ON SYMMETRIC MATRICES WHOSE EIGENVALUES SATISFY LINEAR INEQUALITIES[1]

FRITZ JOHN

The nonnegative symmetric matrices (a_{ik}) of order n are those for which

$$\sum_{i,k=1}^{n} a_{ik}\xi_i\xi_k \geqq 0 \qquad \text{for all } \xi, \tag{1a}$$

at the same time these are also the symmetric matrices whose eigenvalues λ_i satisfy

$$\lambda_1 \geqq 0, \cdots, \lambda_n \geqq 0. \tag{1b}$$

Moreover every such matrix can be represented as a sum of n matrices of the same type that have rank $\leqq 1$ (corresponding to the representation of definite forms as sums of squares of linear forms).

We have here an instance of a system of linear inequalities (1b) for the *eigenvalues* of a matrix that is equivalent to a system of linear inequalities (1a) for the *elements* of the matrix. The present note shows that generally systems of linear inequalities for the eigenvalues of a symmetric matrix (satisfied irrespective of the arrangement of the eigenvalues) are equivalent to suitable systems of linear inequalities for the matrix elements.[2] Instead of solutions of systems of linear inequalities we shall talk of the convex sets formed by such solutions.

The general real nth order matrix $a=(a_{ik})$ will be represented by a point in E_m where $m=n^2$. Let Σ be the set of real symmetric matrices of order n and Ω that of real orthogonal matrices. Each $a\in\Sigma$ gives rise to an unordered set of n real eigenvalues that can be represented by points in E_n. For given $a\in\Sigma$ we denote by Λ_a the set of all points $\lambda=(\lambda_1, \cdots, \lambda_n)$ in E_n whose coordinates λ_i are the eigenvalues of a in some order.[3]

Conversely we denote for a given $\lambda=(\lambda_1, \cdots, \lambda_n)$ in E_n by A_λ the set of all symmetric matrices a whose eigenvalues suitably arranged

Received by the editors October 17, 1965.

[1] This research was supported by the Office of Naval Research under Contract No. Nonr-285 (46). Reproduction in whole or in part is permitted for any purpose of the U. S. Government.

[2] The restriction to *symmetric* matrices is essential. For example positiveness of the eigenvalues of a general matrix is not expressible by *linear* inequalities on the elements.

[3] The set Λ_a has at most $n!$ elements.

are $\lambda_1, \cdots, \lambda_n$. We associate with a point $\lambda=(\lambda_1, \cdots, \lambda_n)$ in E_n the diagonal matrix with diagonal elements $\lambda_1, \cdots, \lambda_n$ which we shall denote by $[\lambda]$. The relations

$$\lambda \in \Lambda_a \quad \text{or} \quad a \in A_\lambda \tag{2a}$$

are then equivalent to the existence of an orthogonal matrix c such that[4]

$$a = c[\lambda]c^T. \tag{2b}$$

The correspondence between λ and a implied by (2a, b) leads naturally to two kinds of mappings of sets in E_n onto subsets of Σ in E_m. Given a set σ of points in E_n we define the sets[5]

$$C(\sigma) = \left\{a \mid a \in \Sigma,\ \Lambda_a \in \sigma\right\},$$

$$D(\sigma) = \left\{a \mid a \in \Sigma,\ a \in A_\lambda \text{ for some } \lambda \in \sigma\right\}.$$

THEOREM. *If σ is a closed convex set in E_n then $C(\sigma)$ is closed and convex. If the closed convex set σ is invariant under all permutations of coordinate axes and is the convex hull of a set τ then $C(\sigma)$ is the convex hull of $D(\tau)$. If here τ is a cone with vertex at the origin then $C(\sigma)$ is the set of all matrices that are representable as sum of n matrices in $D(\tau)$.*

(The representation of nonnegative quadratic forms as sums of n squares is a special case. Here σ is the set in E_n described by (1b). We can choose for τ the subset where all but one of the λ_i vanish. Then $D(\tau)$ consists of the nonnegative matrices of rank $\leqq 1$.)

PROOF OF THE THEOREM. A closed convex set σ can be represented as intersection of closed half-spaces. There exists then a system of linear inequalities

$$\mu_1\lambda_1 + \cdots + \mu_n\lambda_n \geqq p \tag{3}$$

with coefficients $(\mu_1, \cdots, \mu_n, p) = (\mu, p)$ forming a set M in E_{n+1} that completely characterises the points $\lambda=(\lambda_1, \cdots, \lambda_n)$ of σ. For a given (μ, p) the points λ satisfying (3) form a closed half-space $H_{\mu,p}$. We have

---

[4] The superscript T denotes transposition.

