Mathematical Programming 13 (1977) 81–87.
North-Holland Publishing Company

ERGODICITY AND SYMMETRIC MATHEMATICAL PROGRAMS

Arie TAMIR
Tel-Aviv University, Tel-Aviv, Israel

Received 17 February 1976
Revised manuscript received 28 February 1977

The paper provides new conditions ensuring the optimality of a symmetric feasible point of certain mathematical programs. It is shown that these conditions generalize and unify most of the known results dealing with optimality of symmetric policies (e.g. [2, 4, 6, 11]). The generalization is based on certain ergodic properties of nonnegative matrices. An application to a socio-economic model dealing with optimization of a welfare function is presented.

Key words: Symmetric mathematical programs, Ergodicity, S-concavity and majorization, Stochastic matrices, Cyclic symmetry.

1. Introduction

In recent operations research and economic literature we find several decision and optimization models that possess certain symmetry properties. For example, Samuelson [9] and Hadar and Russell [5] have proved that a risk averter, (characterized by a concave utility function) would allocate his funds equally among prospects having a symmetric joint distribution. Bessler and Veinott [2] have studied symmetric networks. Kielson [6, 7] has discussed decision problems associated with random walk models. Berge [1] has applied results on symmetry to quasi-convex programming and Greenberg and Pierskala [4] have shown that certain symmetric nonlinear integer programs can be reduced to corresponding extremal problems in just one variable.

The basic property shared by the above models (and others as well) is a set of convexity and symmetry assumptions that ensure the optimality of a symmetric policy, i.e. one where all the decision variables take on the same value.

The main purpose of this study is to generalize and then unify the known assumptions guaranteeing the optimality of a "symmetric policy". Our generalization is based on certain ergodic properties of nonnegative matrices.

We consider the mathematical program given by

$$\max f(x);\ x \in X,$$

where f is a real-valued function defined on $X \subset \mathbf{R}^n$. Following Berge [1], Greenberg and Pierskala [4], and Tamir [11] we recall several definitions. X is a *(cyclically) symmetric set* if $x \in X$ implies $Px \in X$ for all (cyclic) permutation matrices P. X is *S-convex* if $x \in X$ implies $Sx \in X$ for all doubly stochastic matrices S. Finally X is *cyclically convex* if $x \in X$ implies that $Sx \in X$ for all (doubly stochastic) matrices S which are convex combinations of cyclic permutations.

Turning to the objective function f we say that f is *symmetric (cyclically symmetric)* on a symmetric (cyclically symmetric) set X if $f(Px) = f(x)$ for all permutations P (for all cyclic permutations P). f is *S-concave (cyclically concave)* on an S-convex (cyclically convex) set X if $f(Sx) \geq f(x)$ for all doubly stochastic matrices S (for all matrices S which are convex combinations of cyclic permutations).

We note that S-concavity implies cyclic concavity, since a doubly stochastic matrix can be represented as a convex combination of permutation matrices. It is shown in [1, 4] that S-convex sets contain a symmetric point, i.e. a point $x = (x_i) \in \mathbf{R}^n$, where $x_i = z$ for $i = 1, \ldots, n$. Furthermore, the set of points maximizing an S-concave function on an S-convex set X, is itself S-convex, and hence contains a (maximum) symmetric point.

The above results are extended in [11] to the cyclic concave case. In particular the existence of an optimal symmetric point is demonstrated.

Our purpose is to provide more general conditions on f and X, that will yield the optimality of a symmetric point. A motivation for our results is given in the last section where an application to a socio-economic welfare model is presented.

2. Ergodicity and symmetry

Let A be an $n \times n$ matrix and $X \subset \mathbf{R}^n$. We say that X is A-closed if $x \in X$ implies that the closure of $\{x, Ax, A^2x, \ldots\}$ is in X. f defined on an A-closed set X is A-majorizing if $f(Ax) \geq f(x)$ for all $x \in X$. We observe that the class of nonnegative A-majorizing functions defined on the same A-closed set is closed under multiplication and addition, where these two operations are pointwise. As an example we mention the quasiconcave symmetric functions, that are A-majorizing for any doubly stochastic matrix A, i.e. they are S-concave, (see [4]). Thus, unlike the quasiconcavity property, which is not preserved under addition of symmetric quasiconcave functions the S-majorization is maintained.

To present our main result we recall several definitions from the theory on nonnegative matrices.

