ON A CERTAIN SET OF LINEAR INEQUALITIES

J.G. Kalbfleisch and R.G. Stanton

(received April 20, 1967)

1. Introduction. In this paper we shall discuss the following set of $n+1$ linear inequalities:

$$
(1)\quad
\begin{array}{l}
y_0 + y_1 \geq \binom{n}{0} \\
ny_0 + y_1 + 2y_2 \geq \binom{n}{1} \\
(n-1)y_1 + y_2 + 3y_3 \geq \binom{n}{2} \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
2y_{n-2} + y_{n-1} + ny_n \geq \binom{n}{n-1} \\
y_{n-1} + y_n \geq \binom{n}{n}
\end{array}
$$

If we let $Y = (y_i)$, $C_n = (\binom{n}{i})$, and $Z = (z_i)$ $(i = 0, 1, \ldots, n)$ be $(n+1)$-dimensional column vectors, and define the $n+1$ by $n+1$ tridiagonal matrix $D_n(\phi)$ by

$$
D_n(\phi) = \begin{bmatrix}
\phi & 1 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\
n & \phi & 2 & 0 & 0 & \ldots & 0 & 0 & 0 \\
0 & n-1 & \phi & 3 & 0 & \ldots & 0 & 0 & 0 \\
0 & 0 & n-2 & \phi & 4 & \ldots & 0 & 0 & 0 \\
\cdots & & & & & & & & \cdots \\
0 & 0 & 0 & 0 & 0 & \ldots & \phi & n-1 & 0 \\
0 & 0 & 0 & 0 & 0 & \ldots & 2 & \phi & n \\
0 & 0 & 0 & 0 & 0 & \ldots & 0 & 1 & \phi
\end{bmatrix},
$$

the set of inequalities (1) may be written

$$(2)\qquad A_n Y = C_n + Z$$

where $A_n = D_n(1)$ and $z_i \geq 0$ $(i = 0, 1, \ldots, n)$. In sections 2 and 3, we

Canad. Math. Bull. vol. 11, no. 5, 1968.

https://doi.org/10.4153/CMB-1968-082-6 Published online by Cambridge University Press

consider real solutions of (2), and give expressions for the solution Y corresponding to a specified vector Z of slack variables. The inequalities (1) arise in connection with a current investigation of some covering properties of groups [4], where it is necessary to find all solutions of (1) in non-negative integers y_i with Σy_i specified. In section 4, we give an efficient algorithm for obtaining such solutions.

2. Some properties of the matrix A_n.

We begin by considering the set of $n+1$ linear equations

(3) $$D_n(\phi) X = 0$$

where $X = (x_i)$ $(i = 0, 1, \ldots, n)$ is a column vector. (3) may be written

(4) $$(n+1-i)x_{i-1} + \phi x_i + (i+1)x_{i+1} = 0, \quad -\infty < i < \infty,$$

with boundary conditions

(5) $$x_i = 0 \text{ if } i < 0 \text{ or } i > n.$$

Multiplying (4) by t^i and summing over all i gives

(6) $$nt\, G(t) - t^2 G'(t) + \phi\, G(t) + G'(t) = 0$$

where $G(t) = \Sigma x_i t^i$. The solution of (6) with $G(0) = x_0 = 1$ is

(7) $$G(t) = (1+t)^{\frac{n-\phi}{2}} (1-t)^{\frac{n+\phi}{2}} = \sum_{i=0}^{\infty} t^i \sum_{r=0}^{i} (-1)^r \binom{(n+\phi)/2}{r}\binom{(n-\phi)/2}{i-r}.$$

It follows that

(8) $$x_i = \sum_{r=0}^{i} (-1)^r \binom{(n+\phi)/2}{r} \binom{(n-\phi)/2}{i-r}.$$

The boundary conditions (5) are satisfied if and only if $n+\phi$ is an even integer and $-n \le \phi \le n$; thus (3) has a nontrivial solution X if and only if

(9) $$\phi = \phi_j = n - 2j \quad (j = 0, 1, \ldots, n).$$

The solution vector $X_j = (x_{ij})$ corresponding to ϕ_j is given by

$$(10) \qquad x_{ij} = \sum_{r=0}^{i} (-1)^r \binom{n-j}{r} \binom{j}{i-r}.$$

We may now determine the eigenvalues and eigenvectors of A_n; for the equations $A_n X = \lambda X$ may be written $D_n(\phi) X = 0$ with $\phi = 1 - \lambda$. Thus A_n has eigenvalues

$$(11) \qquad \lambda_j = 1 - \phi_j = 1 - n + 2j \quad (j = 0, 1, \ldots, n)$$

and the eigenvector corresponding to λ_j is $X_j = (x_{ij})$. It also follows that det $A_n = \pi \lambda_j$, which is zero for n odd, and equal to $(-1)^{n/2}(n+1)(n-1)^2(n-3)^2 \ldots 3^2 . 1^2$ for n even. Det A_n may also be obtained simply by direct expansion and recursion. In fact, the determinant of $D_n(\phi)$ was evaluated by J. J. Sylvester [5] as early as 1854.

