(0,1)-Matrices with Different Rows

Hasmik, Sahakyan

Institute for Informatics and Automation Problems
Yerevan, Armenia
e-mail: hasmik@ipia.sci.am

ABSTRACT

We consider (0,1)-matrices with given row and column sums, and with an additional requirement/constraint - restricted number of repeated rows.

Keywords

(0,1)-matrices, hypergraph degree sequences

1. INTRODUCTION

(0,1) matrices with prescribed row and/or column sums are classical objects which appear in many branches of mathematics. Various combinatorial problems from n-cube geometry, hypergraph theory, discrete tomography problems, are effectively modeled in terms of (0,1) matrices.

Consider a family of sets (or a hypergraph) (A, F) where A, $|A| = n$, is a finite set of elements, and F, $|F| = m$, is a collection of subsets of A. Code subsets of A with (0,1) sequences of length n such that i-th component of the sequence equals 1 if and only if i-th element of A is included in the subset. We get a (0,1) matrix of size $m \times n$, where the numbers of 1's in rows are cardinalities of subsets; the number of 1s in j-th column is the number of subsets of F containing j-th element of A. Let $R = (r_1, \cdots, r_m)$ and $S = (s_1, \cdots, s_n)$ denote the row and column sum vectors of the matrix respectively, where r_i is the numbers of 1's in i –th row, and s_j is the numbers of 1's in j –th column. Given a (0,1) matrix, R and S can be easily calculated.

Consider the inverse problem:

Existence/reconstruction of a binary $m \times n$ matrix with the given row sum $R = (r_1, \cdots, r_m)$ and/or column sum $S = (s_1, \cdots, s_n)$.

The Gale-Ryser's theorem is known for solving the problem in polynomial time.

Theorem 1 [1]. Let $R = (r_1, \cdots, r_m)$ and $S = (s_1, \cdots, s_n)$ be non-increasing, positive integer vectors such that $r_i \le n$ for $i = 1, \cdots, m$, and $s_j \le m$ for $j = 1, \cdots, n$. Let $S^* = (s_1^*, \cdots, s_n^*)$ denote the conjugate vector of R: $s_i^* = |\{r_i : r_i \ge j,\ i = 1, \cdots, m\}|$. There exists a binary $m \times n$ matrix with row sum R and column sum S if and only if S is majorized by R^*, that is: $\sum_{i=1}^{k} s_i \le \sum_{i=1}^{k} s_i^*$, for $k = 1, \cdots, n-1$, and $\sum_{i=1}^{n} s_i = \sum_{i=1}^{n} s_i^*$.

We impose an additional requirement/constraint - **restricted number of repeated rows**. Consider the following problems:

Given $R = (r_1, \cdots, r_m)$, $S = (s_1, \cdots, s_n)$, and an integer number k.

(P1).Existence/reconstruction of a (0,1) matrix of size $m \times n$ with the given row sum R and column sum S, and with $\le k$ pairs of repeated rows.

(P2).Existence/reconstruction of a (0,1) matrix of size $m \times n$ with the given column sum S, and with $\le k$ pairs of repeated rows.

Given integer numbers n, m, and k.

(P3).Characterization of $\psi^k(m, n)$, the set of all column sum vectors of (0,1)-matrices of size $m \times n$, which have at most k pairs of repeated rows.

In this paper we restrict our attention to the following values of k: $k = 0$; $k = 1$.

$k = 0$ - is equivalent to the requirement of different rows of matrices in (P1),(P2) and (P3). In this case the problem can be reformulated in terms of hypergraphs, for example, (P2) is equivalent to the existence of simple hypergraphs with the given degree sequence, (P3) is equivalent to the characterization of the set of degree sequences of simple hypergraphs. This is a hypergraph **degree sequence** problem, which is one of the open problems in the hypergraph theory. The complexity of the problem is not known even for 3-uniform hypergraphs. The problem is investigated by several authors, and a number of partial results have been achieved, for example in [2-6].

$k = 1$ - can be considered as a relaxation of the problem, when one pair of repeated rows is allowed.

In Section 2 below we consider particular cases of (P1) and (P2), where row and column sum vectors are homogenous: $R = (r, \cdots, r)$ and $S = (s, \cdots, s)$. We received simple necessary and sufficient conditions for this case.

Section 3 concerns (P3). A simple structural characterization is known for $\psi^0(m, n)$. We prove that $\psi^1(m, n)$ has a similar structure. Moreover, we prove that the problem is not easier for $k = 1$.

2. HOMOGENOUS ROW AND COLUMN SUM VECTORS

Theorem 2. Let $R = (r, \cdots, r)$ and $S = (s, \cdots, s)$ be positive integer vectors of size m and n, respectively, such that $r \le n$ and $\sum_{i=1}^{n} s = \sum_{i=1}^{m} r$. There exists a (0,1)-matrix of size $m \times n$ with row sum R and column sum S, and with different rows if and only if $m \le C_n^r$.

Skeleton of the proof is as follows. The majorization condition from Gale-Ryser's theorem always holds for $R = (r, \cdots, r)$ and $S = (s, \cdots, s)$, if $\sum_{i=1}^{n} s = \sum_{i=1}^{m} r$.

Each (0,1)-matrix of size $m \times n$ with row sum R, and with different rows can be identified with m-subset on the r-th layer of the n-dimensional unit cube. Hence, $m \le C_n^r$ is a necessary and sufficient condition for existence of a such matrix. For providing column sum S, we use the technique of partitioning the n-dimensional unit cube in two directions.