[5] $C(\sigma)$ is the set of symmetric matrices whose eigenvalues in all possible arrangements represent points of σ, while $D(\sigma)$ is the set of symmetric matrices whose eigenvalues when *suitably* arranged represent points in σ. We have $C(\sigma)=D(\sigma)$ when the set σ is invariant under all permutations of coordinates.

$$\sigma = \bigcap_{(\mu,p)\in M} H_{\mu,p}, \qquad C(\sigma) = \bigcap_{(\mu,p)\in M} C(H_{\mu,p}).$$

To show closedness and convexity of $C(\sigma)$ it is sufficient to show that the set $C(H_{\mu,p})$ is closed and convex for each μ, p. It suffices to prove that $C(H_{\mu,p})$ is identical with the set of matrices $a\in\Sigma$ satisfying the system of linear inequalities

$$\text{(4)} \qquad \operatorname{trace}(ab) \geqq p \qquad \text{for all } b \in A_\mu.$$

Let indeed a be a symmetric matrix satisfying (4) and let $\lambda\in\Lambda_a$. Then $a=c[\lambda]c^T$ for some $c\in\Omega$. We have

$$b = c[\mu]c^T \in A_\mu$$

and hence

$$\operatorname{trace}(ab) = \operatorname{trace}([\lambda][\mu]) = \sum_{i=1}^{n} \lambda_i\mu_i \geqq p.$$

Thus $\lambda\in H_{\mu,p}$ whenever $\lambda\in\Lambda_a$ which proves that $a\in C(H_{\mu,p})$.

Conversely assume that

$$\text{(5)} \qquad a \in C(H_{\mu,p}).$$

Let $\lambda\in\Lambda_a$ and $b\in A_\mu$. Then

$$a = c[\lambda]c^T \quad \text{and} \quad b = d[\mu]d^T \quad \text{where} \quad c, d \in \Omega.$$

Denote by $e=(e_{ik})$ the orthogonal matrix $e=c^Td$ and by $f=(f_{ik})$ the matrix with elements $f_{ik}=e_{ik}^2$. Then

$$\begin{aligned}\operatorname{trace}(ab) &= \operatorname{trace}(c[\lambda]c^Td[\mu]d^T)\\ &= \operatorname{trace}([\lambda]e[\mu]e^T) = \sum_{i,k} f_{ik}\lambda_i\mu_k.\end{aligned}$$

Since e is orthogonal the matrix f is doubly-stochastic.[6] By Birkhoff's theorem (see [1, p. 97]) the set of doubly-stochastic matrices forms a convex polyhedron in E_m with the permutation matrices as vertices. Hence there exist nonnegative numbers $\epsilon_1, \cdots, \epsilon_N$ of sum 1 such that

$$\sum_{j=1}^{n} f_{ik}\lambda_i = \sum_{j=1}^{N} \epsilon_j\lambda_k^{(j)},$$

where for each j the numbers $\lambda_1^{(j)}, \lambda_2^{(j)}, \cdots, \lambda_n^{(j)}$ form a permutation

[6] That is the elements of f are nonnegative and those in any row or column add up to 1.

of the values $\lambda_1, \cdots, \lambda_n$. Since $\lambda^{(j)} \in \Lambda_a$ it follows from (5) that

$$\operatorname{trace}(ab) = \sum_{j=1}^{N} \epsilon_j \sum_{k=1}^{n} \mu_k \lambda_k^{(j)} \geqq \sum_{j=1}^{N} p\epsilon_j = p.$$

Thus a matrix $a \in C(H_{\mu,p})$ satisfies (4). We have proved that $C(\sigma)$ is closed and convex.

Assume now that the closed convex set σ is the convex hull of a set τ, and that σ is invariant under all permutations of coordinates. Then $C(\sigma) = D(\sigma)$. Since $\tau \subset \sigma$ we have

$$D(\tau) \subset D(\sigma) = C(\sigma).$$

It follows from the proved convexity of $C(\sigma)$ that the convex hull of $D(\tau)$ is contained in $C(\sigma)$. Conversely let $a \in C(\sigma)$. Then a is of the form $a = c[\lambda]c^T$ with $\lambda \in \sigma$ and $c \in \Omega$. Since λ is in the convex hull of τ we can find nonnegative numbers $\epsilon_1, \cdots, \epsilon_N$ of sum 1 and points $\lambda^{(1)}, \cdots, \lambda^{(N)}$ in τ such that

$$\lambda = \sum_{j=1}^{N} \epsilon_j \lambda^{(j)}.$$

Then

$$a = \sum_{j=1}^{N} \epsilon_j c[\lambda^{(j)}]c^T. \tag{6}$$

Here the matrices $c[\lambda^{(j)}]c^T$ belong to $D(\tau)$, and consequently a lies in the convex hull of $D(\tau)$. It follows that $C(\sigma)$ is the convex hull of $D(\tau)$.

In the special case where τ is a cone[7] with vertex at the origin we can assume that N has the value n, since by Carathéodory's theorem (see [2, p. 35]) any point of the convex hull of a *connected* set in E_n is the centroid of n nonnegative masses located in (not necessarily distinct) points of the set. We can write (6) in the form

$$a = \sum_{j=1}^{n} c[\epsilon_j \lambda^{(j)}]c^T$$

and have represented a as sum of n elements of $D(\tau)$, since with $\lambda^{(j)}$ also $\epsilon_j \lambda^{(j)}$ lies in τ. Actually the sum of any number of matrices in $D(\tau)$ also lies in $C(\sigma)$ since it belongs to the convex hull of $D(\tau)$.

