PROCEEDINGS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

On the Number of Local Maxima of a Constrained Quadratic Form

M. Broom, C. Cannings and G. T. Vickers

Proc. R. Soc. Lond. A 1993 **443**, 573-584
doi: 10.1098/rspa.1993.0163

On the number of local maxima of a constrained quadratic form

By M. Broom[1,2], C. Cannings[1] and G. T. Vickers[2]

[1]*Department of Probability & Statistics and* [2]*Department of Applied & Computational Mathematics, The University, Sheffield S10 2TN, U.K.*

We consider the problem of determining the greatest number of local maxima that a quadratic form can have when the vector is constrained to lie within the unit simplex. Specifically, we investigate the local maxima of

$$V = \boldsymbol{p}^{\mathrm{T}} A \boldsymbol{p},$$

where $\boldsymbol{p} = (p_1, p_2, \ldots, p_n) \in \Delta_n = \{\boldsymbol{x} \in \mathbb{R}^n : x_i \geqslant 0, \Sigma_i x_i = 1\}$ and $A = (a_{ij})$ is a real, symmetric $n \times n$ matrix. Considering the central role played by quadratic forms in the history of mathematics in general and algebra in particular, it is perhaps surprising that this problem does not appear to have received any attention. It is a rather awkward problem because the constraint cannot be readily incorporated. A complete solution to the problem is lacking, but we show that the greatest number of maxima that any $n \times n$ matrix can have increases geometrically with n and also present some results on the lengths (i.e. the number of non-zero elements) of the maximizing vectors.

1. Introduction

Apart from its intrinsic interest, the problem being considered arises naturally in the context of population genetics. If there are n alleles $A_1, A_2, \ldots, A_n$ at a particular locus and a_{ij} is the viability of the genotype $A_i A_j$ then the classic recurrence equation for the vector of allelic frequencies $\boldsymbol{p}$ is

$$p_i' = p_i (A\boldsymbol{p})_i / \boldsymbol{p}^{\mathrm{T}} A \boldsymbol{p} \quad (1 \leqslant i \leqslant n), \tag{1}$$

where a prime denotes the values in the next generation and it is a well-known result that the mean fitness V increases monotonically from one generation to the next (see Kingman (1961*b*) for an attractive proof) and is only constant at an equilibrium point. Thus, provided that all of the p_i are initially positive (i.e. $\boldsymbol{p} \notin \partial\Delta_n$, the boundary of Δ_n, at the first generation) $\boldsymbol{p}$ will always converge to a point at which V has a local maximum and every such vector is a locally stable equilibrium point of equation (1). The determination of these points is a straightforward, if tedious, process for any given A. The determination of the greatest number of stable equilibrium points that a system of n alleles can have is an interesting mathematical problem but also has biological implications. The existence of different allele combinations in different populations of a species does not automatically imply that the populations are subject to different factors. However, one would only expect different maxima to coexist if the populations are isolated. If they are connected spatially for a significant time then travelling waves are likely to develop which will replace one of the maxima by the other. Such a wave does not necessarily replace the smaller fitness by the

Proc. R. Soc. Lond. A (1993) **443**, 573–584
Printed in Great Britain

larger; the outcome may also depend upon the dispersal rates (assuming these to be dependent upon the genotypes involved). Such a situation has been considered by Hutson & Vickers (1992).

The notion of a local maximum of a quadratic form (or symmetric matrix) can be extended to non-symmetric matrices by means of the following:

Definition. If $A = (a_{ij})$ is a real $n \times n$ matrix then $\boldsymbol{p} \in \Delta_n$ is an evolutionarily stable strategy (ESS) of A if (i) $\boldsymbol{p}^{\mathrm{T}}A\boldsymbol{p} \geqslant \boldsymbol{q}^{\mathrm{T}}A\boldsymbol{p}\ \forall\, \boldsymbol{q} \in \Delta_n$ and (ii) if $\boldsymbol{q} \in \Delta_n$, $\boldsymbol{q} \neq \boldsymbol{p}$, $\boldsymbol{p}^{\mathrm{T}}A\boldsymbol{p} = \boldsymbol{q}^{\mathrm{T}}A\boldsymbol{p}$ then $\boldsymbol{p}^{\mathrm{T}}A\boldsymbol{q} > \boldsymbol{q}^{\mathrm{T}}A\boldsymbol{q}$. □

A biological motivation of ESSs is to be found in Maynard Smith & Price (1973) and Maynard Smith (1974). All the results presented here are equally valid for ESSs. It was shown in Vickers & Cannings (1988*b*) that there are constraints upon the supports of the ESS vectors and some of their results will be used here. In what follows the term ESS will normally be used for a local maximum on Δ_n and maximum will be used when the statement applies only to symmetric matrices. The support of $\boldsymbol{p} \in \Delta_n$ is denoted by $R(\boldsymbol{p})$, i.e. $R(\boldsymbol{p}) = \{i : p_i > 0\}$.