Suppose A is a square nonnegative matrix. If by permuting the rows and columns of A in the same way, we obtain a matrix of the form

$$\begin{bmatrix} B & 0 \\ C & D \end{bmatrix}$$

where B is a square matrix, then A is said to be *reducible*. Otherwise it is called *irreducible*. A is called *aperiodic* and *primitive* if A^k is positive for some integer k. Finally, A is substochastic if its row sums do not exceed 1, and it is *strictly substochastic* if for at least one row the sum is less than 1.

Theorem 1. *Let A be an $n \times n$ substochastic matrix such that for some integer k, the elements of A^k are positive. Suppose that $X \subset \mathbf{R}^n$ is A-closed and f, defined on X, is an A-majorizing real function. If $\forall x \in X$, and $\forall$ converging subsequence $\{u^{(i)}\} \in \{x, Ax, A^2x, \ldots\}$ with $\lim_{i\to\infty} f(u^{(i)}) \le f(\lim_{i\to\infty} u^{(i)})$ holds then for any $x \in X$, there exists a symmetric point $y \in X$ (possibly dependent on x) and $f(y) \ge f(x)$. Further, if A is strictly substochastic, then for any $x \in X$, y is the zero vector.*

Proof. Consider first the stochastic case, i.e. when all row sums are equal 1.

From Perron–Frobenius theory for nonnegative matrices it follows that A has exactly one eigenvalue, $\lambda = 1$, of unit modulus, and all other eigenvalues are strictly less than 1 in modulus. Therefore, see [3, 10], if $i \to \infty$, then $A^i \to e\pi$ elementwise, where e is a column vector of 1's, and π is the (positive) row vector of A corresponding to the eigenvalue $\lambda = 1$. (Furthermore, the rate of approach to the above limit is geometric).

Using the A-closedness of X we obtain that for each $x \in X$, $y = e\pi x \in X$. y is clearly a symmetric point having all its components equal to πx. And since f is A-majorizing and continuous by assumption:

$$f(y) = f\Big(\lim_{i\to\infty} A^i x\Big) \ge \lim_{i\to\infty} f(A^i x) \ge f(x).$$

For the strictly substochastic case, the corresponding Perron–Frobenius eigenvalue is less than 1, (see [3, p. 120]). Therefore if $i \to \infty$, then A^i converges geometrically to the zero matrix. It then follows that $f(0) \ge f(x)$ for all $x \in X$.

Remark 1. In terms of Markov chains theory, the assumptions on the matrix A in the stochastic case are equivalent to the ergodicity of the corresponding finite chain, i.e. an irreducible aperiodic stochastic chain.

Remark 2. It is worth pointing out that the assumptions of Theorem 1 are indeed weaker than S-concavity. While S-concavity implies A-majorization of every doubly stochastic matrix, the Theorem requires only that the A-majorization holds for some doubly stochastic matrix A with A^k positive for some k. As an example we establish (see [11]) the class of cyclically concave functions, i.e. functions which are A-majorizing for any doubly stochastic matrix A which is a convex combination of cyclic permutation matrices. (As pointed out in the Introduction the S-concave functions constitute a proper subset of the above).

Remark 3. The result of Theorem 1 for the strictly substochastic case holds for more general circumstances. Given an arbitrary real square matrix A, a necessary and sufficient condition for the convergence of the sequence A^i to the zero matrix is that the moduli of all the eigenvalues of A are less than 1, (see [8, Theorem 11.1.1]). Applying the latter result together with Perron–Frobenius theory for irreducible matrices [3, p. 120], we observe that for the strictly substochastic case of Theorem 1 we can relax the property that A^k is positive for some k by irreducibility to obtain the convergence of $\{A^i\}$ to the zero matrix.

(Simple two dimensional examples illustrate that positivity of A^k cannot be replaced by irreducibility in the stochastic case, while irreducibility cannot be omitted in the strictly substochastic case).

Remark 4. The optimality of a symmetric point in the S-concave [1, 4] and cyclically concave [11] cases are implied by the above theorem when A is a matrix with all the components equal to $1/n$. (The only distinct eigenvalues of A are 0 and 1, the latter being simple).

Theorem 1 also provides a generalization to the following set of conditions for optimality of a symmetric vector, given by Keilson [6].

Theorem 2 (Keilson). *Let $f(x)$ be a symmetric real function defined on the hyperplane $H = \{x \mid \Sigma x_i = 1\}$. Let f be continuous and suppose that for each $x \in H$*

$$f(x_1, x_2, \ldots, x_n) \leq f\left(\frac{x_1 + x_2}{2}, \frac{x_1 + x_2}{2}, x_3, \ldots, x_n\right). \tag{1}$$

Then $f(x) \leq f(1/n, \ldots, 1/n)$ for all $x \in H$.

