Europ. J. Combinatorics (1990) **11**, 609–615

A Diagonal Form for the Incidence Matrices of t-Subsets *vs.* k-Subsets

RICHARD M. WILSON

1. INTRODUCTION AND SUMMARY

Given integers t, k and v, let $W_{tk}(v)$ (or simply W_{tk} if there is no danger of confusion) be the $\binom{v}{t}$ by $\binom{v}{k}$ matrix of 0's and 1's, the rows of which are indexed by the t-subsets T of a v-set X, whose columns are indexed by the k-subsets K of the same set X, and where the entry $W_{tk}(T, K)$ in row T and column K is 1 if $T \subseteq K$ and is 0 otherwise.

These matrices and their algebra have arisen in many combinatorial investigations. Some time ago, N. Linial and B. Rothschild [2], motivated by a result of P. Frankl, investigated the rank of W_{tk} over the field of two elements and derived the following formula for the rank modulo 2:

$$\sum_{f:D\to\mathbb{Z}^+} (-1)^{f(D)}\binom{t - \sum_{x\in D}^{v} f(x)2^x}{}.$$

Here D is the set of positive integers such that $k - t = \sum_{d\in D} 2^d$, and for any mapping f from D to the non-negative integers, $f(D) = \sum_{d\in D} f(d)$. They raise the question of the rank of W_{tk} modulo a general prime number p and give a formula for the rank of W_{tk} modulo 3 in the case $k = t + 1$.

We have an answer for any prime p, but it has a different form from that of the above formula.

THEOREM 1. *For $t \leqslant \min\{k, v - k\}$, the rank of W_{tk} modulo a prime p is*

$$\sum \binom{v}{i} - \binom{v}{i-1}$$

where the sum is extended over those indices i such that p does not divide the binomial coefficient

$$\binom{k-i}{t-i}.$$

In the statement of the theorem, $\binom{v}{-1}$ should be interpreted as zero. This theorem follows immediately from our main theorem (Theorem 2) below. A corollary is that W_{tk} has rank modulo p equals to the number $\binom{v}{t}$ of its rows for all primes $p > k$.

We investigate the modules over the integers $\mathbb{Z}$ generated by the rows of these matrices and give what we call a *diagonal form* (similar to Smith normal form) for W_{tk} in Theorem 2. As another application of the Lemma in Section 3, we give another proof of a result (Theorem 3 of Section 5) which gives necessary and sufficient conditions for a vector to belong to the $\mathbb{Z}$-module generated by the columns of W_{tk}.

We say that a matrix D is a *diagonal form* for a matrix M when D is diagonal and there exist unimodular matrices (square integral matrices which have integral inverses) E and F such that $D = EMF$. We do not require that M and D are square; here 'diagonal' just means that the (i, j) entry of D is 0 if $i \neq j$. (If the diagonal entries

0195-6698/90/060609 + 07 $02.00/0

$d_{11}, d_{22}, \ldots$ of D are non-negative and such that

$$d_{11} \mid d_{22} \mid d_{33} \ldots,$$

then D is called the Smith normal form of M.)

THEOREM 2. *If $t \leq \min\{k, v-k\}$, then W_{tk} has as a diagonal form the $\binom{v}{t} \times \binom{v}{k}$ diagonal matrix with diagonal entries*

$$\binom{k-i}{t-i} \text{ with multiplicity } \binom{v}{i} - \binom{v}{i-1}, \qquad i = 0, 1, \ldots, t.$$

The proof of Theorem 2 is given in Section 4.

2. PRELIMINARIES

In this section, we introduce some notation and summarize some results on modules, etc., that we refer to in the sequel.

For an integer matrix M, we use $\text{row}_{\mathbb{Z}}(M)$ to denote the $\mathbb{Z}$-module generated by the row vectors of M, and $\text{row}_{\mathbb{Q}}(M)$ to denote the span of the rows over the rationals. We define the *index* of an integral matrix M to be the index of $\text{row}_{\mathbb{Z}}(M)$ as a subgroup of the module $Z(M)$ of all integral vectors which belong to $\text{row}_{\mathbb{Q}}(M)$. Thus index 1 means that any integral vector which is a rational linear combination of the rows of M is already an integral linear combination of the rows of M.

