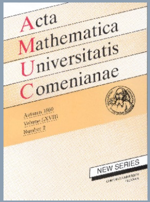


TRANSVERSALS OF RECTANGULAR ARRAYS

S. SZABÓ

Abstract. The paper deals with m by n rectangular arrays whose mn cells are filled with symbols. A section of the array consists of m cells, one from each row and no two from the same column. The paper focuses on the existence of sections that do contain symbols with high multiplicity.

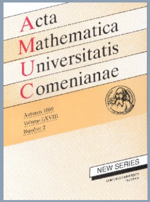
1. Introduction

An n by n array of cells filled with symbols $1, 2, \ldots, n$ such that each symbol appears in each row and each column exactly once is called a Latin square. A section is a set of n cells, one from each row such that no two cells are in the same column. A section is called a transversal if each of its symbols is distinct. H. J. Ryser [5] conjectured that every n by n Latin square has a transversal for odd n. P. W. Shor [6] proved that an n by n Latin square has a section with

$$n - 5.53(\ln n)^2$$

distinct symbols. S. K. Stein [7] showed that if an n by n array is filled with symbols $1, 2, \ldots, n$ such that each symbol appears exactly n times then there is a section with $0.63n$ distinct symbols. P. Erdös and J. H. Spencer [4] proved that if an n by n array is filled with symbols such that each symbol appears at most $(n-1)/16$ times, then the array has a transversal. In this paper we will use the Erdös-Spencer technique to show that m by n arrays have sections in which no symbol appears with high multiplicity.

Received June 27, 2007; revised January 12, 2008.
2000 *Mathematics Subject Classification.* Primary 05D15.
Key words and phrases. latin square; array; transversal; section; probabilistic method; Lovász local lemma.



2. The graph G

Consider an m by n table filled with symbols $1, 2, \dots$ such that each symbol appears at most k times. In order to avoid trivial cases we assume that $2 \le m \le n$. For a given value of m and n there is a large number of such tables. We will work with a fixed table. The symbol in the a-th row and the b-th column is denoted by $f(a, b)$. The s cells

$$[x_1, y_1], \dots, [x_s, y_s]$$

in the table is called an s-clique if

(1) $x_1, \dots, x_s$ are distinct numbers,
(2) $y_1, \dots, y_s$ are distinct numbers,
(3) $f(x_1, y_1) = \dots = f(x_s, y_s)$.

Again to avoid non-desired cases we assume that $2 \le s \le m \le n$. Let T be the set of all s-cliques in the table. We define a graph G in the following way. Let the elements of T be the vertices of G. Two distinct vertices

$$\{[x_1, y_1], \dots, [x_s, y_s]\} \text{ and } \{[x'_1, y'_1], \dots, [x'_s, y'_s]\}$$

are connected if

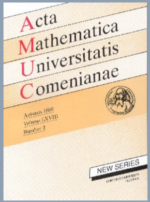
$$\{x_1, \dots, x_s\} \cap \{x'_1, \dots, x'_s\} \ne \emptyset$$

or

$$\{y_1, \dots, y_s\} \cap \{y'_1, \dots, y'_s\} \ne \emptyset.$$

Note that the degree of a vertex of G is at most

$$[s(m-s) + s(n-s) + s^2]\binom{k-1}{s-1}.$$



Go back
Full Screen
Close
Quit

The reason is the following. Choose an s-clique C. Then consider the s rows and s columns of the table that contain a cell from C. These s rows and s columns occupy $s(m-s)+s(n-s)+s^2$ cells of the table. Let us call this the shaded area of the table. Another s-clique C' is connected to C if and only if C' has a cell from the shaded area. There are at most $s(m-s)+s(n-s)+s^2$ choices for such a cell. The common cell contains a symbol. This symbol appears at most k times in the table. So there are at most $\binom{k-1}{s-1}$ choices for the remaining $s-1$ cells of the clique C'.