This result is extended in [11] to the cyclically symmetric case. We now show that Theorem 2 is implied by Theorem 1 when

$$A = \begin{bmatrix} \frac{1}{2}, & \frac{1}{2}, & 0, \ldots & \ldots, 0 \\ 0, & 0, & 1, \ldots & \ldots, 0 \\ \cdot & & & \cdot \\ \cdot & & & 0 \\ \cdot & & & \cdot \\ 0, & 0, & 0, \ldots & \ldots, 1 \\ \frac{1}{2}, & \frac{1}{2}, & 0, \ldots & \ldots, 0 \end{bmatrix}.$$

To show the positivity of A^k we use the property that for a stochastic irreducible matrix A, A^k is positive for some k if and only if the only eigenvalue with unit modulus is 1, [(3, p. 120–124)].

A is clearly stochastic and irreducible since all states of the corresponding markov chain intercommunicate. To see that 1 is the only eigenvalue with unit modulus, we note that $|A - \lambda I| = (-1)^{n-1}[\frac{1}{2}(\lambda + \lambda^{n-1}) - \lambda^n]$, and hence the result follows. To complete our proof we point out that (cyclic) symmetry together with property (1) of Theorem 2 implies that the conditions of Theorem 1 are satisfied for the above matrix A. Also $(1/n, \ldots, 1/n)$ is the unique symmetric point in H.

We conclude this section by observing that the ergodic requirement given in Theorem 1 can be weakened to the following structure:

$$A = \begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix},$$

where A is stochastic, A_1 and A_2 are irreducible matrices, A_1 is ergodic and $A_{21} \neq 0$. (Matrices of this structure are called regular in [10]). As shown in [3, p.

125] a matrix A, fulfilling these properties satisfies

$$\lim_{k\to\infty} A^k = e\pi, \quad \text{where } \pi = (\pi_1, \ldots, \pi_m, 0, \ldots, 0)$$

and

$$\lim_{k\to\infty} A_1^k = \begin{bmatrix} \pi_1, \ldots, \pi_m \\ \pi_1, \ldots, \pi_m \end{bmatrix} \quad (m < n).$$

From the proof of Theorem 1, it follows that for each $x = (x_1, \ldots, x_n) \in X$ there exist $y = (\sum_{i=1}^{m} \pi_i x_i)e$, (certainly symmetric) and $f(y) \geq f(x)$.

3. An application to a socio-economic model

To further motivate our results we describe the following model. Consider a society, consisting of n individuals, which is interested in distributing its total wealth, W, among its members in order to maximize its welfare function. Denoting by x_i, the wealth given to the ith member we assume that the welfare function, $f = f(x_1, \ldots, x_n)$, is indifferent as to the allocation of the wealth vector $(x_1, \ldots, x_n)$ among the n individuals, i.e. $f(x_1, \ldots, x_n)$ is fully symmetric in its arguments. Suppose further that the structure of the society is such that if any p individuals form a coalition and share their corresponding wealth vector then the entire society gains in terms of welfare. (Note that p is a given integer, $2 \leq p \leq n$).

We are interested in finding conditions on the sharing mechanism of p-coalitions that would ensure that maximum society welfare is achieved for a policy distributing the same wealth to each member of the society. Using the results of the previous section we will show that if each p-coalition has the same linear mechanism for dividing its wealth vector then an equal distribution of the total wealth of the society is optimal, provided certain positivity and continuity properties are met.

Specifically, it is assumed that if $(x_{i_1}, \ldots, x_{i_p})$ is the wealth given to an arbitrary p-coalition, $(i_1, \ldots, i_p)$, then after sharing among themselves the members hold the wealth vector

$$(y_{i_1}, \ldots, y_{i_p})' = B(x_{i_1}, \ldots, x_{i_p})'. \tag{3}$$

B is a positive, doubly stochastic $p \times p$ matrix.

The assumption that the society favors p-coalitions is now formulated as follows.

For any permutation $(i_1, \ldots, i_p, \ldots, i_n)$ and wealth distribution $(x_{i_1}, \ldots, x_{i_p}, \ldots, x_{i_n})$ satisfying $x_{i_1} + \cdots + x_{i_n} \leq W$

$$f(x_{i_1}, \ldots, x_{i_p}, \ldots, x_{i_n}) \leq f(y_{i_1}, \ldots, y_{i_p}, x_{i_{p+1}}, \ldots, x_{i_n}) \tag{4}$$

where $(y_{i_1}, \ldots, y_{i_p})$ is given by (3).

