Discrete Mathematics 66 (1987) 303–307
North-Holland

NOTE

ON THE NONUNIFORM FISHER INEQUALITY

László BABAI
Department of Algebra, Eötvös University, Budapest, Hungary H-1088

and

Department of Computer Science, The University of Chicago, Chicago, IL 60637, U.S.A.

Received 18 July 1985
Revised 30 July 1986

Let $\mathscr{F}$ be a family of m subsets (*lines*) of a set of n elements (*points*). Suppose that each pair of lines has λ points in common for some positive λ. The Nonuniform Fisher Inequality asserts that under these circumstances $m \leqslant n$. We examine the case when $m = n$. We give a short proof of the fact that (with the exception of a trivial case) such an $\mathscr{F}$ must behave like a geometry in the following sense: a line must pass through each pair of points. This generalizes a result of de Bruijn and Erdös.

Dedicated to the memory of H.J. Ryser

A *hypergraph* is a pair $\mathscr{H} = (X, \mathscr{F})$ where $X = \{x_1, \ldots, x_n\}$ is a finite set of *points* and $\mathscr{F} = \{F_1, \ldots, F_m\}$ is a family of subsets of X called *lines*. We shall always use n and m to denote the numbers of points and lines, resp. We suggest Lovász' excellent book [7] as a general reference on hypergraphs.

Generalizing R.A. Fisher's celebrated inequality [4], Majumdar [8] and later Isbell [6] found the following result (cf. Lovász [7, Problem 13.15(b)]).

Theorem 1 (Nonuniform Fisher Inequality). *Let $\mathscr{H}$ be a hypergraph with n points and m lines. If every pair of lines has precisely λ points in common for some positive integer λ, then $m \leqslant n$.*

Fisher's original result is equivalent to this theorem for the case of dual BIBD's. (The dual of a hypergraph is obtained by interchanging the roles of points and blocks, preserving the incidence relation.) R.C. Bose found a very elegant proof of Fisher's inquality [1] which actually works for any uniform hypergraph. ($\mathscr{H}$ is *uniform* if $|F_1| = \cdots = |F_m|$). Bose's proof uses elementary linear algebra in a surprising way. Only a slight modification of his proof is needed in order to eliminate the uniformity conditions and thus obtain Theorem 1. No proof, avoiding this linear algebra trick, of Theorem 1 appears to be known. (Fisher's original proof [4] uses direct counting arguments and does not seem to generalize beyond the BIBD case.) We note, however, that an infinite version of Theorem 1 exists (Komjáth [15]).

The Nonuniform Fisher Inequality contains a result of de Bruijn and Erdös [3] as the particular case $\lambda = 1$. For this case, a "purely combinatorial" proof exists ([3], cf. [7, Problems 13.14 and 13.15(a)]). It has been quoted as an advantage of the proof avoiding linear algebra that it permits characterization of the extremal cases ($m = n$). As shown by de Bruijn and Erdös, in the case $\lambda = 1$, equality will hold in the following cases only.

(a) One of the lines is a singleton.

(b) $\mathscr{H}$ is a possibly degenerate projective plane.

In case (a), all the other lines must have 2 points.

A *possibly degenerate projective plane* is a hypergraph satisfying the following axioms:

(i) Through every pair of points there is a line;

(ii) Every pair of lines intersects in precisely one point;

(iii) There exists a triangle (three points not on a line).

We obtain the axioms of a *projective plane* if we strengthen (iii) to

(iv) There exists a quadrilateral (four points, no three of which are on a line).

A *degenerate* plane is one that satisfies (i), (ii), (iii) but not (iv). It is easy to show that there exists precisely one degenerate plane for every $n \geq 3$; it has one line with $n-1$ points and $n-1$ lines with two points each, the latter connecting the points of the long line to the point it misses.

So the crucial step in characterizing the extremal case ($m = n$) in the de Bruijn–Erdös theorem is proving that unless there is a singleton among the lines, the hypergraph must satisfy (i). (If there is no singleton there, it is straightforward to verify (iii).)

Although this fact follows from the results of Ryser [11] and Woodall [14], their arguments are certainly much more involved than those of de Bruijn and Erdös. The aim of this note is to demonstrate that it *is* possible to use the linear algebra method to verify (i) for the extremal hypergraphs in a very simple way. This approach permits a generalization to $\lambda > 1$; we shall prove that for arbitrary λ, if $m = n$ in the Nonuniform Fisher Inequality, then (i) holds.