Let $X = (X_0 | X_1 | \ldots | X_n)$ be the $n+1$ by $n+1$ modal matrix whose columns are the eigenvectors $X_0, X_1, \ldots, X_n$. The matrix X was computed for small values of n, and it was noted that each row of X is orthogonal to all but one of the columns of X. Consequently the following lemma was obtained.

LEMMA. <u>Let</u> $S_{ij} = \sum_{\alpha=0}^{n} x_{i\alpha} x_{\alpha j}$, <u>where</u> x_{ij} <u>is defined by</u> (10). <u>Then</u>

$$(12) \qquad S_{ij} = 0 \text{ if } j \neq n-i; \; S_{i,n-i} = (-1)^i 2^n.$$

<u>Proof</u>: Put $H_j(u) = \sum_i S_{ij} u^i$. Then, using the generating function

$$G_j(t) = \sum_i x_{ij} t^i = (1+t)^j (1-t)^{n-j}$$

corresponding to ϕ_j we get

https://doi.org/10.4153/CMB-1968-082-6 Published online by Cambridge University Press

$$H_j(u) = \sum_i \sum_k x_{ik} x_{kj} u^i = \sum_k x_{kj} (1+u)^j (1-u)^{n-j}$$

$$= (1-u)^n \left(1 + \frac{1+u}{1-u}\right)^j \left(1 - \frac{1+u}{1-u}\right)^{n-j} = 2^n (-u)^{n-j} .$$

By comparing coefficients of powers of u we obtain (12).

Thus the i^{th} row of X is orthogonal to every column of X except the $(n-1)^{th}$ $(i = 0, 1, \ldots, n)$, and

(13) $$X^{-1} = 2^{-n}(X_n | -X_{n-1} | X_{n-2} | -X_{n-3} | \ldots | (-1)^n X_0).$$

Let Λ be the diagonal matrix with diagonal entries $\lambda_0, \lambda_1, \ldots, \lambda_n$, so that $A_n X = X\Lambda$. Then if n is even,

(14) $$A_n^{-1} = X\Lambda^{-1} X^{-1} .$$

We are indebted to Professor D. A. Sprott for pointing out a connection with probability theory. If n is even, $\frac{1}{n} D_n(0)$ is the matrix of transition probabilities for the Ehrenfest Model. A discussion of some of the properties of this matrix appears in [1] and [3]. The arguments given here are somewhat simpler than theirs because the Lemma makes the derivation of X^{-1} almost trivial.

3. Real Solutions of $A_n Y = C_n + Z$. If n is even, A_n is nonsingular, and there will be a unique solution Y of (2) corresponding to each vector Z. Since $C_n = X_n$, the eigenvector corresponding to $\lambda_n = n+1$, we have

(15) $$Y = A_n^{-1}(C_n + Z) = \frac{1}{n+1} C_n + A_n^{-1} Z .$$

If the slack variables z_i are all non-negative, Y is a solution of (1).

If n is odd, say $n = 2k+1$, A_n is singular and the situation is slightly more complicated. Since $X_0, X_1, \ldots, X_n$, are linearly independent, every vector Y may be expressed as a linear combination of them. If $Y = \Sigma a_j X_j$ is a solution,

$$A_n Y = \Sigma a_j A_n X_j = \Sigma a_j \lambda_j X_j = C_n + Z .$$

Since $C_n = X_n$ and $\lambda_k = 0$, Z must lie in the subspace spanned by $X_0, X_1, \ldots, X_{k-1}, X_{k+1}, \ldots, X_n$. But, by the Lemma,

$$\sum_\alpha x_{k+1, \alpha} x_{\alpha i} = 0, \quad i \neq k ,$$

and therefore the slack variables z_α must satisfy

$$\sum_\alpha z_\alpha x_{k+1, \alpha} = 0 . \tag{16}$$

If Z is any vector satisfying (16), then $Z = \Sigma b_j X_j$ with $b_k = 0$, and corresponding to Z are solutions

$$Y = \frac{1}{n+1} C_n + \sum_{j \neq k} \frac{b_j}{\lambda_j} X_j + t X_k \tag{17}$$

where t is an arbitrary real number. Let Λ^* be the diagonal matrix with diagonal entries $\lambda_0, \lambda_1, \ldots, \lambda_{k-1}, 1, \lambda_{k+1}, \ldots, \lambda_n$, and let B be the column vector (b_j). Then $Z = XB$, and thus $B = X^{-1} Z$. We may now rewrite (17) in the form

$$Y = \frac{1}{n+1} C_n + X \Lambda^{*-1} X^{-1} Z + t X_k . \tag{18}$$

If all the slack variables z_α are non-negative, Y is a solution of (1).