3. The probability space Ω

Let ω be an injective map from $\{1,\dots,m\}$ to $\{1,\dots,n\}$. The set of cells

$$[i,\omega(i)],\ 1\le i\le m$$

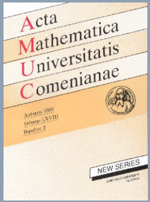
is called a section of the table. Intuitively a section consists of m cells of the table such that no two cells are in the same row and no two cells are in the same column.

Let Ω be the probability space consisting of all sections of the table. Clearly,

$$|\Omega| = n(n-1)\cdots(n-m+1).$$

We assign the same probability to each element of Ω. For an element $\{[x_1,y_1],\dots,\dots,[x_s,y_s]\}$ of T we define $A([x_1,y_1],\dots,[x_s,y_s])$ to be the subset of Ω which contains all ω with $\omega(x_1)=y_1,\dots,\omega(x_s)=y_s$. Intuitively, $A([x_1,y_1],\dots,[x_s,y_s])$ is the set of all sections that contain the cells $[x_1,y_1],\dots,[x_s,y_s]$. For notational convenience we number the elements of T by $1,2,\dots,\mu$ and identify the elements of T by their numbers. If the vertex $\{[x_1,y_1],\dots,[x_s,y_s]\}$ is numbered by i, then $A([x_1,y_1],\cdots,[x_s,y_s])$ will be denoted by A_i. As an example suppose that $\{[1,1],\dots,[s,s]\}$ is a vertex of G and is numbered by 1. The event A_1 consists of all the ω for which

$$\omega(1)=1,\ \omega(2)=2,\dots,\omega(s)=s.$$



$$\begin{aligned}
\Pr[A_1] &= \frac{[n-s][n-s-1]\cdots[n-s-(m-s)+1]}{n(n-1)\cdots(n-m+1)} \\
&= \frac{1}{n(n-1)\cdots(n-s+1)} \\
&= p.
\end{aligned}$$

In general $\Pr[A_i] = p$ for all i, $1 \le i \le \mu$.

4. The conditional probabilities

The content of this section is the following lemma.

Lemma 1. *Suppose that the vertex* 1 *is not adjacent to any of the vertices* $2, \dots, t$ *in the graph* G *and that* $\Pr[\overline{A}_2 \cdots \overline{A}_t] > 0$*. Then* $\Pr[A_1 | \overline{A}_2 \cdots \overline{A}_t] \le p$*.*

Proof. By definition

$$\Pr[A_1 | \overline{A}_2 \cdots \overline{A}_t] = \frac{\Pr[A_1 \overline{A}_2 \cdots \overline{A}_t]}{\Pr[\overline{A}_2 \cdots \overline{A}_t]}.$$

The event $A_1 \overline{A}_2 \cdots \overline{A}_t$ is the set of all ω for which

$$\omega \in A_1,\ \omega \notin A_2, \dots, \omega \notin A_t.$$

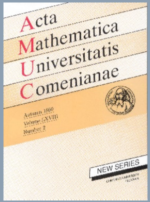
Intuitively $A_1 \overline{A}_2 \cdots \overline{A}_t$ is the set of all sections that contain the clique $\{[1,1], \dots, \dots, [s,s]\}$ associated with A_1 and do not contain any of the cliques associated with the events $A_2, \dots, A_t$. Let $S(y_1, \dots, y_s)$ be the set of all ω with

$$\omega(1) = y_1, \dots, \omega(s) = y_s,\ \omega \notin A_2, \dots, \omega \notin A_t.$$

Intuitively $S(y_1, \dots, y_s)$ is the set of all sections that contain the clique

$$\{[1, y_1], \dots, [s, y_s]\}$$

Go back
Full Screen
Close
Quit



and do not contain any of the cliques associated with $A_2, \dots, A_t$. Clearly, $S(1, \dots, \dots, s) = A_1\overline{A}_2 \cdots \overline{A}_t$ and the sets $S(y_1, \dots, y_s)$ form a partition of the set $\overline{A}_2 \cdots \overline{A}_t$ as $y_1, \dots, y_s$ vary over the possible $n(n-1)\cdots(n-s+1)$ values. Next we try to establish that $|S(1, \dots, s)| \le |S(y_1, \dots, y_s)|$. If $S(1, \dots, s) = \emptyset$, then $|S(1, \dots, s)| \le |S(y_1, \dots, y_s)|$ holds. So we may assume that $S(1, \dots, s) \neq \emptyset$. Choose an ω from $S(1, \dots, s)$. Consider the cells $[1, y_1], \dots, [s, y_s]$. Then define the sets A, B, C in the following way. Let