Theorem 2 (Ryser, Woodall). *Let $\mathscr{H}$ be a hypergraph with n points and $m = n$ lines. Assume no line is a singleton. If every pair of lines has precisely λ points in common for some positive λ, then through every pair of points there is a line.*

If we add uniformity ($|F_1| = \cdots = |F_n| \stackrel{\text{def}}{=} k$) to the conditions of Theorem 2, then, by a result of H.J. Ryser [9], $\mathscr{H}$ must also be *regular* of degree k (every point must belong to precisely k lines) and there must be precisely λ lines through each pair of points. This means the hypergraph is a *symmetric design* [10, Ch. 8].

The nonuniform extremal configurations are called *λ-designs* [11]. Ryser [11] and Woodall [14] prove that the points of a λ-design have only two different degrees, r_1 and r_2, say. (The *degree* or "replication number" of a point is the number of lines through the point.) It is no longer true that every pair of points is linked by λ lines. An example of de Witte [13] (cf. [11]) shows that there exists a λ-design with $n = m = 7$, $\lambda = 2$ where some pairs of points are linked by 1, others by 2, still others by 3 lines.

Only one class of λ-designs is known ("point-complemented symmetric block designs" for $\lambda \geqslant 2$, cf. [2]). That all λ-designs are of this type is the *λ-design conjecture*. Bridges [2] shows that a λ-design is of this type if and only if every pair of points of different degrees is linked by precisely λ lines. Seress [16] proves the λ-design conjecture assuming $r_1 r_2 = \lambda(n-1)$. (It is known that $r_1 r_2 \leqslant \lambda(n-1)$.)

The proof of Theorem 2 is entirely self-contained and makes no reference to the results quoted above. It incorporates the (known) compact proof of Theorem 1 (see the fourth paragraph of the proof).

Proof of Theorem 2. Let us not assume for a moment that $m = n$. If there is a line with fewer than λ points then $m = 1$. If some line, say F_1, has precisely λ points, then the sets $F_i - F_1$ must be pairwise disjoint and therefore $m \leqslant n - \lambda + 1$. This is always $\leqslant n$ with equality only if $\lambda = 1$ and thus one of the lines is a singleton. Henceforth we assume that the numbers $l_i \stackrel{\text{def}}{=} |F_i| - \lambda$ are all positive.

Let M be the indicence matrix of $\mathscr{H}$. M is an $m \times n$ matrix; its columns correspond to the points and its rows to the lines. The entry $M[i, j]$ is 1 if point x_j belongs to line F_i.

The *intersection matrix* of $\mathscr{H}$ is the $m \times m$ matrix $A = MM^{\mathrm{T}}$. Our intersection conditions are summarized in the matrix equation

$$MM^{\mathrm{T}} = \lambda J + L, \tag{1}$$

where J is the $m \times m$ matrix with all entries equal to 1 and L is the diagonal matrix $L = \operatorname{diag}(l_1, \ldots, l_m)$.

Since both J and L are positive semidefinite Hermitian matrices and L is positive definite, A is positive definite as well. In particular A is nonsingular, therefore rank $M = m$, proving Theorem 1 ($m \leqslant n$).

Let us henceforth assume $m = n$. Notice that M is now a nonsingular $n \times n$ matrix.

Theorem 3. *Assume M is a nonnegative square matrix satisfying* (1), *where $\lambda > 0$ and L is a diagonal matrix with positive diagonal entries. Then all entries of $M^{\mathrm{T}}M$ are positive.*

Our objective is to prove that *all entries of the matrix $M^{\mathrm{T}}M$ are positive.*

We begin with three simple observations. All matrices below are n by n, with real entries.

Let B and C be two matrices. We write $B > C$ ($B \geqslant C$, resp.) if all entries of $B - C$ are positive (nonnegative, resp.).

Observation 1. *If $B > 0$, $C \geqslant 0$ and no column of C is zero then $BC > 0$. If no row of C is zero, then $CB > 0$.*

Observation 2. *Any matrix of the form BJ can be written as $BJ = DJ$, where $D = \operatorname{diag}(d_1, \ldots, d_n)$ is a diagonal matrix.*

(As before, J is the all-ones matrix.)