The expected number of local maxima (and the lengths of them) in random matrices is a problem which is considered in Haigh (1988, 1989) and Kingman (1988, 1989). For example, Kingman (1989) contains the result that the number of stable polymorphisms with just two alleles is asymptotically

$$\tfrac{1}{4}\pi(\tfrac{1}{2}n\pi)^{\frac{1}{2}},$$

when the a_{ij} are independent and chosen from a uniform distribution.

2. Bounds on the number of maxima

Let U_n be the greatest number of ESSs that any $n \times n$ matrix can have. It is shown in Cannings & Vickers (1988) that if $n = 3r+s$ (where $s = 2$, 3 or 4) then

$$U_n \geqslant s3^r.$$

This follows from a consideration of the graph theoretic notion of cliques and provides a lower bound which we improve upon later. This result was discovered by several authors (see Hofbauer & Sigmund 1988), and rests upon a result of Moon & Moser (1965). Furthermore the supports of the vectors which form the local maxima constitute an antichain, because of the non-inclusion result of Bishop & Cannings (1976), and so

$$U_n \leqslant \binom{n}{[\frac{1}{2}n]},$$

which is a classic result of Sperner (1928). Here, and throughout this paper, square brackets will be used to denote the integer part of an expression. It is thus natural to conjecture that U_n increases geometrically at a rate between $3^{\frac{1}{3}}$ and 2. The following theorem is used to establish this result. Define $u_n(r)$ to be the greatest number of ESSs (whose supports have length r) achievable by any $n \times n$ matrix.

Theorem 1. (i) $u_n(r)\,u_m(s) \leqslant u_{n+m}(r+s)$ $(r < n, s < m)$, *and* (ii) $U_n U_m \leqslant U_{m+n}$.

Proof. It is sufficient to prove that if $\boldsymbol{p}$ is an ESS of the $n \times n$ matrix A with

$$(A\boldsymbol{p})_i \begin{cases} = \alpha, & i \in R(\boldsymbol{p}) \\ < \alpha, & i \notin R(\boldsymbol{p}) \end{cases}$$

and if $\boldsymbol{q}$ is an ESS of the $m \times m$ matrix B with

$$(B\boldsymbol{q})_i \begin{cases} = \beta, & i \in R(\boldsymbol{q}) \\ < \beta, & i \notin R(\boldsymbol{q}) \end{cases}$$

then
$$\boldsymbol{r}^{\mathrm{T}} = \left(\frac{(M-\beta)\,\boldsymbol{p}^{\mathrm{T}}}{2M-\alpha-\beta}, \frac{(M-\alpha)\,\boldsymbol{q}^{\mathrm{T}}}{2M-\alpha-\beta} \right)$$

is an ESS of the $(n+m) \times (n+m)$ matrix $C = (c_{ij})$ where

$$c_{ij} = \begin{cases} a_{ij}, & 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant n, \\ b_{i-n\,j-n}, & n+1 \leqslant i \leqslant n+m, n+1 \leqslant j \leqslant n+m, \\ M, & \text{otherwise}, \end{cases}$$

and M is any number which exceeds all the elements of A and B. Necessary and (barring degeneracies) sufficient conditions for a local maximum are given in Kingman (1961a) and for an ESS they are given in Haigh (1975). For the purpose of this paper, it suffices to know that if $\boldsymbol{p} \in \Delta_n$ is such that

$$(A\boldsymbol{p})_i \begin{cases} = \lambda, i \in R(\boldsymbol{p}), \\ < \lambda, i \notin R(\boldsymbol{p}), \end{cases}$$

and $(\boldsymbol{p}-\boldsymbol{q})^{\mathrm{T}} A (\boldsymbol{p}-\boldsymbol{q}) < 0$ whenever $\boldsymbol{q} \in \Delta_n$, $\boldsymbol{q} \neq \boldsymbol{p}$, $R(\boldsymbol{q}) \subset R(\boldsymbol{p})$ then $\boldsymbol{p}$ is an ESS of A. Now

$$(C\boldsymbol{r})_i = \begin{cases} \dfrac{(M-\beta)\,(A\boldsymbol{p})_i + (M-\alpha)\,M}{2M-\alpha-\beta}, & 1 \leqslant i \leqslant n, \\[2ex] \dfrac{(M-\beta)\,M + (M-\alpha)\,(B\boldsymbol{q})_{i-n}}{2M-\alpha-\beta}, & n+1 \leqslant i \leqslant n+m \end{cases}$$

or
$$(C\boldsymbol{r})_i \begin{cases} = \dfrac{M^2-\alpha\beta}{2M-\alpha-\beta}, & i \in R(\boldsymbol{r}), \\[2ex] < \dfrac{M^2-\alpha\beta}{2M-\alpha-\beta}, & i \notin R(\boldsymbol{r}). \end{cases}$$