Given an arbitrary distribution wealth vector $(x_1, \ldots, x_n)$ we combine the symmetric property with property (4) to have

$$f(x_1, \ldots, x_p, \ldots, x_n) \le f(y_1, \ldots, y_p, x_{p+1}, \ldots, x_n)$$
$$= f(y_2, \ldots, y_p, x_{p+1}, \ldots, x_n, y_1)$$

where $(y_1, \ldots, y_p)$ are defined by (3). Using $B_{.j}$ to denote the j-th row of B, we define the $n \times n$ matrix $A = (a_{ij})$ by

$$A = \begin{array}{c} \begin{array}{cc} p & n-p \end{array} \\ \left[\begin{array}{c|c} B_{.2} & \\ \vdots & 0 \\ B_{.p} & \\ \hline 0 & I_{n-p} \\ \hline B_{.1} & 0 \end{array}\right] \end{array} \begin{array}{l} p-1 \\ \\ n-p \\ \\ \end{array}$$

I_{n-p} is the identity of order $n-p$. It is clear that A is doubly stochastic and that the set of wealth vectors $(x_1, \ldots, x_n)$ satisfying $x_1 + \cdots + x_n \le W$ is A-closed. Furthermore the welfare function is A-majorizing. The next lemma proves that A given by (5) is primitive.

Lemma 3. *Let $A = (a_{ij})$ be a square matrix defined by* (5) *where $B_{.j}$ $(j = 1, \ldots, p)$ is a positive vector. Then, $k = 2(n-p)+1$ is the smallest power such that A^k is a positive matrix.*

Proof. Since we are concerned only with sign properties of A^k where A is a nonnegative matrix, we may replace each positive entry of A^k by 1. Thus it is easily verified that

$$A^{(n-p)+1} = \begin{array}{c} p \\ \left[\begin{array}{ccccc} 1 & \cdots & 1 & \cdots & 1 \\ & & & & \\ 1 & \cdots & 1 & \cdots & 1 \\ \hline 1 & \cdots & 1\ 0 & \cdots & 0 \\ & & \ddots & & \\ 1 & \cdots & 1 & \cdots & 1\ 0 \\ \hline 1 & \cdots & 1 & \cdots & 1 \end{array}\right] \end{array} \begin{array}{l} p-1 \\ \\ n-p \\ \\ \end{array} .$$

Moreover, for any integer $1 \le k < 2(n-p)+1$ the element (p, n) of A^k is zero. Noting that for $k = 2(n-p)$ the (p, n) element is the only zero entry, we conclude that $A^{2(n-p)+1}$ is a positive matrix; thus completing the proof.

To conclude the discussion we prove our main assertion, regarding the above model.

Theorem 4. *Let $f(x_1, \ldots, x_n)$ be the welfare function of the society when the n individuals are distributed the wealth vector $(x_1, \ldots, x_n)$. Suppose that $f(x_1, \ldots, x_n)$ is continuous and symmetric. If the society prefers p-coalitions in the sense of* (3)–(4) *then the maximum welfare of the entire society is achieved for a distribution where each member obtains the same wealth.*

Proof. Using Lemma 3 we observe that the assumptions of Theorem 1 are met. Thus a symmetric distribution is optimal.

Extensions of the above model have been developed and will be reported elsewhere.

References

[1] C. Berge, *Topological spaces* (Macmillan, New York, 1963).

[2] S.A. Bessler and A.F. Veinott Jr., "Optimal policy for a dynamic multi-echelon inventory model", *Naval Research Logistic Quarterly* 13 (1966) 355–387.

[3] D.R. Cox and H.D. Miller, *The theory of stochastic processes* (Methuen, London, 1970).

[4] H.J. Greenberg and W.P. Pierskalla, "Symmetric mathematical programs", *Management Science* 16 (1970) 309–312.

[5] S. Hadar and W.R. Russell, "Diversification of interdependent prospects", *Journal of Economic Theory* 7 (1974) 231–240.

[6] J. Keilson, "On global extrema for a class of symmetric functions", *Journal of Mathematical Analysis and Applications* 18 (1967) 218–228.

[7] J. Keilson, "A theorem on optimum allocation for a class of symmetric multilinear return functions", *Journal of Mathematical Analysis and Applications* 15 (1966) 269–272.

[8] L. Mirski, *An introduction to linear algebra* (Oxford University Press, Oxford, 1961).

[9] P.A. Samuelson, "General proof that diversification pays", *Journal of Financial Quantitative Analysis* 2 (1967) 66–84.

[10] E. Seneta, *Non-negative matrices; introduction to theory and applications* (Wiley, New York, 1973).

[11] A. Tamir, "Symmetry and diversification of interdependent prospects", Tech. Rept. Northwestern University (1975).