A convenient algorithm for numerical computation of real solutions Y of the equations $AY = B$, where Y and B are column vectors and A is an arbitrary nonsingular tridiagonal matrix, is described by Henrici [2, page 350]. This algorithm depends upon a rather interesting factorization of a tridiagonal matrix into two "bidiagonal" matrices.

4. Solutions of (1) in non-negative integers. The inequalities (1) arise in connection with a covering problem [4] where y_α represents the number of elements of a certain type in a covering set. Consequently

https://doi.org/10.4153/CMB-1968-082-6 Published online by Cambridge University Press

the y_α's and z_α's must be non-negative integers. The problem is to construct a covering set with as few members as possible - that is, with Σy_α as small as possible. Adding the equations (2) gives

$$(n+1)\Sigma y_\alpha = 2^n + \Sigma z_\alpha$$

so that $\Sigma y_\alpha \geq 2^n/(n+1)$. This is not, however, a sufficient condition for the existence of a solution of (1). Furthermore not every solution Y corresponds to a covering set. Thus it is often necessary to consider several totals Σy_α, beginning with the least integer greater than or equal to $2^n/(n+1)$. It is, however, necessary to consider only solutions with $y_0 = 1$ since any covering set will be isomorphic to one with $y_0 = 1$.

In this section we give an efficient algorithm for finding all solutions of (1) with $y_0 = 1$ and $\Sigma y_\alpha \leq m$. The latter condition is equivalent to insisting that the total slack Σz_α be at most $T = 2^n - m(n+1)$. The algorithm is easily programmed for a computer, and may be generalized to yield the non-negative integer solutions of many sets of linear inequalities whose matrices of coefficients are tridiagonal.

The algorithm is represented pictorially by the directed graph in Figure 1. Vertices represent operations, and edges indicate the order in which they are performed. First, we give the operations corresponding to the vertices and the rules for moving from one to another. Then we shall explain the algorithm and give an example.

S (start) : Put $y_{-1} = 0$, $y_0 = 1$, $z_0 = T$, $k = 0$. Go to A.

A: Calculate $R_k = \binom{n}{k} + z_k - y_k - (n-k+1)y_{k-1}$. If $R_k < 0$, go to C; if $R_k \geq 0$, go to B.

B: Calculate $y_{k+1} = [R_k/(k+1)]$, $z_k = (k+1)y_{k+1} + y_k + (n-k+1)y_{k-1} - \binom{n}{k}$ and $z_{k+1} = T - z_1 - z_2 - \ldots - z_k$. If $z_k < 0$, go to C; if $z_k \geq 0$ and $k < n-1$, go to G; if $z_k \geq 0$ and $k = n-1$, go to F.

C: Select the largest $j < k$ for which $z_j \neq 0$ and go to D. If $z_j = 0$ for all $j < k$, go to E.

https://doi.org/10.4153/CMB-1968-082-6 Published online by Cambridge University Press

D: Decrease z_j by 1 and put $z_{j+1} = T - z_0 - z_1 \dots - z_j$. Put

$k = j$ and go to A.

E: Algorithm terminates.

F: Put $z_n = y_{n-1} + y_n - 1$. Y is a solution with slack Z. Go to C.

G: Increase k by 1 and go to A.

The algorithm examines all possible slack vectors Z with $\Sigma z_\alpha \leq T$. We begin with $Z = (T, 0, 0, \dots, 0)$, and change Z in such a way that

$$||Z|| = z_0(T+1)^n + z_1(T+1)^{n-1} + \dots + z_{n-1}(T+1) + z_n$$

is steadily decreasing. Suppose that after a number of steps we reach A with $Z = (z_0, z_1, \dots, z_k, 0, \dots, 0)$, and have solved the first k equations of (2) for $y_0, y_1, \dots, y_k$. The k^{th} equation (numbering from 0) is

$$(k+1)y_{k+1} + y_k + (n-k+1)y_{k-1} = \binom{n}{k} + z_k ,$$

and we wish to solve this for y_{k+1}. Since z_k contains all the available slack, $(k+1)y_{k+1}$ cannot be larger than R_k. Thus if R_k is negative, we cannot solve for y_{k+1}, and must alter the values of $z_{k-1}, z_{k-2}, \dots$ (step C). If $R_k \geq 0$, we must choose y_{k+1} so that $||Z||$ will be as large as possible (step B). This means that we want as much slack as possible in the k^{th} equation, and so y_{k+1} must be chosen as large as possible. We must now recalculate z_k, which becomes the actual slack in the k^{th} equation, and place all remaining slack at z_{k+1}.