The index of M is the absolute value of the product of the non-zero entries of any diagonal form of M. (More generally, if a diagonal form has non-zero entries $d_1, d_2, d_3, \ldots$, then the finite group $Z(M)/\text{row}_{\mathbb{Z}}(M)$ is isomorphic to the direct sum of cyclic groups of orders $d_1, d_2, d_3, \ldots$.)

Given two matrices M_1 and M_2 (of the same shape), there exist unimodular matrices E and F such that $M_1 = EM_2F$ (we say M_1 and M_2 are *integrally equivalent* and write $M_1 \sim M_2$ in this case) iff M_1 can be obtained from M_2 by a sequence of integral elementary row and column operations.

COROLLARY 1. *Let M be an integral matrix. Then M has index 1 iff $M = MAM$ for some integral matrix A.*

PROOF. Suppose $M = MAM$ and let $\mathbf{x}$ be an integral vector in $\text{row}_{\mathbb{Q}}(M)$, say $\mathbf{x} = \mathbf{y}M$, where $\mathbf{y}$ is rational. Then

$$\mathbf{x} = \mathbf{y}M = \mathbf{y}MAM = (\mathbf{x}A)M = \mathbf{z}M$$

where $\mathbf{z}$ is integral; so $\mathbf{x} \in \text{row}_{\mathbb{Z}}(M)$ and this shows that M has index 1.

Conversely, suppose $EMF = D$, where E, F are unimodular and D is diagonal with entries 0 and 1. Say M is $m \times n$. If $m \leq n$, let $F' = F$ and let E' be obtained from E by adjoining $(n-m)$ rows of zeros; if $m \geq n$, let $E' = E$ and let F' be obtained from F by adjoining $(m-n)$ columns of zeros. In either case, $MF'E'M = M$. □

Proposition 1 shows that an integral $r \times n$ matrix M of rank r will have an integral right inverse iff M has index 1 (because if the rows of M are independent, $MA = I$ is equivalent to $MAM = M$).

We also apply the term *index* to submodules $\mathcal{M}$ of $\mathbb{Z}^n$ in an analogous manner. The index of $\mathcal{M}$ is the index of $\mathcal{M}$ as a subgroup of

$$\bar{\mathcal{M}} = \{\mathbf{x} \in \mathbb{Z}^n : c\mathbf{x} \in \mathcal{M} \text{ for some } c \in \mathbb{Z}, c \neq 0\}.$$

We emphasize that we mean $[\bar{\mathcal{M}}: \mathcal{M}]$ and not $[\mathbb{Z}^n: \mathcal{M}]$, the latter of which will be infinite unless $\bar{\mathcal{M}} = \mathbb{Z}^n$. If $\mathcal{M}_1$ and $\mathcal{M}_2$ are two modules of the same dimension, same index, and if $\mathcal{M}_1 \subseteq M_2$, then $\mathcal{M}_1 = \mathcal{M}_2$.

If $\mathcal{M}_1 \subseteq \mathcal{M}_3 \subseteq \mathbb{Z}^n$ and $\mathcal{M}_1$ has index 1, then the quotient $\mathcal{M}_3/\mathcal{M}_1$ is a free $\mathbb{Z}$-module (because if $\mathbf{a} + \mathcal{M}_1$ has finite order, then $\mathbf{a} \in \bar{\mathcal{M}}_1 = \mathcal{M}_1$). The same argument as used for vector spaces yields the following proposition.

PROPOSITION 2. *Let $\mathcal{M}_1$, $\mathcal{M}_2$ be submodules of $\mathbb{Z}^n$. If $\mathcal{M}_1$ has index* 1, *then any $\mathbb{Z}$-basis of $\mathcal{M}_1$ can be extended to a $\mathbb{Z}$-basis of $\mathcal{M}_1 + \mathcal{M}_2$ by adjoining elements of $\mathcal{M}_2$.*

We also need:

PROPOSITION 3. *If $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r$ is a $\mathbb{Z}$-basis for a module $\mathcal{M} \subseteq \mathbb{Z}^n$ of index* 1, *then the matrix the rows of which are $d_1\mathbf{e}_1, d_2\mathbf{e}_2, \ldots, d_r\mathbf{e}_r$ has as a diagonal form the $r \times n$ diagonal matrix with diagonal entries $d_1, d_2, \ldots, d_r$ and, in particular, has index $d_1 d_2, \ldots, d_r$ if all d_i are non-zero.*

3. An Important Lemma

Of fundamental importance is the equation

$$W_{jt} W_{tk} = \binom{k-j}{t-j} W_{jk} \tag{3.1}$$

which holds for $0 \leq j \leq t \leq k \leq v$. This holds because for a j-subset S and a k-subset K of our v-set X,

$$(W_{jt} W_{tk})(S, K) = \sum_T W_{jt}(S, T) W_{tk}(T, K),$$

where the sum is extended over all t-subsets T of X. The right-hand side is the number of t-subsets T such that $S \subseteq T \subseteq K$, and this number is $\binom{k-j}{t-j}$ if $S \subseteq K$, and 0 otherwise.

Given $n_i \times m$ matrices A_i, $i = 0, 1, \ldots, l$, we denote by

$$\bigcup_{i=0}^{l} A_i$$

the $(n_0 + n_1 + \cdots n_l) \times m$ matrix obtained by stacking the matrices $A_0, A_1, \ldots, A_l$ one on top of the other. For $0 \leq t \leq k \leq v$, define

$$M_{tk} = \bigcup_{i=0}^{t} W_{ik}.$$

Equation (3.1) shows that $\mathrm{row}_{\mathbb{Q}}(W_{jk}) \subseteq \mathrm{row}_{\mathbb{Q}}(W_{tk})$ for $j \leq t$ and hence $\mathrm{row}_{\mathbb{Q}}(M_{tk}) = \mathrm{row}_{\mathbb{Q}}(W_{tk})$. In particular, M_{tk} has rank at most $\binom{v}{t}$, the number of rows of W_{tk}.

LEMMA 1. *For non-negative integers t, k, v with $t \leq k \leq v - t$, the matrix M_{tk} has rank $\binom{v}{t}$ and index* 1.

PROOF. The proof will be by induction, on the sum $t + k + v$, say.

We first establish the validity of the lemma whenever $k = v - t$. To this end, we claim that for any $t \leq k \leq v$,

$$\sum_{i=0}^{l} (-1)^i W_{ik}^T W_{ik} = I \tag{3.2}$$

where $l = \min\{k, v-k\}$, I is the identity of order $\binom{v}{k}$, and where $\bar{W}_{ik}$ is the $\binom{v}{i} \times \binom{v}{k}$ matrix defined by

$$\bar{W}_{ik}(S, K) = \begin{cases} 1 & \text{if } S \cap K = \varnothing \\ 0 & \text{otherwise} \end{cases}$$

for i-subsets S and k-subsets K. To prove this, just note that the entry in row A and column B on the right-hand side of (3.2) is

$$\sum_{i=0}^{l} (-1)^i \binom{|B| - |A \cap B|}{i} = \begin{cases} 0 & \text{if } A \neq B; \\ 1 & \text{if } A = B. \end{cases}$$

In the case $k = v - t$, the upper limit l of the sum is equal to t and (3.2) can be written

$$\bar{M}_{tk}^T M_{tk} = I \qquad \left(\text{of order } \binom{v}{t}\right)$$

where

$$\bar{M}_{tk} = \bigcup_{i=0}^{t} (-1)^i \bar{W}_{ik}.$$

Since M_{tk} has an integral left inverse, it has rank $\binom{v}{t}$ and index 1.

We now assume $k < v - t$.