Theorem 3. Let $S = (s , \cdots , s)$ be a positive integer vector of size n; m be a positive integer, and $m \leq 2^n$.There exists a (0,1)-matrix of size $m \times n$, with the column sum vector $S = (s , \cdots , s)$, and with different rows, if and only if

$$m - \sum_{i=0}^{k} \left((n-i) \cdot C_n^{n-i} \right) + \left\lfloor \frac{\delta(n-k-1)}{n} \right\rfloor \leq s \leq \sum_{i=0}^{k} \left((n-i) \cdot C_n^{n-i} \right) + \lfloor \delta(n-k-1)/n \rfloor,$$

where k and δ are parameters from the canonical representation form of m*:*

$m = C_n^n + C_n^{n-1} + \cdots + C_n^{n-k} + \delta,\ \delta < C_n^{n-k-1}$.

We prove the theorem using several results from [5-6]. We compose also a matrix with column sum S and with different rows.

3. STRUCTURAL CHARACTERIZATION

For a given m and n, let $\psi(m,n)$ denote the set of all column sum vectors of (0,1)-matrices of size $m \times n$, which have different rows.

We consider $\psi(m,n)$ as a subset in the following ranked partially ordered set (poset):

$\Xi_{m+1}^n = \{(s_1, \cdots, s_n) : 0 \leq s_i \leq m\}$,

with the component-wise partial order: $(s_1, \cdots, s_n) \leq (q_1, \cdots, q_n)$ if and only if $s_i \leq q_i$, and with the rank of elements defined as: $r(s_1, \cdots, s_n) = s_1 + \cdots + s_n$.

Thus, $\psi(m,n) \subseteq \Xi_{m+1}^n$.

We define an upper area (upper subposet) of Ξ_{m+1}^n:

$\widehat{H} = \{(a_1, \cdots, a_n) \in \Xi_{m+1}^n | a_i \geq m_{mid}\ for\ all\ i\}$, where $m_{mid} = (m+1)/2$, for odd m, and $m_{mid} = m/2$ for even m.

Theorem 4 below reduces the characterization area from Ξ_{m+1}^n to $\widehat{H}$. Theorem 5 describes the structure.

Theorem 4 [5].$\psi(m,n)$ can be easily derived from $\psi(m,n) \cap \widehat{H}$, its part in $\widehat{H}$.

Theorem 5 [5] $\psi(m,n) \cap \widehat{H}$ is an ideal in $\widehat{H}$.

An illustration is given in Figure 1 below.

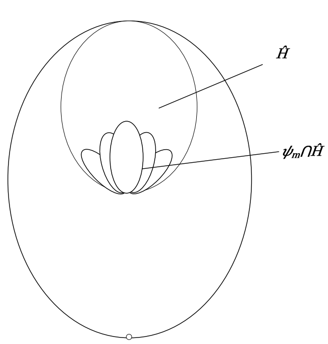

Figure 1

4. RELAXATION OF THE PROBLEM

For a given m and n, let $\psi^1(m,n)$ denote the set of all column sum vectors of (0,1)-matrices of size $m \times n$, which have at most one pair of repeated rows.

Obviously, $\psi(m,n) \subseteq \psi^1(m,n)$.

Theorem 6. $\psi^1(m,n)$ can be easily derived from $\psi^1(m,n) \cap \widehat{H}$, its part in $\widehat{H}$.

Theorem 7. $\psi^1(m,n) \cap \widehat{H}$ is an ideal in $\widehat{H}$.

An illustration is given in Figure 2 below.

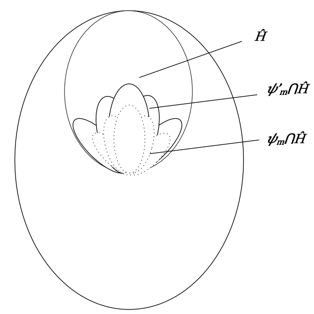

Figure 2

Theorem 6 and Theorem 7 are analogous of Theorem 4 and Theorem 5, respectively, and have been proved with the same technique.

Theorem 8. The question of characterizing $\psi^1(m,n)$ has the same complexity as the question of characterizing $\psi(m,n)$.

4. CONCLUSION

We considered (0,1)-matrices with prescribed row/column sums and with different rows, and investigate the questions of existence/construction, and characterization of the set of all solutions. The problem has its analogue in hypergraph theory, discrete tomography, and others. The complexity is not known, and this is one of the known open problems. We solved the problem for homogenous row and column sum vectors. Further we investigate a relaxation of the problem, when one pair of repeated rows is allowed.

REFERENCES

[1]. H.J. Ryser. Combinatorial Mathematics. Carus Math. Monograph # 14, Math.Assoc. of America, Providence, 1964.

[2]. D.Billington, Conditions for degree sequences to be realisable by 3-uniform hypergraphs", 1988.

[3]. N.L. Bhanu Murthy And Murali K. Srinivasan, The polytope of degree sequences of hypergraphs. 2002.

[4]. W.Kocay, P. C., Li, On 3-hypergraphs with equal degree sequences, 2007.

[5] H. A. Sahakyan, Numerical characterization of n-cube subset partitioning, Discrete Applied Mathematics 157 (2009) 2191-2197.4

[6] H. A. Sahakyan, L.A. Aslanyan, Essential points of n-cube subsets partitioning characterization, ODSA 2010, Rostock.