Reprinted from JOURNAL OF COMBINATORIAL THEORY, Series A
Vol. 24, No. 1, January 1978
Printed in Belgium

Existence of Submatrices with All Possible Columns

J. MICHAEL STEELE

Department of Mathematics, University of British Columbia, Vancouver, Canada

Communicated by the Managing Editors

Received September 10, 1976

Let M be a matrix with entries from $\{1, 2, ..., s\}$ with n rows such that no matrix M' formed by taking k rows of M has s^k distinct columns. Let $f(k; n, s)$ be the largest integer for which there is an M with $f(k; n, s)$ distinct columns. It is proved that $f(k; n, s) = s^n - \sum_{j=k}^{n} \binom{n}{j} (s-1)^{n-j}$. This result is related to a conjecture of Erdös and Szekeres that any set of $2^{k-2} + 1$ points in R^2 contains a set of k points which form a convex polygon.

1. INTRODUCTION

The theorems provided in this note are motivated by questions like the following:

> Suppose an n set $x_1, x_2, ..., x_n$ is colored by s colors in m distinct ways. How large need m be to guarantee that there is a k set colored in all possible (i.e., s^k) ways? (1.1)

> Suppose that S is a class of subsets of a set X and that $\{x_1, x_2, ..., x_n\}$ is an n-element subset of X for which m of the sets $A \cap \{x_1, x_2, ..., x_n\}$, $A \in S$, are distinct. How large need m be to guarnatee that there is a k-element set $\{x_{i_1}, x_{i_2}, ..., x_{i_k}\} \subset \{x_1, x_2, ..., x_n\}$ for which there are 2^k distinct sets $A \cap \{x_{i_1}, x_{i_2}, ..., x_{i_k}\}$, $A \in S$? (1.2)

The first of these questions is new, but the second has been considered previously. It has in fact been solved quite precisely by Sauer [4] in response to a query of Erdös. An earlier independent solution was given in [5] in connection with a probabilistic application, but the result of [5] was not the best possible. In Section 2 of this note Theorem 2.1 gives a general result by a new method which implies these earlier results and covers the fresh ground indicated by question (1.1).

The third section gives a geometrical interpretation to a special case of Theorem 2.1, and shows the relationship of the present work to a long-standing conjecture of Erdös and Szekeres (see [1, p. xxi]).

0097-3165/78/0241-0084$02.00/0

2. Main Results

Let M be a matrix with entries from an s-symbol alphabet $\{1, 2,..., s\}$. Now let $f(k; n, s)$ be the largest integer such that there is a matrix M with n rows and $f(k; n, s)$ distinct columns such that no matrix M' formed by taking k of the rows of M has s^k distinct columns.

To note the relationship of $f(k; n, s)$ to question (1.1) one defines a correspondence between matrices and sets of colorings as follows: $M = (a_{ij})$, where $a_{ij} = b$ and b is the color of x_i in the jth coloring of $\{x_1, x_2,..., x_n\}$. For any subset of elements $\{x_{i_1}, x_{i_2},..., x_{i_k}\} \subset \{x_1, x_2,..., x_n\}$ there is a corresponding subset of k rows of M which forms a submatrix M'. Further, since any coloring of $\{x_{i_1}, x_{i_2},..., x_{i_k}\}$ corresponds to a column of M, the number of distinct colorings of $\{x_{i_1}, x_{i_2},..., x_{i_k}\}$ equals the number of distinct columns of M'. In the notation of (1.1) we therefore have $m = f(k; n, s) + 1$.

The main result can now be stated quite succinctly.

Theorem 2.1.

$$f(k; n, s) = s^n - \sum_{j=k}^{n} \binom{n}{j}(s-1)^{n-j}. \tag{2.1}$$

Proof. First it will be shown by construction that $f(k; n, s) \geqslant s^n - \sum_{j=k}^{n} \binom{n}{j}(s-1)^{n-j}$, and then the opposite inequality will be proved afterward by relating the general case to the first construction.

Define M to be the matrix consisting of all columns such that no column contains k or more ones. Since $\sum_{j=k}^{n} \binom{n}{j}(s-1)^{n-j}$ is precisely the number of columns with k or more ones, we see that M has $s^n - \sum_{j=k}^{n} \binom{n}{j}(s-1)^{n-j}$ columns. But since no k-row submatrix of M contains the column of all ones we have $f(k; n, s) \geqslant s^n - \sum_{j=k}^{n} \binom{n}{j}(s-1)^{n-j}$.

