency

ABSTRACT

Pooling designs have many applications in molecular biology. In this paper, we firstly construct a family of error-correcting pooling designs using the containment relationship of subspaces of vector spaces. Then by comparing with the test efficiencies, the new design is better than that in D'yachkov et al. (2005). At last, we analyze how the related parameters influence the test efficiency $\frac{t}{s}$.

1. Introduction

Given a set of c items with some defects, the group test is applicable to an arbitrary subset of items with two possible outcomes: a negative outcome indicates that all items in the subset are negative, while a positive outcome indicates that there is at least one item in the subset that is positive. The group testing problem is asking to identify all defects with the minimum number of tests, each of which is on a subset of items, called a pool, and the test-outcome is negative when the pool does not contain any defect and positive when the pool contains a defect at least. In other words, a pooling design is a specification of all tests such that they can be performed simultaneously, with the goal being to identify all positive items with a small number of tests in [2–4,9]. A group testing algorithm is non-adaptive if all tests could be specified without knowing the outcomes of other tests. A mathematical model of the non-adaptive group testing design is a d-disjunct matrix, which is also called a pooling design. Designing a good error-tolerant pooling design is the central problem in the area of non-adaptive group testing.

The pooling design can be represented by a binary matrix whose columns are indexed with items and rows are indexed with pools. An entry at cell (i, j) is 1 if and only if the ith pool is contained by the jth item, and 0, otherwise. In practice, test-outcomes may contain errors, to make pooling design error tolerant, one introduced the concept of d^z-disjunct matrix in [11]. A binary matrix M is said to be d^z-disjunct if given any $d+1$ columns of M with one designated, there are $z+1$ rows with a 1 in the designated column and 0 in each of the other d columns. The concept of fully d^z-disjunct matrix was given in the paper [7]. A d^z-disjunct matrix is fully d^z-disjunct if it is not $d_1^{z_1}$-disjunct whenever $d_1 > d$ or $z_1 > z$. A d^z-disjunct matrix can detect $z-1$ errors and correct $\lfloor \frac{z}{2} \rfloor$ errors in [6].

Pooling designs reduce costs for many applications in molecular biology, such as DNA library screening, and nonunique probe selection in [5]. There are also several constructions of d^z-disjunct matrices in the literature in [1,8,10,12].

* Corresponding author.
E-mail address: xm-liu771216@163.com (X. Liu).

http://dx.doi.org/10.1016/j.disc.2015.01.003

In this paper, we construct a new family of inclusion matrix associated with subspaces under vector space V, and by comparing with the test efficiencies, we illustrate that the new design gives better ratio of efficiency than the former one. Moreover, we do some parametric analysis and analyze how the related parameters influence the test efficiency $\frac{t}{s}$.

2. Parameters

In order to understand the following contents better, in this section, we will introduce the concepts of vector spaces and relevant counting formulas.

Let $\mathbb{F}_q$ be a finite field with q elements, where q is a prime power. For a positive integer n, let V be the n-dimensional row vector space over a finite field $\mathbb{F}_q$, and fix an $(n-b)$-subspace W of V, let $P = \{A|A \text{ is a subspace of } V, A+W=V\}$. Denote by $GL_n(\mathbb{F}_q)$ the set of all the $n \times n$ nonsingular matrices over $\mathbb{F}_q$, then $GL_n(\mathbb{F}_q)$ forms a group under matrix multiplication, called the general linear group of degree n over $\mathbb{F}_q$. Clearly, V admits an action of $GL_n(\mathbb{F}_q)$ defined as follows:

$$V \times GL_n(\mathbb{F}_q) \longrightarrow V$$
$$((x_1, x_2, \cdot, x_n), T) \longmapsto (x_1, x_2, \ldots, x_n)T.$$

We assume that U is an m-subspace of V with a basis $u_1, u_2, \ldots, u_m$, then the $m \times n$ matrix

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix}$$

is said to be a matrix representation of U. We usually denote a matrix representation of the m-subspace U still by U. The above action is transitive on the set of all the subspaces with the same dimension by Wan (2002, Theorem 1.3).

Denote the set of all i-subspaces U of vector space V satisfying $U+W=V$ by $M(i;n,b)$, where W is a fixed $(n-b)$-subspace of V, and let

$$N(i;n,b) = |M(i;n,b)|.$$

Let P_2 be a fixed l_2-subspace of vector space V satisfying $P_2+W=V$. Denote by $M(l_1,l_2;n,b)$ the set of l_1-subspaces contained in P_2, and denote by $M'(l_1,l_2;n,b)$ the set of l_2-subspaces containing a fixed l_1-subspace P_1 of V satisfying $P_1+W=V$. Let

$$N(l_1,l_2;n,b) = |M(l_1,l_2;n,b)|.$$
$$N'(l_1,l_2;n,b) = |M'(l_1,l_2;n,b)|.$$

Proposition 2.1 ([13]). *If $b \le i \le n$, let V denote the n-dimensional row vector space over a finite field $\mathbb{F}_q$, and fix an $(n-b)$-subspace W of V. Then the number of i-subspace U of V satisfying $W+U=V$ is $N(i;n,b) = q^{b(n-i)} \begin{bmatrix} n-b \\ i-b \end{bmatrix}_q$.*

Proposition 2.2 ([13]). *If $b \le l_1 \le l_2 \le n$, let V denote the n-dimensional row vector space over a finite field $\mathbb{F}_q$, and fix an $(n-b)$-subspace W of V. For a given l_2-subspace U_2 of V satisfying $U_2+W=V$, $N(l_1,l_2;n,b)$ denote the number of l_1-subspace U_1 of V satisfying $U_1+W=V$ and $U_1 \subseteq U_2$. Then*

$$N(l_1,l_2;n,b) = q^{b(l_2-l_1)} \begin{bmatrix} l_2-b \\ l_1-b \end{bmatrix}_q.$$

Proposition 2.3. *If $b \le l_1 \le l_2 \le n$, let V denote the n-dimensional row vector space over a finite field $\mathbb{F}_q$, and fix an $(n-b)$-subspace W of V. For a given l_1-subspace U_1 of V satisfying $U_1+W=V$, $N'(l_1,l_2;n,b)$ denote the number of l_2-subspace U_2 of V satisfying $U_2+W=V$ and $U_1 \subseteq U_2$. Then*

$$N'(l_1,l_2;n,b) = q^{b(l_2-l_1)} \begin{bmatrix} n-l_1 \\ l_2-l_1 \end{bmatrix}_q.$$

Proof. We know $N'(l_1,l_2;n,b)$ is independent of the particular l_1-subspace U_1 chosen. Define a (0, 1)-matrix, whose rows are labeled by the l_1-subspaces U_1 of V satisfying $U_1+W=V$, whose columns are labeled by the l_2-subspace U_2 of V satisfying $U_2+W=V$, and with a 1 or 0 in the (i,j) position of the matrix, if the ith l_1-subspace is or is not contained in the jth l_2-subspace, respectively. If we count the number of 1's in the matrix by rows, we get $N'(l_1,l_2;n,b)N(l_1;n,b)$, where $N'(l_1,l_2;n,b)$ is the number of 1's in each row and $N(l_1;n,b)$ is the number of rows. If we count the number of 1's in the matrix by columns, we get $N(l_1,l_2;n,b)N(l_2;n,b)$, where $N(l_1,l_2;n,b)$ is the number of 1's in each column and $N(l_2;n,b)$ is the number of columns. Thus we have

$$N'(l_1,l_2;n,b)N(l_1;n,b) = N(l_1,l_2;n,b)N(l_2;n,b)$$