Proof. Make d_i equal to the sum of the ith row of B. □

Observation 3. *Let B and C be arbitrary $n \times n$ matrices. If the matrices JC and BJC are positive, then so is BJ.*

Proof. By Observation 2, $BJ = DJ$ for some diagonal matrix $D = \operatorname{diag}(d_1, \ldots, d_n)$. We have to prove that all the d_i are positive. Let r denote a row of JC. Then the ith row of $BJC = DJC$ is $d_i r$. Since both JC and BJC are positive, we infer $d_i > 0$. □

We proceed with the proof of Theorem 2.

By (1) we have

$$M^{\mathrm{T}} = M^{-1}(\lambda J + L). \tag{2}$$

Consequently,

$$M^{\mathrm{T}} L^{-1} = M^{-1}(\lambda J L^{-1} + I), \tag{3}$$

where I is the identity matrix. Multiplying by J from the right we obtain

$$M^{\mathrm{T}} L^{-1} J = M^{-1} J(\lambda L^{-1} J + I). \tag{4}$$

The left hand side is positive by Observation 1 since $M^{\mathrm{T}} L^{-1}$ is nonnegative and non-singular. Setting $B = M^{-1}$ and $C = \lambda L^{-1} J + I$, we notice that C and therefore JC are positive, and BJC is positive because the left hand side of (4) is. Using Observation 3 we conclude that $BJ > 0$, i.e.,

$$M^{-1} J > 0. \tag{5}$$

Multiplying both sides of (3) by M from the right, we obtain

$$M^{\mathrm{T}} L^{-1} M = \lambda M^{-1} J L^{-1} M + I. \tag{6}$$

Here, $\lambda M^{-1} J > 0$ by (5) and $L^{-1} M$ is nonnegative and nonsingular. Consequently (by Observation 1) the right hand side is positive. We conclude that

$$M^{\mathrm{T}} M = M^{\mathrm{T}} I M \geqslant M^{\mathrm{T}} L^{-1} M > 0,$$

completing the proof of Theorem 2. □

References

[1] R.C. Bose, A note on Fisher's inequality for balanced incomplete block designs, Ann. Math. Stat. 20 (1949) 619–620.
[2] W.G. Bridges, A characterization of type-1 λ-designs, J. Combin. Theory Ser. A 22 (1977) 361–367.
[3] N.G. de Bruijn and P. Erdös, On a combinatorial problem, Proc. Konink. Nederl. Akad. Wetensch. Ser. A 51 (1948), 1277–1279 (=Indagationes Math. 10 (1948) 421–423.)
[4] R.A. Fisher, An examination of the different possible solutions of a problem in incomplete blocks, Ann. Eugenics (London) 10 (1940), 52–75.
[5] M. Hall Jr., Combinatorial Theory (Blaisdell, New York, 1967).
[6] J.R. Isbell, An inequality for incidence matrices, Proc. Amer. Math. Soc. 10 (1959) 216–218.
[7] L. Lovász, Combinatorial Problems and Exercises (Akadémiai Kiadó, Budapest and North-Holland, Amsterdam, 1979).
[8] K.N. Majumdar, On some theorems in combinatorics relating to incomplete block designs, Ann. Math. Stat. 24 (1953) 377–389.
[9] H.J. Ryser, A note on a combinatorial problem, Proc. A.M.S. 1 (1950) 422–424.
[10] H.J. Ryser, Combinatorial Mathematics, Carus Math. Monograph 14, Math. Assoc. of America (1963).
[11] H.J. Ryser, An extension of a theorem of de Bruijn and Erdös on combinatorial designs, J. Algebra 10 (1968) 246–261.
[12] S.S. Shrikhande and N.M. Singhi, On the λ-design conjecture, Utilitas Math. 9 (1976) 301–313.
[13] P. de Witte, Some new properties of semi-tactical λ-spaces, Bull. Soc. Math. Belg. 19 (1967) 13–24.
[14] D.R. Woodall, Squares λ-linked designs, Proc. London Math. Soc. 20 (1970) 669–687.
[15] P. Komjáth, An infinite version of Ryser's inequality, to appear.
[16] A. Seress, On the λ-design conjecture, in preparation.