COLLOQUIUM MATHEMATICUM

VOL. XXIII 1971 FASC. 1

MINIMIZATION OF ± 1 *MATRICES UNDER LINE SHIFTS*

BY

THOMAS A. BROWN AND JOEL H. SPENCER (SANTA MONICA, CALIF.)

Consider the following device, which is due to E. R. Berlekamp of Bell Labs and Professor David Gale (University of California, Berkeley): an $n \times n$ array of lights is controlled by $2n$ switches, one for each row and one for each column. When a switch is thrown, all lights in the corresponding column which are off turn on, and those which are on turn off. This device may be thought of as a code with $2n-1$ information bits and minimum Hamming distance n. It is easy to enumerate the codewords by weight (i.e., the number of ways of getting K lights on starting from an all-off position), but it is not so easy to find the distribution of co-set leaders (i.e., the minimum number of lights left on starting from an arbitrary position). Finding co-set leaders for $n = 15$ was the object of a computer experiment of Gleason [1]. Moon and Moser [3] showed that, as n becomes large, most co-set leaders have weights close to $n^2/2$. The purpose of this note is to derive upper bounds on the weights of co-set leaders.

It is convenient to consider the array of lights as a matrix of ± 1s ($+1$: light on; -1: light off). Then throwing a switch simply corresponds to multiplying the corresponding row or column by -1. In this guise the problem we are addressing was posed by Tusnády and Van Lint. There is no inherent reason to restrict our attention to square matrices.

Let $C(m, n)$ denote the set of all $m \times n$ matrices with entries ± 1. Let $A = [a_{ij}] \in C(m, n)$. We are allowed to choose a subset of rows and change the signs of all the entries in these rows. We are also allowed to do the same with the columns. More formally we define

$$A \cong B \quad \text{iff } B = D_m A D_n,$$

where D_m, D_n are square diagonal matrices of orders m, n and entries ± 1 on the diagonal. Clearly $\cong$ is an equivalence relation. Let $\{A\}$ denote the equivalence class containing A. Set

$$d(A) = \sum_{i,j} a_{ij},$$

$$l(A) = \tfrac{1}{2}[d(A) + mn] = \{\#\text{ of } a_{ij} = +1\}.$$

We wish to find

$$g(m, n) = \max_{A \in C(m,n)} \min_{B \in \{A\}} l(B),$$

$$f(n) = g(n, n)$$

Komlós and Sulyok [2] found (for m, n sufficiently large) the value of

$$\max_{A \in C(m,n)} \min_{B \in \{A\}} |d(A)|.$$

They observed that the moment method (which we apply below) shows that for some c_1, c_2

$$\frac{n^2}{2} - c_1 n^{3/2} \leqslant f(n) \leqslant \frac{n^2}{2} - c_2 n^{3/2}.$$

Our main result is the following:

THEOREM.

$$\frac{n^2}{2} - \frac{n^{3/2}}{2} + o(n^{3/2}) \leqslant f(n) \leqslant \frac{n^2}{2} - \frac{n^{3/2}}{\sqrt{2\pi}} + o(n^{3/2}).$$

Part I. $f(n) \leqslant \dfrac{n^2}{2} - \dfrac{n^{3/2}}{\sqrt{2\pi}} + o(n^{3/2})$.

Proof. Define $r_i(A) = \sum_j a_{ij}$, $s_i(A) = |r_i(A)|$, and $l_i(A) = \frac{1}{2}[r_i(A) + n] = [\#$ of $a_{ij} = +1$ in the ith row$]$. Fix $A \in C(n, n)$. Let σ range over the 2^n possible column shifts. Then, regardless of A, the i-th row of A runs over all 2^n possible values. Thus

$$E_\sigma\big(s_i(A^\sigma)\big) = E(X),$$

where X is the distance from the origin after a random walk of n steps of ± 1. By a straightforward integration involving the normal curve we find that

$$E_\sigma\big(s_i(A^\sigma)\big) = \sqrt{n}\sqrt{\frac{2}{\pi}} + o(n^{1/2}).$$

Thus

$$E_\sigma\Big(\sum_{i=1}^{n} s_i(A^\sigma)\Big) = n^{3/2}\sqrt{\frac{2}{\pi}} + o(n^{3/2}).$$

So for some fixed σ

$$\sum_{i=1}^{n} s_i(A^\sigma) \geqslant n^{3/2}\sqrt{\frac{2}{\pi}} + o(n^{3/2}).$$

Upon A^σ we then apply operation τ, a shifting of each row to achieve the minimal number of $+$1s in that row. We have

$$\begin{aligned} l_i(A^{\sigma\tau}) &= \min[l_i(A^\sigma), n - l_i(A^\sigma)] \\ &= \tfrac{1}{2}[n - s_i(A^\sigma)] \end{aligned}$$

Thus

$$\begin{aligned} l(A^{\sigma\tau}) = \sum_{i=1}^{n} l_i(A^{\sigma\tau}) &= \sum_{i=1}^{n} \tfrac{1}{2}[n - s_i(A^\sigma)] \\ &\leqslant \frac{n^2}{2} - \frac{n^{3/2}}{\sqrt{2\pi}} + o(n^{3/2}). \end{aligned}$$

(Note: $1/\sqrt{2\pi} \sim .4$)

Part II. $\dfrac{n^2}{2} - \dfrac{n^{3/2}}{2} + o(n^{3/2}) \leqslant f(n)$.