If the new value of z_k is negative, we have reached a contradiction and must go to C where the values of $z_{k-1}, z_{k-2}, \dots$ will be altered. If $z_k \geq 0$ and $k = n-1$, we have a solution Y (step F). The actual slack in the n^{th} equation will be $y_{n-1} + y_n - 1$, and the solution has total slack $T - z_n + y_{n-1} + y_n - 1$. To obtain other solutions we go to C where $z_{n-1}, z_{n-2}, \dots$ are changed. Finally, if $z_k \geq 0$ and $k < n-1$, we replace k by $k+1$ (step G), and then return to A to solve the next equation for the next Y value.

https://doi.org/10.4153/CMB-1968-082-6 Published online by Cambridge University Press

When we arrive at C from A, B, or F, we must alter the first k slack variables $z_0, z_1, \ldots, z_{k-1}$ in such a way that $||Z||$ decreases, but does so as little as possible. This is accomplished by decreasing by one the last nonzero z_j. Thus in step C we select the largest $j < k$ with $z_j \neq 0$. If no such j exists (as must eventually happen) the algorithm terminates at E. Otherwise (step D) we decrease z_j by one, put all remaining slack at z_{j+1}, and return to A. Since $z_0, z_1, \ldots, z_{j-1}$ have not been changed, $y_0, y_1, \ldots, y_j$ also remain fixed, and we begin our calculations with y_{j+1} - that is, with $k = j$ in step A.

As an example, we take $n = 5$, $\Sigma y_\alpha \leq 7$, and total allowable slack $T = 10$. The first column of Table 1 names the step, and the remaining columns give the values of k, R_k, j, Z, and Y at that step. Opposite a step we have entered only the values which are calculated at that step, and other variable values remain unchanged from the preceding step. The first few iterations are rather uninteresting and no solutions are obtained. We omit them, and begin at A with $z_0 = 2$, $z_1 = 8$, $y_0 = 1$, $y_1 = 2$, and $k = 1$. We give several steps during which a solution $Y = (1, 2, 2, 0, 1, 1)$, $Z = (2, 6, 0, 0, 1, 1)$ is obtained (F). If the algorithm is continued to completion, a total of seventeen solutions of (1) with $y_0 = 1$ and $\Sigma y_\alpha \leq 7$ are obtained, two with $\Sigma y_\alpha = 6$, and fifteen with $\Sigma y_\alpha = 7$. The computations required only a few seconds on the IBM 7040 computer.

Step	k	R_k	j	z_0	z_1	z_2	z_3	z_4	z_5	y_0	y_1	y_2	y_3	y_4	y_5
	1			2	8					1	2				
A		6													
B					8	0						3			
G	2														
A		-1													
C			1												
D	1				7	1									
A		5													
B					6	2						2			
G	2														
A		2													
B						0	2						0		
G	3														
A		6													
B							0	2						1	
C	4														
A		6													
B								1	1						1
F									1						
C			4												
D	4							0	2						
A		4													
B								-4	6						0
C			1												
D	1				5	3									
A		3													

Table 1

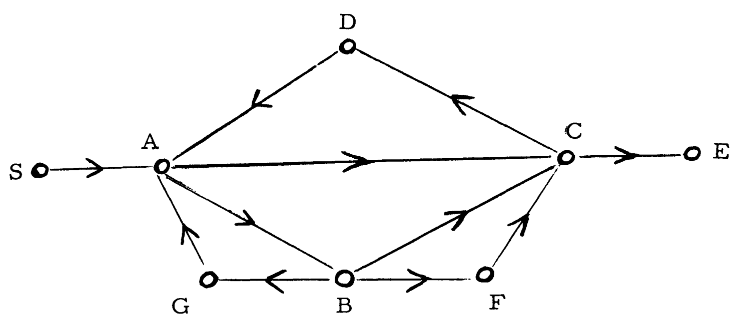

Figure 1

https://doi.org/10.4153/CMB-1968-082-6 Published online by Cambridge University Press

REFERENCES

1. D.R. Cox and H.D. Miller, The theory of stochastic processes. (Methuen and Co. Ltd., London, 1965) 129-132.

2. P. Henrici, Discrete variable methods in ordinary differential equations. (Wiley, 1962).

3. Mark Kac, Random walk and the theory of Brownian motion, American Mathematical Monthly 54 (1947), 369-391.

4. R.G. Stanton and J.G. Kalbfleisch, Covering problems for dichotomized matchings. Aequationes Mathematicae 1 (1968) 94-103.

5. J.J. Sylvester, Théorème sur les déterminants de M. Sylvester, Nouv. Annales de Math. 13 (1854) 305.

York University
Toronto

https://doi.org/10.4153/CMB-1968-082-6 Published online by Cambridge University Press