---

[7] The set τ is a cone with vertex at the origin if $(\lambda_1, \cdots, \lambda_n) \in \tau$ implies $(\gamma\lambda_1, \cdots, \gamma\lambda_n) \in \tau$ for any real $\gamma \geqq 0$. Cones are connected sets.

EXAMPLE. Let σ be the set of points $\lambda=(\lambda_1, \cdots, \lambda_n)$ described by the inequalities

$$0 \leqq \lambda_i \leqq \tfrac{1}{2}(\lambda_1 + \cdots + \lambda_n) \qquad \text{for } i = 1, 2, \cdots, n. \tag{7}$$

(For $n=3$ the system (7) is equivalent to the triangle inequalities

$$\lambda_1 \leqq \lambda_2 + \lambda_3, \qquad \lambda_2 \leqq \lambda_3 + \lambda_1, \qquad \lambda_3 \leqq \lambda_1 + \lambda_2.) \tag{8}$$

Obviously σ is a closed convex set and invariant under permutations of the λ_i. Let τ be the set consisting of the points

$$\lambda = (t, t, 0, 0, \cdots, 0) \qquad \text{with } t \geqq 0 \tag{9}$$

and all those obtained from (9) by permutation of coordinates. Clearly τ is a cone with vertex at the origin contained in σ. Hence the convex hull η of τ also is contained in σ. Actually $\eta=\sigma$, for, as is easily seen, η is closed. If $\eta\neq\sigma$ there would exist a point $\lambda=(\lambda_1, \cdots, \lambda_n)$ satisfying all inequalities (7) and also satisfying an additional inequality

$$\mu_1\lambda_1 + \cdots + \mu_n\lambda_n \geqq p \tag{10}$$

which does not hold for any point in τ. This implies that

$$t(\mu_i + \mu_k) < p \qquad \text{for } i \neq k \text{ and all } t \geqq 0$$

and consequently

$$\mu_i + \mu_k \leqq 0 < p \qquad \text{for } i \neq k. \tag{11}$$

If none of the quantities μ_i were positive relation (10) could not hold since $p>0$ and $\lambda_i\geqq 0$ for all i. If, say, $\mu_1>0$ it follows from (7), (11) that

$$\mu_1(\lambda_1 - \lambda_2 - \lambda_3 - \cdots - \lambda_n) + (\mu_1 + \mu_2)\lambda_2 + (\mu_1 + \mu_3)\lambda_3 + \cdots + (\mu_1 + \mu_n)\lambda_n \leqq 0 < p$$

contrary to (10). Thus σ is the convex hull of τ.

From our theorem we find then that every matrix of $C(\sigma)$ is representable as sum of n matrices in $D(\tau)$. A matrix b belongs to $D(\tau)$ iff it is of the form $b=c[\lambda]c^T$ with $c\in\Omega$ and λ given by (9). We can write $[\lambda]=-g^2$ where $g=(g_{ik})$ is the skew matrix with elements

$$g_{ik} = t^{1/2}(\delta_{i2}\delta_{k1} - \delta_{i1}\delta_{k2}).$$

Then $b=c[\lambda]c^T=-(cgc^T)^2$ where cgc^T is again a skew matrix of rank $\leqq 2$. Hence

Every symmetric matrix of order n whose eigenvalues λ_i satisfy (7) *can be represented as the negative of a sum of n squares of real skew matrices of rank* $\leqq 2$.

Conversely we can show

The negative of any sum of squares of skew matrices is a symmetric matrix with eigenvalues satisfying (7). For any real skew matrix g has eigenvalues that are either zero or conjugate pure imaginary in pairs. Then $-g^2$ is symmetric and has eigenvalues that are either 0 or positive and equal in pairs. The eigenvalues of $-g^2$ satisfy (7) so that $-g^2 \in C(\sigma)$. Since $C(\sigma)$ is a convex cone with vertex at the origin any sum of matrices of the form $-g^2$ where g is skew also lies in $C(\sigma)$.

Added in proof. An instance of equivalence of inequalities for matrix elements and for eigenvalues of the matrix is due to R. Hill [cf. C. Truesdell and R. A. Toupin, *Correction to our paper "Static grounds for inequalities in finite strain of elastic materials,"* Arch. Rational Mech. Anal. **19** (1965), 407]. Hill finds that the third order symmetric matrices a for which

$$\text{trace}\,(gag) = \text{trace}\,(ag^2) < 0$$

for all skew symmetric $g \neq 0$ are exactly those whose eigenvalues λ_i satisfy $\lambda_i + \lambda_k > 0$ for $i \neq k$. This corresponds to our inequalities (3), (4) with $p = 0$ and $\mu = (t, t, 0)$ where $t > 0$.

References

1. M. Marcus and H. Minc, *A survey of matrix theory and matrix inequalities*, Allyn and Bacon, Boston, Mass., 1964.

2. H. G. Eggleston, *Convexity*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 47, Cambridge Univ. Press, New York, 1958.

Courant Institute of Mathematical Sciences, New York University