Proof. A Hadamard matrix of order n is an $n \times n$ matrix H with entries ± 1 such that

$$HH^T = nI,$$

where I is the idetity matrix. Hadamard matrices have been studied extensively and we shall only use some of the elementary properties. We note that

$$(D_m H D_n)(D_m H D_n)^T = D_m H D_n D_n^T H^T D_m^T = nI,$$

so that $\{H\}$ consists solely of Hadamard matrices. First we show: If H is a Hadamard matrix of order n, $l(H) \geqslant n^2/2 - n^{3/2}/2$. For consider the rows of H as vectors $\vec{u}_1, \ldots, \vec{u}_n \in R^n$. They are mutually orthogonal of length $\sqrt{n}$. Set $\vec{v} = (1, 1, \ldots, 1, 1) \in R^n$. To find a lower bound $l(H)$ we first find a lower bound for

$$\sum_{i=1}^{n} \vec{v} \cdot \vec{u}_i.$$

Multiplying $\vec{v}$ and all the $\vec{u}_i$ by $n^{-1/2}$ affects this sum by a factor of n^{-1}. By a rotation of R^n we send

$$\begin{aligned} &\vec{u}_i \to \vec{\epsilon}_i = (0, 0, \ldots, 0, 1, 0, \ldots, 0) \quad (1 \text{ in } i\text{-th position}), \\ &\vec{v} \to \vec{v}^* = (v_1, v_2, \ldots, v_n) \quad |\vec{v}^*| = 1. \end{aligned}$$

By elementary calculus the minimum value of $\sum \vec{v}^* \cdot \vec{\varepsilon} = \sum v_i$ subject to the condition $|\vec{v}^*| = 1$ is achieved when $v_i = -n^{-1/2}$ for all i. Thus

$$\sum_{i=1}^{n} \vec{v}^* \cdot \vec{\varepsilon}_i \geqslant -\sqrt{n},$$

$$d(H) = \sum_{i=1}^{n} \vec{v} \cdot \vec{u}_i \geqslant -n^{3/2},$$

and so

$$l(H) \geqslant \frac{n^2}{2} - \frac{n^{3/2}}{2}.$$

This would complete the proof of Part II except for the fact that Hadamard matrices do not exist for all orders. It is known that Hadamard matrices of order $4^i 12^j$ exist. By the theory of simultaneous approximations, given $\delta > 0$ we can find an n_0 such that, if $n > n_0$, then there exists a pair of integers (i, j) such that $0 < n - 4^i 12^j < \delta n$. Thus we can find matrices of any sufficiently high order which are "almost" Hadamard. But in order to complete our proof we must show that the non-Hadamard portions of these matrices make a substantial contribution to the number of $+1$s under any combination of row and column shifts. To do this we employ a method used by Moon and Moser [3].

Consider an $s \times t$ matrix A, where $s \leqslant t$. It is not difficult to show that there are 2^{s+t-1} matrices in the equivalence class of A. Thus there are $2^{st-(s+t-1)}$ equivalence classes altogether. Each contains a matrix B such that $l(B) \leqslant g(s, t)$. Thus

$$2^{st-(s+t-1)} \leqslant \sum_{i=0}^{g(s,t)} \binom{st}{i}.$$

An application of Stirling's formula shows that

$$g(s, t) \geqslant \frac{st}{2} - \sqrt{st(s+t)\frac{\log 2}{2}} + o(s^{1/2}t)$$

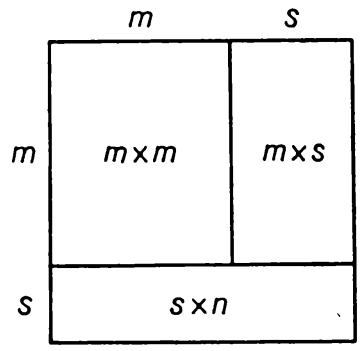
$$\geqslant \frac{st}{2} - ts^{1/2}\sqrt{\log 2} + o(s^{1/2}t).$$

Thus there exists a constant K such that

$$g(s, t) = g(t, s) \geqslant \frac{st}{2} - Ks^{1/2}t \quad \text{for all } t \geqslant s.$$

Given an $\varepsilon > 0$ pick n_0 such that for $n > n_0$ there exists an m such that an $m \times m$ Hadamard matrix exists and $m + s = n$, where $0 < s < \varepsilon^2 n/4K$.