Let A', B', C' be the submatrices of A, B, C which correspond to the supports of $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}$ and let the lengths of these supports be $s, t, s+t$ respectively. Since $\boldsymbol{p}$ maximizes the quadratic form $\boldsymbol{x}^{\mathrm{T}} A' \boldsymbol{x}$ ($\boldsymbol{x} \in \Delta_n$) we have

$$\boldsymbol{x}^{\mathrm{T}} A' \boldsymbol{x} < \left(\sum_{k=1}^{s} x_k \right)^2 \alpha \,\forall\, \boldsymbol{x} \in \mathbb{R}^s \text{ not parallel to } \boldsymbol{p}$$

and similarly
$$\boldsymbol{y}^{\mathrm{T}} B' \boldsymbol{y} < \left(\sum_{k=1}^{t} y_k \right)^2 \beta \,\forall\, \boldsymbol{y} \in \mathbb{R}^t \text{ not parallel to } \boldsymbol{q}.$$

Let
$$\boldsymbol{z} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} \quad \text{with} \quad \sum_{i=1}^{s+t} z_i = 0.$$

Then
$$\boldsymbol{z}^{\mathrm{T}} C' \boldsymbol{z} = \boldsymbol{x}^{\mathrm{T}} A' \boldsymbol{x} + 2M \left(\sum_{k=1}^{s} x_k \right) \left(\sum_{k=1}^{t} y_k \right) + \boldsymbol{y}^{\mathrm{T}} B' \boldsymbol{y} \leqslant \left(\sum_{k=1}^{s} x_k \right)^2 (\alpha - 2M + \beta) \leqslant 0,$$

with equality only if $\boldsymbol{z} = \boldsymbol{0}$. This completes the proof.

This theorem, together with a result of Pólya & Szego (1972) and the bound of Sperner, gives the following.

Theorem 2. *The sequence* $\{U_n\}$ *is such that* $\lim_{n\to\infty} U_n^{1/n}$ *exists. Furthermore, denoting this limit by* γ, $U_n^{1/n} \leqslant \gamma$ *for all* n.

Note that $U_n^{1/n}$ is not necessarily monotone increasing. It follows from these results that $2 \geqslant \gamma \geqslant 3^{\frac{1}{3}}$. It is natural to conjecture that $\lim_{n\to\infty}(U_{n+1}/U_n)$ also exists. We have not been able to prove this, but clearly if the limit does exist then its value is also γ. The lower bound for γ will be improved upon later.

Obviously

$$u_n(1) = n \quad \text{and} \quad u_n(n) = 1$$

and it is shown in Vickers & Cannings (1988*b*) that

$$u_n(n-1) = 2$$

and in Cannings & Vickers (1988) that

$$u_n(2) = \begin{cases} \frac{1}{4}n^2 & n \text{ even}, \\ \frac{1}{4}(n^2-1), & n \text{ odd}. \end{cases} \tag{2}$$

This result uses a classic theorem of Turan (1954) on the number of edges that can be present in a triangle-free graph. Now suppose that we have an $n \times n$ matrix which has $u_n(r)$ ESSs of length r. Then each of $1, 2, \ldots, n$ is in (on average) $u_n(r)\,r/n$ different ESSs. Remove the index which is in the least number of supports of the ESSs. Then we see that

$$u_{n-1}(r) \geqslant u_n(r) - (r/n)\,u_n(r),$$

which implies

$$u_n(r) \leqslant \left[\frac{nu_{n-1}(r)}{(n-r)}\right].$$

This shows that

$$u_n(n-2) \leqslant n$$

and later we will see that this upper bound is attainable. Consider now the set of numbers $v_n(r)$, with $n \geqslant r+1$, defined by

$$v_n(r) = [nv_{n-1}(r)/(n-r)], \quad v_{r+1}(r) = 2, \tag{3}$$

so that $v_n(r) \geqslant u_n(r)$ and one can show that $v_n(2) = u_n(2)$. It follows from this last equation that

$$nv_{n-1}(r)/(n-r) - 1 \leqslant v_n(r) \leqslant nv_{n-1}(r)/(n-r)$$

from which it can be shown that

$$\frac{2(r-2)}{(r^2-1)}\binom{n}{r} + \frac{n-r+1}{r-1} \leqslant v_n(r) \leqslant \frac{2}{r+1}\binom{n}{r} \tag{4}$$

and so $v_n(r)$ is of order n^r when $n \gg r$. The recurrence relation (3) is a very interesting one and a discussion of its properties will be published elsewhere. One special case is the result

$$v_n(3) = \left[\frac{n^3 - 3n^2 + 6n - 13}{13}\right]$$

and so $\lim_{n\to\infty}(v_n(3)/n^3) = \frac{1}{13}$. Quite generally, $\lim_{n\to\infty}(v_n(r)/n^r)$ exists and is less than $2/(r+1)!$ because $(n-r)!\,v_n(r)/n!$ decreases as n increases. When n and r are large the inequality (4) shows that $v_n(r)$ behaves like

$$\frac{2}{r}\binom{n}{r}.$$

Thus for large n, $v_n(r)$ is a maximum when r is approximately $\frac{1}{2}n$ and