To obtain the opposite inequality we suppose that a matrix M has no k-row submatrix with s^k columns. To describe the columns which are missing from M, let $C_1, C_2,..., C_\tau$ where $\binom{n}{k} = \tau$ be a list of the k-element subsets of the row indices. For each $i = 1, 2,..., \tau$ there is a submatrix M_i formed by the C_i rows of M. Also by the hypothesis there is a k-vector v_i which is not a column of M_i. Now for each such v_i let Z_i be the set of columns of the $n \times s^n$ matrix which equal v_i when restricted to the index set C_i. Finally observe that none of the columns of $Z = \bigcup_{i=1}^{\tau} Z_i$ is a column of M.

If ν denotes the number of columns of M then $\nu \leqslant s^n - |\bigcup_{i=1}^{\tau} Z_i|$, (where $|\bigcup_{i=1}^{\tau} Z_i|$ denotes the number of the columns in the union $\bigcup_{i=1}^{\tau} Z_i$).

The proof will be completed by obtaining a lower bound on $|\bigcup_{i=1}^{\tau} Z_i|$. To do this we define a function on column vectors $w = (w_1, w_2,..., w_n)$ as follows:

$$\Phi(w) = w', \qquad \text{where } w' = (w_1', w_2',..., w_n') \tag{2.2}$$

and

$$\begin{aligned} w_j' &= 1 \qquad \text{if } w \in Z_i \text{ and } j \in C_i \text{ for some } i = 1, 2,\ldots, \tau, \\ &= w_j \qquad \text{otherwise.} \end{aligned} \tag{2.3}$$

The function Φ has several elementary but valuable properties which we first note and then prove:

$$|\Phi(Z)| \leqslant |Z| \qquad \text{for } Z = \bigcup_{i=1}^{\tau} Z_i . \tag{2.4}$$

$\Phi(Z_i)$ contains all columns of the $n \times s^n$ matrix which when restricted to C_i equal the k-column vector (1, 1,..., 1). (2.5)

$\Phi(Z)$ contains all n-columns which contain k or more ones. (2.6)

$$|\Phi(Z)| \geqslant \sum_{j=k}^{n} \binom{n}{j}(s-1)^{n-j}. \tag{2.7}$$

The proof of (2.4) is immediate since Φ is a function, and (2.5) is just a consequence of (2.3). To prove (2.6) note that if w has k or more ones, then there is a C_i, restricted to to which w has all ones, and hence $w \in \Phi(Z_i)$, by (2.3) and the definition of Z_i. Finally (2.7) comes from (2.6) and easy counting.

The last calculation is that

$$\nu \leqslant s^n - |Z| \leqslant s^n - |\Phi(z)| \leqslant s^n - \sum_{j=k}^{n} \binom{n}{j}(s-1)^{n-j}, \tag{2.8}$$

which completes the proof.

The preceding method also permits a precise understanding of those extreme matrices which lack k-row submatrices with a complete column set. Such matrices are characterized by a "missing" column vector.

THEOREM 2.2. *Suppose M is an n-row matrix with $s^n - \sum_{j=k}^{n} \binom{j}{n}(s-1)^{n-j}$ distinct columns and which has no k-row submatrix with s^k distinct columns. Then there is an n vector v such that for each column w of M one has $w_i \neq v_i$ for at least k values of the index i.*

Proof. In the notation of the previous proof, we note that if there is no v as required above then there are v_i and v_j such that $C_i \cap C_j \neq \varnothing$ yet v_i and v_j are not equal on $C_i \cap C_j$. By the definition of Φ and Z_i we therefore have $|\Phi(Z_i \cup Z_j)| < |Z_i \cup Z_j|$. Consequently, we have $|\Phi(Z)| < |Z|$.

But, since M has $s^n - \sum_{j=k}^{n} \binom{n}{j}(s-1)^{n-j}$ distinct columns, we note $|Z| = \sum_{j=k}^{n} \binom{n}{j}(s-1)^{n-j}$. However, by (2.7) we know $|\Phi(Z)| \geqslant \sum_{j=k}^{n} \binom{n}{j}(s-1)^{n-j}$ so the inequality $|\Phi(Z)| < |Z|$ yields a contradiction.

3. Relevance to a Famous Conjecture

Is it true that out of every $2^{k-2}+1$ points in the plane one can always select k points so that they form a convex n-sided polygon? This problem, posed in the winter of 1932–1933, published in 1935, promulgated daily, is still unsolved for $k \geqslant 6$ [1, pp. xxi, 42; 2; 3].

The results of Section 2 are relevant to this conjecture of Erdös and Szekeres, since they provide a sufficient condition that a set contain a convex polygon.