Imagine constructing an $n \times n$ matrix from three "pieces", as indicated below:



Clearly

$$f(n) \geqslant f(m)+g(m, s)+g(s, n).$$

Applying our results above gives

$$f(n) \geqslant \frac{m^2}{2} - \frac{m^{3/2}}{2} + \frac{ms}{2} - Kms^{1/2} + \frac{ns}{2} - Kns^{1/2}$$

$$\geqslant \frac{(m+s)^2}{2} - \frac{m^{3/2}}{2} - K(m+n)s^{1/2}$$

$$\geqslant \frac{n^2}{2} - \frac{n^{3/2}}{2} - \varepsilon n^{3/2}.$$

This concludes the proof of part II.

Applying these methods we have found

$$f(1) = 0, \quad f(4) = 4, \quad 32 \leqslant f(10) \leqslant 37,$$
$$f(2) = 1, \quad f(5) = 7, \quad 72 \leqslant f(15) \leqslant 89.$$
$$f(3) = 2, \quad 22 \leqslant f(8) \leqslant 23,$$

We now study the values of $g(m, n)$ for n fixed and $m \to +\infty$.

THEOREM.

$$g(2^{n-1}, n) = \begin{cases} \sum_{i=0}^{k} i\binom{n}{i} & \text{if } n = 2k+1, \\ \sum_{i=0}^{k-1} i\binom{n}{i} + \frac{k}{2}\binom{n}{k} & \text{if } n = 2k. \end{cases}$$

Proof. We may think of an $A \in C(2^{n-1}, n)$ as corresponding to an ordered family of 2^{n-1} (possibly non-distinct) subsets of n-element set. In particular, $A \leftrightarrow (S_1, \ldots, S_{2n-1})$, where $S_j = \{i: a_{ij} = +1\}$. Define A by letting the 2^{n-1} sets run over all S, $|S| \leqslant k$ if $n = 2k+1$ and if $n = 2k$ all S, $|S| < k$ and all S, $|S| = k$, $1 \in S$.

A column shift preserves the family of S, except possibly if $n = 2k$ shifting from those S with $|S| = k$, $1 \in S$ to those with $1 \notin S$. Looking at Part I of our main theorem we note that $\sum_{i=1}^{n} s_i(A^\sigma)$ is independent of σ. As $E[\sum_{i=1}^{n} s_i(A^\sigma)]$ is independent of A this particular A gives the minimal possible $\max_\sigma \sum_{i=1}^{n} s_i(A^\sigma)$ and thus yields the minimal $g(n, 2^{n-1})$ which is as desired, q.e.d.

By the same argument

$$g(2^{n-1}k, n) = kg(2^{n-1}, n) \quad \text{for all integral } k \geqslant 1.$$

Setting

$$a_{n,p}(K) = g(2^{n-1}K+p, n) - Kg(2^{n-1}, n),$$

we can show

$$g(p, n) \leqslant a_{n,p}(K) \leqslant g(2^{n-1}, n)$$

and

$$a_{n,p}(K) \leqslant a_{n,p}(K+1),$$

As a is integral valued there exists $b_{n,p}$ such that

$$a_{n,p}(K) = b_{n,p} \quad \text{for all } K \geqslant K_{n,p}.$$

Thus, for fixed n, $g(m, n)$ is given for all sufficiently large m by 2^{n-1} linear equations of the form

$$g(2^{n-1}K+p, n) = Kg(2^{n-1}, n) + b_{n,p}.$$

For example

$$\begin{aligned}
g(2K, 2) &= K, \\
g(2K+1, 2) &= K, \\
g(4K, 3) &= 3K, \\
g(4K+1, 3) &= 3K, \\
g(4K+2, 3) &= 3K+1, \\
g(4K+3, 3) &= 3K+2, \\
g(8K, 4) &= 10K, \\
g(8K+1, 4) &= 10K, \\
g(8K+2, 4) &= 10K+2, \\
g(8K+3, 4) &= 10K+3, \\
g(8K+4, 4) &= 10K+4, \\
g(8K+5, 4) &= 10K+5, \\
g(8K+6, 4) &= 10K+7, \\
g(8K+7, 4) &= 10K+8.
\end{aligned}$$

These equations hold for all $K \geqslant 0$. It is possible that in general $g(2^{n-1}K+p, n) = Kg(2^{n-1}, n) + g(p, n)$ but we have not been able to prove this (**P 733**).

REFERENCES

[1] A. Gleason, *A search problem in the n-cube*, Proceedings of the Symposium on Applied Mathematics 10, p. 175 - 187.

[2] E. Komlós and M. Sulyok, *On the sum of elements of* ± 1 *matrices*, Proceedings of the Combinatorial Colloquium in Balatonfüred, Hungary, 1969.

[3] J. W. Moon and L. Moser, *An extremal problem in matrix theory*, Matematički Vesnik 3 (18) (1966), p. 209 - 211.

Reçu par la Rédaction le 11. 2. 1970