$$\lim_{n\to\infty}\max_r v_n(r)^{1/n} = 2.$$

This suggests that the value of γ might be 2. We have no contradictory information. A more accurate approximation to $u_n(r)$ is provided by $w_n(r)$, where $w_n(r)$ is the maximum number of subsets of size r that can be chosen from n objects when no collection of $(r+1)$ objects contains more than 2 of the subsets. Then

$$v_n(r) \geqslant w_n(r) \geqslant u_n(r).$$

The essence of the latter part of this relationship is Theorem 4 of Vickers & Cannings (1988*b*) which asserts that if $\{1,2,3\}\subset X$ then there cannot be three ESSs with supports $X\backslash\{1\}, X\backslash\{2\}, X\backslash\{3\}$. Clearly $w_n(2)$ is just the maximum number of edges in a triangle-free graph and is the same as $u_n(2)$ given by (2) above. When the complementary sets are considered, we see that $w_n(r)$ is the maximum number of subsets of size $(n-r)$ that can be formed from a set of n objects when no collection of $(n-r-1)$ objects may belong to more than two of the subsets. This formulation of the problem suggests a fairly close link between the evaluation of $w_n(r)$ and block designs. Following Anderson (1989), a block design (b,n,s,k,λ) is a family of b subsets of a set of n elements such that each subset has k elements and each pair of elements belongs to exactly λ subsets. Also

$$s(k-1) = \lambda(n-1) \quad \text{and} \quad bk = ns.$$

The correspondence between the problems is only exact when $(n-r-1) = 2$, $\lambda = 2$ and $k = n-r = 3$. So if the block design $(b,n,s,3,2)$ exists then b is $\frac{1}{3}n(n-1)$ and this is also $w_n(n-3)$. Note that

$$v_n(n-3) = [\tfrac{1}{3}n(n-1)].$$

It follows that $w_7(4)$ is 14 because the subsets

124	235	346	457	561	672	713
126	237	341	452	563	674	715

constitute $(14,7,6,3,2)$. Another classical problem which is closely allied to both $w_n(r)$ and block designs is that of determining the existence of Steiner systems. Again following Anderson (1989), a Steiner system $S(l,m,n)$ is a collection of m-element subsets of an n-element set such that every l-element subset lies in exactly one of the m-element subsets. If the Steiner system $S(l,m,n)$ exists then it contains

$$\binom{n}{l}\bigg/\binom{m}{l}$$

subsets. For the problem $w_n(r)$, m is $(n-r)$ and l is $(n-r-1)$. Hence

$$w_n(r) \leqslant \frac{2}{r+1}\binom{n}{r} \tag{5}$$

and there is equality only if an optimal solution exists, that is one in which every collection of $(n-r-1)$ objects belongs to exactly two of the subsets of size $(n-r)$. The factor of 2 is to allow for the different definition of Steiner systems from that of $w_n(r)$. For example, $S(3,4,8)$ exists and a realization of it is

$$\begin{array}{ccccccc} 1234 & 1256 & 1358 & 1278 & 1367 & 1457 & 1468 \\ 5678 & 3478 & 2467 & 3456 & 2458 & 2368 & 2357. \end{array}$$

The cyclic permutation $1 \to 2 \to \ldots \to 8 \to 1$ gives another realization in which no subset is the same as the first. Thus the union of the two shows that $w_8(4)$ is 28.

The Steiner system $S(2,3,n)$, usually called a Steiner triple system, exists if and only if $n \equiv 1$ or 3 (mod 6). Hence

$$w_n(n-3) = \tfrac{1}{3}n(n-1) \tag{6}$$

for such n provided that there is a permutation which does not change the label of one subset (of length 3, i.e. a triple) so as to be the same as an original label. From the construction of $S(2,3,n)$ given by Anderson (1989) it is easy to see that such a permutation exists. The result (6) can be extended to give the following.

Theorem 3.

$$w_n(n-3) = [\tfrac{1}{3}n(n-1)].$$

Proof. When $n \equiv 3$ (mod 6) a Steiner triple system exists and furthermore it contains a subsystem of $\frac{1}{3}n$ disjoint triples (see Anderson 1989). It is easy to re-order the elements of the base set and so construct another Steiner system in which all of the triples of the new system are different from the old ones and so it also has a new subsystem of disjoint triples.

Let $n \equiv 4$ (mod 6). Construct the two Steiner systems, as above, on $(n-1)$ elements. Remove the $\frac{1}{3}(n-1)$ triples of one of the subsystems and form the $(n-1)$ triples using the nth element and each pair contained in the triple of the subsystem. This produces

$$\tfrac{1}{3}(n-1)(n-2) - \tfrac{1}{3}(n-1) + (n-1) = \tfrac{1}{3}(n-1)\,n$$

triples and each pair of triples belongs to exactly two of them. Thus the result is true for such n.