Given $1 \leq j \leq k$, choose a point x_0 in the v-set. Then the rows (j-subsets) and columns (k-subsets) of $W_{jk}(v)$ are partitioned according to whether or not they contain x_0. This gives us a block decomposition of $W_{jk}(v)$:

$$W_{jk}(v) = \begin{array}{|c|c|} \hline W_{j-1,k-1}(v-1) & 0 \\ \hline W_{j,k-1}(v-1) & W_{jk}(v-1) \\ \hline \end{array}$$

After permuting rows, we find that $M_{tk}(v)$ is integrally equivalent to

$$\begin{array}{|c|c|} \hline M_{t-1,k-1}(v-1) & 0 \\ \hline M_{t,k-1}(v-1) & M_{tk}(v-1) \\ \hline \end{array}$$

In view of the induction hypothesis applied to $M_{t-1,k-1}(v-1)$ and $M_{tk}(v-1)$, we can use elementary integral row and column operations to reduce the above matrix to

$$\begin{array}{|cc|cc|} \hline I_1 & 0 & & \\ 0 & 0 & \multicolumn{2}{c}{0} \\ \hline & & I_2 & 0 \\ \multicolumn{2}{c|}{?} & 0 & 0 \\ \hline \end{array} \tag{3.3}$$

where I_1 and I_2 are identity matrices of orders $\binom{v-1}{t-1}$ and $\binom{v-1}{t}$, respectively. Then surely $\operatorname{rank}(M_{tk}(v)) \geq \binom{v-1}{t-1} + \binom{v-1}{t} = \binom{v}{t}$, so that we now know that the rank of $M_{tk}(v)$ is $\binom{v}{t}$.

Further row operations on the matrix in (3.3) can be used to create an identity of

order $\binom{v}{t}$ as a submatrix of some $M \sim M_{tk}(v)$. At that point, since $\binom{v}{t}$ is the rank of $M_{tk}(v)$, all other entries of M must be zeros. □

We prefaced the statement of the lemma by asserting that the rank of M_{tk} was at most $\binom{v}{t}$. The argument applied to the rational field, but the lemma proves that M_{tk} has rank exactly $\binom{v}{t}$ *over any field.* The lemma, of course, implies the well known fact that W_{tk} has rank$\binom{v}{t}$ over the rationals.

4. Proof of Theorem 2

Theorem 2 follows easily from the next two propositions. When $t \leq k \leq v - t$, Propositions 4 and 5 assert the existence of a $\binom{v}{t} \times \binom{v}{k}$ integral matrix E, the rows of which form a $\mathbb{Z}$-basis for an index 1 module (namely, M_{tk}) and such that the rows of DE form a $\mathbb{Z}$-basis for W_{tk}, where D is a (square) diagonal matrix with $\binom{v}{i} - \binom{v}{i-1}$ occurrences of $\binom{k-i}{t-i}$ on the diagonal, $i = 0, 1, \ldots, t$. A reference to Proposition 3 completes the proof of Theorem 2. □

Proposition 4. *Let $k \leq v$ be given and put $l = \min\{k, v-k\}$. There exist integral matrices $E_{0k}, E_{1k}, \ldots, E_{lk}$ such that E_{ik} is a $\binom{v}{i} - \binom{v}{i-1}$ by $\binom{v}{k}$ matrix, the rows of which are in row$_{\mathbb{Z}}(W_{ik})$ and such that for each $t \leq l$, the rows of $E_{0k} \cup E_{1k} \cup \cdots \cup E_{tk}$ form a $\mathbb{Z}$-basis for row$_{\mathbb{Z}}(M_{tk})$.*

Proof. Let $E_{0k} = W_{0k}$. Inductively, when $E_{0k}, E_{1k}, \ldots, E_{ik}$ $(i < l)$ have been defined, use Proposition 2 to extend the rows of $E_{0k} \cup \cdots \cup E_{ik}$ (a $\mathbb{Z}$-basis for row$_{\mathbb{Z}}(M_{ik})$ which has index 1 by Lemma 1) to a $\mathbb{Z}$-basis of

$$\text{row}_{\mathbb{Z}}(M_{i+1,k}) = \text{row}_{\mathbb{Z}}(M_{ik}) + \text{row}_{\mathbb{Z}}(W_{i+1,k})$$

by adding $\binom{v}{i+1} - \binom{v}{i}$ vectors (these are to be the rows of $E_{i+1,k}$) from row$_{\mathbb{Z}}(W_{i+1,k})$. □

Proposition 5. *Let $E_{0k}, E_{ik}, \ldots, E_{lk}$ $(l = \min\{k, v-k\})$ be as in Proposition 4. Then for $t \leq l$, a $\mathbb{Z}$-basis for row$_{\mathbb{Z}}(W_{th})$ is provided by the rows of*

$$\binom{k}{t}E_{0k} \cup \binom{k-1}{t-1}E_{1k} \cup \binom{k-2}{t-2}E_{2k} \cup \cdots \cup E_{tk}.$$

Proof. The proof is by induction on k. The case $k = 0$ is trivial. Fix $k > 0$. There is nothing to prove if $t = k$ since the assertion reduces to Proposition 4, so we will assume $t < k$.