To see this let X be the plane and S the class of convex subsets of X. Next define

$$\Delta(x_1, x_2, \ldots, x_n) = |\{\{x_1, x_2, \ldots, x_n\} \cap A;\ A \in S\}| \tag{3.1}$$

that is, $\Delta(x_1, x_2, \ldots, x_n)$ is the number of subsets $\{x_{i_1}, x_{i_2}, \ldots, x_{i_t}\} \subset \{x_1, x_2, \ldots, x_n\}$ such that $\{x_{i_1}, x_{i_2}, \ldots, x_{i_t}\} = \{x_1, x_2, \ldots, x_n\} \cap A$ for some $A \in S$. Let A_j, $j = 1, 2, \ldots, \Delta(x_1, x_2, \ldots, x_n)$, be elements of S such that each of the sets $\{x_1, x_2, \ldots, x_n\} \cap A_j$ is distinct. These A_j define a $n \times \Delta(x_1, x_2, \ldots, x_n)$ matrix as follows:

$$\begin{aligned} a_{ij} &= 1 \quad \text{if} \quad x_i \in A_j, \\ &= 0 \quad \text{if} \quad x_i \notin A_j. \end{aligned} \tag{3.2}$$

By the definition of the A_j we know that $M = (a_{ij})$ has $\Delta(x_1, x_2, \ldots, x_n)$ distinct columns so

$$\Delta(x_1, x_2, \ldots, x_n) \leqslant f(k; n, 2) \tag{3.3}$$

unless M has k rows which have 2^k distinct columns. But since $\Delta(x_{i_1}, x_{i_2}, \ldots, x_{i_k}) = 2^k$ if and only if the set $\{x_{i_1}, x_{i_2}, \ldots, x_{i_k}\}$ forms a convex polyhedron, we have proved the following:

Theorem 3.1. *A sufficient condition that the set* $\{x_1, x_2, \ldots, x_n\} \subset R^2$ *contains k points which form a convex polygon is that*

$$\Delta(x_1, x_2, \ldots, x_n) > \sum_{j=0}^{k-1} \binom{n}{j}. \tag{3.4}$$

To prove the Erdös–Szekeres conjecture it thus suffices to show that (3.4) holds when $n = 2^{k-2}+1$. Of course, condition (3.5) has only been proved sufficient and quite possibly the Erdös–Szekeres conjecture can be true without (3.4) being met. Still, there are several possible uses of $\Delta(x_1, x_2, \ldots, x_n)$ in this problem and (3.4) pinpoints the most direct one.

To gain another view of Theorem 3.1 one should note that it is possible to give a more geometrical proof which avoids invoking the full strength of Theorem 2.1. For this proof, suppose $B \in \{\{x_1, x_2, \ldots, x_n\} \cap A \colon A \in S\}$

and let ∂B denote the subset of B equal to the elements of B on the boundary of the convex hull of B. We note that $|\partial B| \leqslant k-1$ if $\{x_1, x_2, ..., x_n\}$ contains no k-element convex polygon, since, indeed, ∂B is convex polygon. Next note that there are precisely $\sum_{j=0}^{k-1} \binom{n}{j}$ subsets of $\{x_1, x_2, ..., x_n\}$ with fewer than k elements. Since ∂B uniquely determines B we have

$$\Delta(x_1, x_2, ..., x_n) \leqslant \sum_{j=0}^{k-1} \binom{n}{j} \tag{3.5}$$

unless $\{x_1, x_2, ..., x_n\}$ contains a k-element subset which forms a convex polygon. This completes a second proof of Theorem 3.1.

4. A Closely Related Problem

In connection with the results given here and the Erdös–Szekeres conjecture the following question seems quite interesting:

> What is the minimum value of $\Delta(x_1, x_2, ..., x_n)$ given that $\{x_1, x_2, ..., x_n\}$ contains a k-set which forms a convex polygon? (4.1)
> (The x_i are assumed noncolinear.)

If this value is called $g(n, k)$, it is trivial that $g(n, k) \geqslant 2^k$, but a substantial improvement on this seems difficult. Still, by consideration of this problem it may be possible to make progress of the yet unreachable conjecture of Erdös and Szekeres.

References

1. P. Erdös, "The Art of Counting" (J. Spencer, Ed.), MIT Press, Cambridge, Mass., 1973.
2. P. Erdös and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* **2** (1935), 463–470.
3. P. Erdös and G. Szekeres, On some extremum problems in elementary geometry, *Ann. Univ. Sci. Budapest. Eötvös. Sect. Math.* **3-4** (1960–1961), 53–62.
4. N. Sauer, On the density of families of sets, *J. Combinational Theory A* **13** (1972), 145–147.
5. V. N. Vapnik and A. Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theor. Probability Appl.* **16** (1971), 264–280.

Printed by the St Catherine Press Ltd., Tempelhof 37, Bruges, Belgium.