Let $n \equiv 5$ (mod 6). The construction above on $(n-1)$ elements will produce a set of $\frac{1}{3}(n-1)(n-2)$ triples which contain a subsystem of $\frac{1}{3}(n-2)$ disjoint triples. (The construction of 2 Steiner systems on $(n-2)$ elements had two such subsystems; one has been removed.) Repeat the procedure of the last paragraph to obtain

$$\begin{aligned} \tfrac{1}{3}(n-1)(n-2) - \tfrac{1}{3}(n-2) + (n-2) &= \tfrac{1}{3}(n-2)(n+1) \\ &= \tfrac{1}{3}n(n-1) - \tfrac{2}{3} \\ &= [\tfrac{1}{3}n(n-1)] \end{aligned}$$

triples of the required type. The pair of elements $(n-1, n)$ will not belong to any triple.

When $n \equiv 1$ (mod 6) a Steiner triple system exists but it only contains $\frac{1}{6}n$ disjoint triples, using the construction of Anderson (1989). It is thus necessary to re-number the elements to produce a second Steiner system in which not only is every triple different from the old but the union of the new disjoint triple is the complement of the old. This is in fact easy to do and the proof for $n \equiv 2$ (mod 6) now follows the same lines as that for $n \equiv 5$ (mod 6).

The case $n \equiv 0 \pmod 6$ is the most awkward. We first show the subsidiary result that if

$$w_n(n-3) = [\tfrac{1}{3}n(n-1)]$$

when $n = k$ then it also holds for $n = 3k$. Suppose first that $k \equiv 0$ or $1 \pmod 3$. Divide the $3k$ elements into 3 blocks of k:

$$\begin{array}{cccc} 1 & 2 & \ldots & k \\ k+1 & k+2 & \ldots & 2k \\ 2k+1 & 2k+2 & \ldots & 3k. \end{array}$$

Within each block we may select $\frac{1}{3}k(k-1)$ triples where each pair belongs to exactly two of the triples. Now choose triples with one element from each block as follows;

$$\left\{\begin{array}{l} (i \quad k+1 \quad 2k+i) \quad (i \quad k+2 \quad 2k+i+1) \ldots (i \quad 2k \quad 2k-i-1), \\ (i \quad k+1 \quad 2k+i-2) \quad (i \quad k+2 \quad 2k+i-1) \ldots (i \quad 2k \quad 2k-i-3), \end{array}\right. \quad 1 \leqslant i \leqslant k.$$

This gives a total of

$$k(k-1)+2k^2 = k(3k-1) = \tfrac{1}{3}(3k)(3k-1)$$

triples as required, with every pair of elements in exactly two triples.

Suppose now that $k \equiv 2 \pmod 3$. Since $k \equiv 2$ or $5 \pmod 6$ we can choose $\frac{1}{3}(k-2)(k+1)$ triples in each block of k element and all pairs bar one will belong to exactly two of the triples. The exceptional pair belongs to none of the triples. Choose the three exceptional pairs (one from each block) to be

$$(1 \quad 2) \quad (k+1 \quad k+2) \quad (2k+1 \quad 2k+2).$$

Now choose triples, one element from each block, exactly as above except that the following four triples are not selected;

$$\begin{array}{ll} (1 \quad k+1 \quad 2k+1) & (1 \quad k+2 \quad 2k+2) \\ (2 \quad k+1 \quad 2k+2) & (2 \quad k+2 \quad 2k+1). \end{array}$$

Now form the triples

$$\begin{array}{lll} (1 \quad 2 \quad k+1) & (k+1 \quad k+2 \quad 2k+1) & (2k+1 \quad 2k+2 \quad 1) \\ (1 \quad 2 \quad k+2) & (k+1 \quad k+2 \quad 2k+2) & (2k+1 \quad 2k+2 \quad 2) \end{array}$$

to give as the total number of triples

$$(k-2)(k+1)+(2k^2-4)+6 = k(3k-1).$$

A little checking will confirm that each pair does indeed belong to exactly 2 of the triples.

We now know that the theorem is true if $n \not\equiv 0 \pmod 6$ and that if it is true for k then it is also true for $3k$. This only leaves the case $n = 2.3^a$ and since the theorem is true when $n = 6$ it is true for all n. □

Theorem 3 is not new and proofs of the results can be found in Street & Street (1987) and Stevenson & Wallis (1983). However, the proof given here is much more direct, being based upon an explicit construction rather than Latin squares, and is consequently more appropriate to our problem.

It is also known that $S(3,4,n)$ exists if and only if $n \equiv 2$ or $4 \pmod 6$. Hence, for such n,

$$w_n(n-4) = \tfrac{1}{12}n(n-1)(n-2), \tag{7}$$

because a suitable permutation will again exist. Unfortunately equation (7) cannot be extended to apply for all n simply by inserting square brackets. This is because we have been able to show that $w_7(3)$ is only 15 whereas the formula (7) gives $17\frac{1}{2}$.

According to Anderson (1989) only 14 Steiner systems are known with l greater than 3. We are only concerned with those for which m is $l+1$ and this reduces the number to 5. Specifically, we can say that the values of

$$w_{12}(6), \quad w_{24}(18), \quad w_{48}(42), \quad w_{72}(66) \quad \text{and} \quad w_{84}(78)$$

are all given by the formula

$$w_n(r) = \frac{2}{r+1}\binom{n}{r}$$

because again a suitable permutation will exist.