$$\begin{aligned} A &= \{y_1, \dots, y_s\}, \\ B &= \{a : \ a \in A, \ a \le s\}, \\ C &= \{a : \ a \in A, \ a > s, \ a \in \text{ range of } \omega\}. \end{aligned}$$

Suppose that C has u elements, say $j_1, \dots, j_u$. Then $\{1, \dots, s\} \setminus B$ has at least u elements, say $i_1, \dots, i_v$. There are $x_1, \dots, x_u$ such that $\omega(x_1) = j_1, \dots, \omega(x_u) = j_u$. Clearly, $x_1, \dots, x_u \ge s + 1$. Define ω^* by

$$\begin{array}{lllllll} \omega^*(1) & = & y_1 & , \dots, & \omega^*(s) & = & y_s, \\ \omega^*(x_1) & = & i_1 & , \dots, & \omega^*(x_u) & = & i_u \end{array}$$

and $\omega^*(x) = \omega(x)$ for all x, $s + 1 \le x \le m$, $x \notin \{x_1, \dots, x_u\}$. Note that $\omega^* \in S(y_1, \dots, y_s)$. From a given ω^* we can reconstruct ω without any ambiguity. Namely setting

$$\begin{array}{lllllll} \omega(1) & = & 1 & , \dots, & \omega(s) & = & s, \\ \omega(x_1) & = & j_1 & , \dots, & \omega(x_u) & = & j_u \end{array}$$

and $\omega(x) = \omega^*(x)$ for all x, $s + 1 \le x \le m$, $x \notin \{x_1, \dots, x_u\}$. Thus the map $* : S(1, \dots, s) \to S(y_1, \dots, y_s)$ defined by $\omega \to \omega^*$ is injective. This gives that $|S(1, \dots, s)| \le |S(y_1, \dots, y_s)|$. Table 1 illustrates our consideration in the $s = 8$, $u = 3$, $v = 4$ special case. The cells $[1, \omega(1)], \dots, [m, \omega(m)]$ are marked with "$\times$" and the cells $[1, y_1], \dots, [s, y_s]$ are marked with "$\bullet$".

Go back
Full Screen
Close
Quit

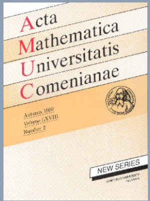


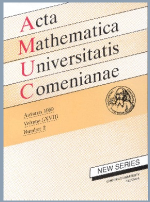
Table 1. An illustration in the $s = 8$, $u = 3$, $v = 4$ case.

	i_1		i_2				i_3	i_4			j_1	j_2	j_3	
		y_1		y_3	y_4	y_7			y_2		y_6	y_8	y_5	
	*										×			x_1
			*									×		x_2
										×				
							*						×	x_3
8								×				•		
7						•	×							
6						×					•			
5					×								•	
4				×	•									
3			×	•										
2		×							•					
1	×	•												
	1	2	3	4	5	6	7	8						

Now turn back to the probability estimations.

$$\Pr[A_1\overline{A}_2\cdots\overline{A}_t] = \frac{|S(1,\dots,s)|}{|\Omega|}.$$

Go back
Full Screen
Close
Quit



If $|S(1,\dots,s)| = 0$, then $\Pr[A_1|\overline{A}_2\cdots\overline{A}_t] = 0 \le p$ and we are done. So we may assume that $|S(1,\dots,s)| \ne 0$.

$$\begin{aligned}\Pr[\overline{A}_2\cdots\overline{A}_t] &= \frac{\sum |S(y_1,\dots,y_s)|}{|\Omega|}\\ &\ge \frac{1}{|\Omega|}[n(n-1)\cdots(n-s+1)]|S(1,\dots,s)|.\end{aligned}$$

Thus

$$\Pr[A_1|\overline{A}_2\cdots\overline{A}_t] \le \frac{1}{n(n-1)\cdots(n-s+1)} = p.$$

□

5. Applications

We quote a version of the Lovász local lemma. For more details see [1].