Since the rows of E_{ik} are contained in row$_{\mathbb{Z}}(W_{ik})$, (3.1) shows that the rows of $\binom{k-i}{t-i}W_{ik}$ are contained in row$_{\mathbb{Z}}(W_{tk})$. The matrix

$$\bigcup_{i=0}^{t} E_{ik}$$

has index 1, so by Proposition 3, the rows of

$$\bigcup_{i=0}^{t} \binom{k-i}{t-i}E_{ik}$$

generate a submodule $\mathcal{M}$ of row$_{\mathbb{Z}}(W_{tk})$ which has rank $\binom{v}{t}$ and index

$$N = \prod_{i=0}^{t} \binom{k-i}{t-i}^{\binom{v}{i}-\binom{v}{i-1}}. \tag{4.1}$$

To complete the proof, we will show that $\text{row}_{\mathbb{Z}}(W_{tk})$ itself has index N defined by (4.1).

We have $2t \leq v$. Let $E_{0t}, E_{1t}, \ldots, E_{tt}$ be the $\binom{v}{i} - \binom{v}{i-1}$ by $\binom{v}{t}$ matrices as in Proposition 4. Define integral matrices A_{itk} for $0 \leq i \leq t$ by

$$E_{it}W_{tk} = \binom{k-i}{t-i}A_{itk}.$$

That is, each A_{itk} is a $\binom{v}{i} - \binom{v}{i-1}$ by $\binom{v}{k}$ matrix, the rows of which are contained in $\text{row}_{\mathbb{Z}}(W_{ik})$. Since the union of the rows of $E_{0t}, \ldots, E_{tt}$ form a $\mathbb{Z}$-basis for $\text{row}_{\mathbb{Z}}(M_{tt})$ which consists of all integral vectors of length $\binom{v}{t}$, it is clear that the rows of

$$\bigcup_{i=0}^{t} \binom{k-i}{t-i}A_{itk}$$

form a $\mathbb{Z}$-basis for $\text{row}_{\mathbb{Z}}(W_{tk})$. We now claim that the rows of

$$A = \bigcup_{i=0}^{t} A_{itk}$$

form a $\mathbb{Z}$-basis for $\text{row}_{\mathbb{Z}}(M_{tk})$. Since M_{tk} has index 1, Proposition 3 will show that $\text{row}_{\mathbb{Z}}(W_{tk})$ has index N and will complete the proof.

The rows of A are contained in $\text{row}_{\mathbb{Z}}(M_{tk})$ since they are integral vectors which are rational linear combinations of the rows of W_{tk} (and hence M_{tk}) and M_{tk} has index 1. To show that they generate $\text{row}_{\mathbb{Z}}(M_{tk})$, we need to show that for $j \leq t$, the rows of W_{jk} are integral linear combinations of these rows.

Now by our induction hypothesis, $\text{row}_{\mathbb{Z}}(W_{jt})$ has $\mathbb{Z}$-basis consisting of the rows of

$$\bigcup_{i=0}^{j} \binom{t-i}{j-i}E_{it}.$$

Since $\binom{k-j}{t-j}W_{jk} = W_{jt}W_{tk}$, the rows of $\binom{k-j}{t-j}W_{jk}$ are integral linear combinations of the rows of

$$\left(\bigcup_{i=0}^{j} \binom{t-i}{j-i}E_{it}\right)W_{tk} = \bigcup_{i=0}^{j} \binom{t-i}{j-i}\binom{k-i}{t-i}A_{itk}$$

$$= \binom{k-j}{t-j}\left(\bigcup_{i=0}^{j} \binom{k-i}{j-i}A_{itk}\right).$$

This shows that $\text{row}_{\mathbb{Z}}(W_{jk}) \subseteq \text{row}_{\mathbb{Z}}(A)$ and completes the proof. □