The numbers $w_n(r)$ satisfy the same inequality as that demonstrated for the $u_n(r)$ in theorem 1. This is because if X is a typical member of the subsets of length r of a set of size n whose elements are labelled $1, 2, \ldots, n$ and if Y is likewise a typical member of the subsets of size s of a set of m elements labelled $n+1, n+2, \ldots, n+m$ then the collection of all sets like $X \cup Y$ will form subsets of length $r+s$ of $n+m$ elements. For example, $w_8(4)$ being 28 implies that $w_{23}(12)$ is at least 21952. Perhaps of more significance is that the above inequality implies that the limit

$$\lim_{n\to\infty} \max_r w_n(r)^{1/n}$$

exists and does not exceed 2. This result is improved upon in the next theorem.

Theorem 4.

$$\lim_{n\to\infty} \max_r w_n(r)^{1/n} = \lim_{n\to\infty} w_{2n}(n)^{1/2n} = 2.$$

Proof. Let $t_n(r)$ be the maximum number of subsets of size r that can be formed from a base set of n elements when each pair of subsets differs by at least 2 elements. In such a family of subsets, no collection of $(r+1)$ elements can contain more than one subset. Hence $t_n(r)$ certainly cannot exceed $w_n(r)$. It is first shown that

$$t_{nq}(np) \geqslant \binom{q}{p}^{n-1}. \tag{8}$$

Divide a base set of $2q$ elements into two equal blocks. There are ${}^qC_p$ possible subsets of size p that can be formed from each block. Arbitrarily associate the subsets from one block with those formed from the other. This will give ${}^qC_p$ subsets of length $2p$ from a base set of $2q$ elements and clearly any two of the subsets will differ by at least two elements. For example, with $p=2$, $q=3$ we can choose

1245
2356
1346.

Thus $t_{2q}(2p)$ is at least ${}^qC_p$. Now form the ${}^qC_p$ different solutions by simply permuting the subsets from the second block. For the example above this would give the following:

1245	1246		1256
2356,	2345	and	2346.
1346	1356		1345

Let a third block of q elements now join the base set and form the ${}^qC_p$ subsets of length p. Assign each of these to one of the solutions above, e.g.

$$\begin{matrix} 124578 \\ 235678, \\ 134678 \end{matrix} \quad \begin{matrix} 124689 \\ 234589 \\ 135689 \end{matrix} \quad \text{and} \quad \begin{matrix} 125679 \\ 234679. \\ 134579 \end{matrix}$$

This will give $({}^qC_p)^2$ subsets of length $3p$ from a base set of $3q$ elements which preserve the required amount of difference. Thus

$$t_{3q}(3p) \geqslant \binom{q}{p}^2.$$

We may now form different solutions by permuting the subsets from the third block. This will give ${}^qC_p$ different solutions to the three-block problem. Fourth, and subsequent blocks, may be added in exactly the same way. This establishes the inequality (8). When combined with a weak form of inequality (4) we obtain

$$\binom{nq}{np} \geqslant v_{nq}(np) \geqslant w_{nq}(np) \geqslant t_{nq}(np) \geqslant \binom{q}{p}^{n-1}$$

and so
$$\binom{nql}{npl}^{1/nl} \geqslant (w_{nql}(npl))^{1/nl} \geqslant \binom{ql}{pl}^{1/l-1/nl}.$$

Letting $n = l \to \infty$ gives

$$\lim_{n\to\infty} (w_{nq}(np))^{1/n} = q^q/(p^p(q-p)^{q-p}), \tag{9}$$

where n is restricted to being a square. However, we know that

$$w_{n+m}(r+s) \geqslant w_n(r)\, w_m(s)$$

and if we define
$$a_n = w_{nq}(np)$$

for fixed p and q it follows that $a_{n+m} \geqslant a_n\, a_m$ and so, using the result of Pólya & Szego (1972) again, we know that the limit

$$\lim_{n\to\infty} (a_n)^{1/n} = \lim_{n\to\infty} (w_{nq}(np))^{1/n}$$

exists. This shows that the limit (9) exists without any restriction on the values of n. In particular,

$$\lim_{n\to\infty} w_{2n}(n)^{1/2n} = 2.$$

Since we already know that

$$\lim_{n\to\infty} \max_r w_n(r)^{1/n}$$

exists and does not exceed 2, its value must also be 2.

3. Symmetric circulant matrices

The analysis for arbitrary (symmetric) matrices is in general rather difficult. It is thus natural to seek a simpler class of matrices for which it is possible to find all the local maxima. It was these considerations which prompted the study of matrices with $a_{ii} = 0$ and $a_{ij} = \pm 1$ $(i \neq j)$ in Cannings & Vickers (1988). However, it was shown in Vickers & Cannings (1988*a*) that such matrices will not in general provide the greatest number of maxima. This was shown by demonstrating that the 7×7 symmetric circulant matrix with first row given by

$$[0 \quad 8 \quad 13 \quad 2 \quad 2 \quad 13 \quad 8]$$

has 14 local maxima (or ESSS) rather than the 12 which is all that clique matrices can provide. The general problem of finding all maxima for such matrices would again seem to be rather complex but we present some results for special cases.