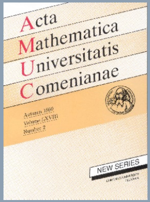
Lemma 2. *Let $A_1,\dots,A_\mu$ be events in a probability space Ω such that $\Pr[A_1] = \cdots = \Pr[A_\mu] = p$. Let G be a graph on $\{1,\dots,\mu\}$ such that each vertex in G has degree at most d. Suppose that $\Pr[A_i|\overline{A}_{j(1)}\cdots\overline{A}_{j(t)}] \le p$ whenever i is not adjacent to any of the vertices $j(1),\dots,j(t)$. Then $4dp \le 1$ implies $\Pr[\overline{A}_1\cdots\overline{A}_\mu] > 0$.*

Let us turn to the applications.

(a) In the $s = 2$ case $d = 2(m+n-2)(k-1)$, $p = 1/[n(n-1)]$. If $k-1 \le [n(n-1)]/[8(m+n-2)]$, then the $4dp \le 1$ condition holds and the Lovász local lemma guarantees the existence of a transversal. When $m = n$, this reduces to a result similar to that of Erdös and Spencer.

In the remaining part we consider only n by n arrays, that is, we will assume that $m = n$.

Go back
Full Screen
Close
Quit



(b) In the $s = 3$ case $d = (6n-9)(k-1)(k-2)/2$, $p = 1/[n(n-1)(n-2)]$. If

$$\frac{n(n-1)(n-2)}{2(6n-9)(k-1)(k-2)} \geq 1$$

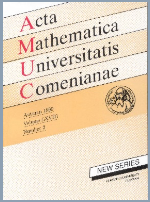
then the condition $4dp \leq 1$ holds and by the Lovász local lemma there is a section in which each symbol appears at most twice. We can say that for large n if each symbol appears at most $0.28n$ times in the table, then there is a section in which no symbol appears more than twice.

We would like to point out that P. J. Cameron and I. M. Wanless [2] show that every Latin square of order n contains a section in which no symbol occurs more than twice.

We single out one more special case. In this case each symbol appears at most n times in an n by n table. So the conditions are similar to the conditions of Stein's result described in the introduction.

(c) In the $s = 6$ case $d = (12n-36)(k-1)\cdots(k-5)/120$, $p = 1/[n(n-1)\cdots(n-5)]$. If $k = n$, then the condition $4dp \leq 1$ holds and by the Lovász local lemma there is a section in which each symbol appears at most 5 times.

Acknowledgement. Thanks for Professor Sherman K. Stein to the stimulating ideas and correspondence on the subject. Also thanks for the anonymous referee to the valuable suggestions.



1. Alon N. and Spencer J. H., *The Probabilistic Method*, 2nd Ed. John Wiley and Sons, Inc. 2000.
2. Cameron P. J. and Wanless I. M., *Covering radius for sets of permutations*, Discrete Math. **203** (2005), 91–109.
3. Erdös P., Hickerson D. R., Norton D. A. and Stein S. K., *Has every Latin square of order n a partial transversal of size $n-1$?*, Amer. Math. Monthly **95** (1988), 428–430.
4. Erdös P. and Spencer J. H., *Lopsided Lovász local lemma and latin transversals*, Discrete Appl. Math. **30** (1991), 151–154.
5. Ryser H. J., *Neuere Probleme der Kombinatorik*, im "Vorträge über Kombinatorik", Mathematischen Furschugsinstitute, Oberwolfach 1968.
6. Shor P. W., *A lower bound for the length of partial transversals in a latin square.* J. Combin. Theory Ser A. **33** (1982), 1–8.
7. Stein S. K., *Transversals of latin squares and their generalizations*, Pacific J. Math. **59** (1975), 567–575.

S. Szabó, Institute Mathematics and Informatics, University Pécs, Hungary,

e-mail: `sszabo7@hotmail.com`