5. The Column Module

The following Theorem 3 was first proved in [3]. When applied to a constant vector, it shows that the column vector of height $\binom{v}{t}$ all of whose co-ordinates are λ belongs to the module $\text{col}_{\mathbb{Z}}(W_{tk})$ iff

$$\lambda\binom{v-i}{t-i} \equiv 0 \quad \left(\text{mod}\binom{k-i}{t-i}\right)$$

for $i = 0, 1, \ldots, t$; this corollary was also proved in [1]. These congruences are recognizable as the necessary conditions for the existence of a t-design $S_\lambda(t, k, v)$.

Theorem 3. *Let $t \leq k \leq v - t$. A (column) vector $\mathbf{x}$ of height $\binom{v}{t}$ belongs to the*

module $\mathrm{col}_{\mathbf{Z}}(W_{tk})$ *generated by the columns of* W_{tk} *iff*

$$\frac{1}{\binom{k-i}{t-i}} W_{it}\mathbf{x} \tag{5.1}$$

is integral for $i = 0, 1, \ldots, t$.

PROOF. If $\mathbf{x} \in \mathrm{col}_{\mathbf{Z}}(W_{tk})$, then $\mathbf{x} = W_{ty}\mathbf{y}$ for some integral $\mathbf{y}$ and then, using (3.1),

$$\frac{1}{\binom{k-i}{t-i}} W_{it}\mathbf{x} = \frac{1}{\binom{k-i}{t-i}} W_{it}W_{tk}\mathbf{y} = W_{ik}\mathbf{x},$$

which is integral. Thus the conditions (5.1) are necessary.

Now assume that $\mathbf{x}$ satisfies the conditions (5.1). By Proposition 1 and Lemma 1, there exists an integral matrix A of size $\binom{v}{k}$ by $1+\binom{v}{1}+\cdots+\binom{v}{t}$ such that $M_{tk} = M_{tk}AM_{tk}$, which implies

$$W_{tk} = W_{tk}AM_{tk} = W_{tk}\sum_{i=0}^{t} A_i W_{ik},$$

where A_i is a $\binom{v}{k} \times \binom{v}{i}$ submatrix of A. Then, by (3.1),

$$W_{tk} = W_{tk}\left(\sum_{i=0}^{t} \frac{1}{\binom{k-i}{t-i}} A_i W_{it}\right) W_{tk}.$$

Since W_{tk} has rank $\binom{v}{t}$ equal to the numbers of rows, we can cancel it on the right to obtain

$$I = W_{tk}\sum_{i=0}^{t} \frac{1}{\binom{k-i}{t-i}} A_i W_{it}.$$

Then

$$\mathbf{x} = W_{tk}\left(\sum_{i=0}^{t} A_i\right) \frac{1}{\binom{k-i}{t-i}} W_{it}\mathbf{x} = W_{tk}\sum_{i=0}^{t} A_i\mathbf{y}_i = W_{tk}\mathbf{z}$$

where the $\mathbf{y}_i$'s and $\mathbf{z}$ are integral. □

ACKNOWLEDGEMENTS

This research was supported in part by NSF Grant DMS-8703898. This manuscript was written while the author was a visiting member of the Institute for Mathematics and its Applications, University of Minnesota.

REFERENCES

1. J. Graver and W. B. Jurkat, The module structure of integral designs, *J. Comb. Theory, Ser. A.* **15** (1973), 75–90.
2. N. Linial and B. Rothschild, Incidence matrices of subsets—a rank formula, *SIAM J. Alg. Discr. Meth.* **2** (1981), 333–340.
3. R. M. Wilson, The necessary conditions for t-designs are sufficient for something, *Util. Math.* **4** (1973), 207–217.

Received 23 June 1988 and in revised form 28 November 1988

RICHARD M. WILSON

Department of Mathematics, California Institute of Technology, Pasadena, CA 91125, U.S.A.