Theorem 5.

$$u_n(n-2) = n.$$

The proof of this theorem is given as an appendix and consists in showing that the $n \times n$ symmetric circulant matrix with first row

$$[-2\cos\theta \quad 1 \quad 0 \ldots 0 \quad 1]$$

has n local maxima when

$$2\pi/n < \theta < 2\pi/(n-1).$$

Since $v_n(n-2) = n$ it follows that $w_n(n-2)$ is also n.

Theorem 6.

$$\gamma \geqslant 30^{1/9} (> 14^{1/7} > e^{1/e} > 3^{1/3}).$$

This depends upon the principal 9×9 submatrix of the 11×11 symmetric circulant matrix with first row

$$[0 \quad 15 \quad 0.01 \quad 14 \quad 1.1 \quad 7.6 \quad 7.6 \quad 1.1 \quad 14 \quad 0.01 \quad 15]$$

having 30 local maxima each of length 3. The supports of these maxima are

$$1\ 2\ 3 \quad 1\ 2\ 5 \quad 1\ 2\ 7 \quad 1\ 2\ 9 \quad 1\ 4\ 7,$$

and their cyclic companions (on 1 to 11), but omitting any with 10 or 11. The result then follows from Theorem 2.

We are grateful to I. Anderson for valuable comments made on an earlier version of this paper.

Appendix

Lemma 1. (This is a slightly stronger form of a result of Kingman (1961*a*).) *Suppose that $A = (a_{ij})$ is a symmetric, non-singular $n \times n$ matrix and let the equations*

$$(A\boldsymbol{u})_i = w \quad (1 \leqslant i \leqslant n), \qquad \sum_{i=1}^{n} u_i = 1$$

have a solution with $w > 0$. Then the following statements are equivalent: (i) $\boldsymbol{y}^{\mathrm{T}} A \boldsymbol{y} < 0$ *whenever $\sum_i \boldsymbol{y}_i = 0$ and $\boldsymbol{y} \neq \boldsymbol{0}$*, (ii) *$A$ has precisely one positive eigenvalue.*

Proof. Let $X = (x_{ij})$ be the matrix of eigenvectors and λ_i be the eigenvalues (if the eigenvalues are not all distinct, the eigenvectors can still be chosen to be linearly independent) so that

$$(AX)_{ij} = \lambda_j x_{ij} \quad (1 \leqslant i \leqslant n, \quad 1 \leqslant j \leqslant n), \qquad X^{\mathrm{T}} X = I,$$

where I is the $n \times n$ identity matrix. Since the eigenvectors span $\mathbb{R}^n$ we can write

$$y_i = \sum_j \alpha_j x_{ij}, \quad u_i = \sum_j \beta_j x_{ij}$$

and so

$$(A\boldsymbol{y})_i = \sum_j \alpha_j \lambda_j x_{ij} \quad \text{and} \quad \boldsymbol{y}^{\mathrm{T}} A \boldsymbol{y} = \sum_j \lambda_j \alpha_j^2.$$

We may suppose that the columns of X are ordered so that $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$. Suppose first that $\boldsymbol{y}$ is a linear combination of the first two columns of X. Then since $\boldsymbol{y}^{\mathrm{T}} A \boldsymbol{y}$ is negative it follows that λ_2 is negative. Furthermore, $\boldsymbol{u}^{\mathrm{T}} A \boldsymbol{u}$ is positive and so λ_1 is positive. Hence statement (i) implies (ii). Now suppose that A has just one positive eigenvalue and that $\sum_i \boldsymbol{y}_i = 0$. Then

$$(\boldsymbol{u}^{\mathrm{T}} A X)_i = (X^{\mathrm{T}} A \boldsymbol{u})_i$$

implies that

$$\lambda_i \beta_i = w \sum_k x_{ki} \quad \text{and} \quad \sum_i \alpha_i \beta_i \lambda_i = 0.$$

Thus, for any θ,

$$\sum_j \lambda_j (\alpha_j - \theta \beta_j)^2 = \sum_j \lambda_j \alpha_j^2 + \theta^2 \sum_j \lambda_j \beta_j^2$$

and also

$$\sum_{ij} \beta_j x_{ij} = 1 \quad \text{implies that} \quad \sum_j \lambda_j \beta_j^2 = w.$$

If we now choose θ to be α_1/β_1 then we see that $\boldsymbol{y}^{\mathrm{T}} A \boldsymbol{y}$ is negative as required.

Lemma 2. *The eigenvalues of the $k \times k$ tridiagonal matrix B with first row*

$$[-2\cos\theta \quad 1 \quad 0 \quad 0 \ldots 0]$$

are $-2\cos\theta - 2\cos(r\pi/(k+1))$ *for* $r = 1, 2, \ldots, k$.

Lemma 3. *The system of k equations* $(Bx)_i = \text{const.}$ *has a solution*

$$x_i = \cos(\tfrac{1}{2}(k+1) - i)\,\theta - \cos(\tfrac{1}{2}(k+1)\,\theta) \quad (1 \leqslant i \leqslant k).$$

Also the value of the constant is $\cos\tfrac{1}{2}(k+1)\,\theta(1 - 1/\cos\theta)$ *and if* $2\pi/k > \theta > 0$ *then all the* x_i *are positive.*

Proof. It is a straightforward matter to check that the exhibited x_i do satisfy the equations. Furthermore, if x_1 is positive then all of the x_i will be positive and

$$x_1 = 2\sin(\tfrac{1}{2}k\theta)\sin(\tfrac{1}{2}\theta).$$

Proof of Theorem 5. The $n \times n$ symmetric circulant with first row

$$[-2\cos\theta \quad 1 \quad 0 \quad 0 \ldots 0 \quad 1]$$

will have n local maxima, each with a support of length $n-2$, provided that: (a) the $(n-2) \times (n-2)$ principal submatrix, C, has just one positive eigenvalue; (b) the solution to $(Cx)_i = w$, $\sum_i x_i = 1$, has $x_i > 0\,\forall\, i$ and w is positive; and (c) the solution satisfies the condition of being proof against invasion, i.e.

$$-2x_1\cos\theta + x_2 \geqslant x_{n-2} \quad \text{and} \quad -2x_1\cos\theta + x_2 \geqslant x_1.$$

The condition (a) reduces to $2\pi/(n-1) > \theta > \pi/(n-1)$ and this range of θ also satisfies the condition (b). Moreover the condition (c) requires that $4\pi/n > \theta > 2\pi/n$ and so the final conclusion is that the exhibited circulant matrix will have n local maxima, each with support of length $n-2$, provided that $2\pi/(n-1) > \theta > 2\pi/n$.

References

Anderson, I. 1989 *A first course in combinatorial mathematics*, 2nd edn (*Oxford appl. Math. and Comp. Ser.*). Oxford University Press.

Bishop, T. & Cannings, C. 1976 Models of animal conflict. *Adv. appl. Prob.* **8**, 616–621.

Cannings, C. & Vickers, G. T. 1988 Patterns of ESS's II. *J. theor. Biol.* **132**, 409–420.

Haigh, J. 1975 Game theory and evolution. *Adv. appl. Prob.* **7**, 8–11.

Haigh, J. 1988 The distribution of evolutionarily stable strategies. *J. appl. Prob.* **25**, 233–246.

Haigh, J. 1989 How large is the support of an ESS? *J. appl. Prob.* **26**, 164–170.

Hofbauer, J. & Sigmund, K. 1988 *The theory of evolution and dynamical systems.* Cambridge University Press.

Hutson, V. C. L. & Vickers, G. T. 1992 Travelling waves and dominance of ESS's. *J. math. Biol.* **30**, 457–471.

Kingman, J. F. C. 1961*a* A mathematical problem in population genetics. *Proc. Camb. phil. Soc.* **57**, 574–582.

Kingman, J. F. C. 1961*b* A matrix inequality. *Quart. J. Math.* **12**, 78–80.

Kingman, J. F. C. 1988 Typical polymorphisms maintained by selection at a single locus. *J. appl. Prob.* A **25**, 113–125.

Kingman, J. F. C. 1989 Maxima of random quadratic forms on a simplex. In *Probability, statistics and mathematics* (ed. T. W. Anderson *et al.*). Academic Press.

Maynard Smith, J. 1974 The theory of games and the evolution of animal conflict. *J. theor. Biol.* **47**, 209–221.

Maynard Smith, J. & Price, G. R. 1973 The logic of animal conflict. *Nature, Lond.* **246**, 15–18.

Moon, J. W. & Moser, L. 1965 On cliques in graphs. *Israel J. Maths.* **3**, 23–38.

Pólya, G. & Szego, G. 1972 *Problems and theorems in analysis.* vol. 1, problem 98. Springer-Verlag.

Sperner, E. 1928 Ein Satz über Untermenge einer endlichen Menge. *Math. Z.* **27**, 544–548.

Stevenson, D. R. & Wallis, W. D. 1983 Two-fold triple systems without repeated blocks. *Discrete Math.* **47**, 125–8.

Street, A. P. & Street, D. J. 1987 *Combinatorics of experimental design.* Oxford University Press.

Turan, P. 1954 On the theory of graphs. *Colloq. Math.* **3**, 19–30.

Vickers, G. T. & Cannings, C. 1988*a* On the number of stable equilibria in a one locus, multiallelic system. *J. theor. Biol.* **131**, 273–277.

Vickers, G. T. & Cannings, C. 1988*b* Patterns of ESS's I. *J. theor. Biol.* **132**, 387–408.

Received 18 January 1993; accepted 16 April 1993