NOTE ON THE FUNDAMENTAL THEOREM ON IRREDUCIBLE NON-NEGATIVE MATRICES

by HANS SCHNEIDER

(Received 12th February 1958)

1. Let $A=[a_{ij}]$ be an n-th order irreducible non-negative matrix. As is very well-known, the matrix A has a positive characteristic root ρ (provided that $n>1$), which is simple and maximal in the sense that every characteristic root λ satisfies $|\lambda|\leqslant\rho$, and the characteristic vector x belonging to ρ may be chosen positive. These results, originally due to Frobenius, have been proved by Wielandt (4) by means of a strikingly simple basic idea. Recently, a variant of Wielandt's proof has been given by Householder (2).

We shall sketch part of the proof. For each non-negative column vector y we set

$$\rho_*(y)=\sup r : ry\leqslant Ay, \qquad (1)$$

$$\rho^*(y)=\inf r : ry\geqslant Ay. \qquad (2)$$

For *strictly positive* y, we may replace (1) and (2) by

$$\rho_*(y)=\min_i \frac{(Ay)_i}{y_i}=\min_i \frac{\Sigma_{j=1}^n a_{ij}y_j}{y_i}, \qquad (1')$$

$$\rho^*(y)=\max_i \frac{(Ay)_i}{y_i}=\max_i \frac{\Sigma_{j=1}^n a_{ij}y_j}{y_i}. \qquad (2')$$

Let P be the section of the non-negative cone (all $y_i\geqslant 0$) by the plane $\Sigma_{i=1}^n y_i=1$. Since $\rho_*(\lambda y)=\rho_*(y)$, for all positive λ, the supremum of $\rho_*(y)$ over P equals the supremum of $\rho_*(y)$ over all $y\geqslant 0$. On P, $\rho_*(y)$ attains this supremum, say $\rho=\sup\rho_*(y)=\rho_*(x)$, where x is on P. It is then shown that $Ax=\rho x$, that ρ is simple and maximal, and that $x>0$. Similarly the infimum of $\rho^*(y)$ over all $y\geqslant 0$ is attained on P.

2. In this argument there arises a dilemma :

(i) *Either* the whole of P is considered, in which case $\rho_*(y)$ (or $\rho^*(y)$) may have singularities and discontinuities at vectors y which have zero elements ;

(ii) *Or*, the subset P_1 of P, consisting of all positive y on P, is considered, in which case $\rho_*(y)$ (and $\rho^*(y)$) are everywhere continuous on P_1, but P_1 is not closed.

In either case, some justification is required for the assertion that $\rho_*(y)$ attains its supremum (or $\rho^*(y)$ its infimum).

For an example of a discontinuity in $\rho_*(y)$, examine the case of

$$A=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & 0 & 2 \end{bmatrix}$$

https://doi.org/10.1017/S0013091500010816 Published online by Cambridge University Press

At $z=(0, 0, 1)$, we have $\rho_*(z)=2$. But for all y of form $(\alpha, \alpha, 1-2\alpha)$ with $0<\alpha \leqslant \frac{1}{4}$, $\rho_*(y)=1$.

It may however be shown that $\rho_*(y)$ is upper semi-continuous on P, and hence $\rho_*(y)$ attains its supremum on P. For $\rho^*(y)$ slightly more complex reasoning is required. If R_1 denotes the real line with infinity adjoined, topologised in the normal way, then $\rho^*(y)$ is a lower semi-continuous function into R_1, for all non-negative A, and is even continuous when A is irreducible. Again it follows that $\rho^*(y)$ attains its infimum, on P.

3. In this note we shall demonstrate an alternative method of resolving the dilemma. We begin by proving an inequality, which is of some intrinsic interest when applied to the characteristic vector x.

An inequality : Let A *be an n-th order non-negative irreducible matrix,* κ *a least diagonal element of* A, *and* λ *a least* non-vanishing *non-diagonal element. Let y be a positive column vector, and suppose that* $y_1 \geqslant y_2 \geqslant \ldots \geqslant y_n$. *If* $\rho^*(y) \leqslant$ M, *then*

$$\frac{y_n}{y_1} \geqslant \left(\frac{\lambda}{M-\kappa}\right)^{n-1} \quad \ldots\ldots(3)$$

Proof : Let $1 \leqslant k < n$. Then for all $i > k$,

$$My_i \geqslant \Sigma_{j=1}^{n} a_{ij} y_j \geqslant \Sigma_{j=1}^{k} a_{ij} y_j + a_{ii} y_i$$

whence

$$(M-\kappa) y_{k+1} \geqslant (M-\kappa) y_i \geqslant (\Sigma_{j=1}^{k} a_{ij}) y_k. \quad \ldots\ldots(4)$$

Since A is irreducible, here is at least one $i > k$ for which $\Sigma_{j=1}^{k} a_{ij} \geqslant \lambda > 0$. Hence it follows from (4) that

$$\frac{y_{k+1}}{y_k} \geqslant \frac{\lambda}{M-\kappa}. \quad \ldots\ldots(5)$$

The inequality (3) is obtained from (5) by setting $k=1, \ldots, n-1$ and multiplying.

Corollary : If, in addition, $\Sigma_{i=1}^{n} y_i = 1$, *then for all i*

$$y_i \geqslant \frac{1}{n}\left(\frac{\lambda}{M-\kappa}\right)^{n-1}. \quad \ldots\ldots(6)$$

For

$$1 = \Sigma_{i=1}^{n} y_i \leqslant n y_1 \leqslant n \left(\frac{M-\kappa}{\lambda}\right)^{n-1} y_n.$$

4. We now turn to the proof of the fundamental theorem. We shall consider $\rho^*(y)$ only. Set

$$R = \max_i \Sigma_{j=1}^{n} a_{ij},$$

and

$$\delta = \frac{1}{n}\left(\frac{\lambda}{R-\kappa}\right)^{n-1}.$$

Clearly $R \geqslant k + \lambda$, whence

$$\delta \leqslant \frac{1}{n}. \quad \ldots\ldots(7)$$

https://doi.org/10.1017/S0013091500010816 Published online by Cambridge University Press

Let P_2 be that part of the plane section of the non-negative cone defined by $\Sigma_{i=1}^{n} y_i = 1$, and $y_i \geqslant \delta$, for all i. Evidently P_2 is bounded and closed, $\rho^*(y)$ is continuous everywhere on P_2, and hence $\rho^*(y)$ attains its infimum over P_2, say at the vector x on P_2. Set $\rho = \rho^*(x) = \inf \rho^*(y)$ over P_2.

We note that ρ is also the infimum of $\rho^*(y)$ over all positive y. It is sufficient to consider the infimum on the plane section P_1, and to show that there exists a z on P_2 for which $\rho^*(z) < \rho^*(y)$ for all y on P_1 which do not lie on P_2. By (7), the vector $z = \frac{1}{n}(1, 1, \ldots, 1)$ belongs to P_2, while by (6),

$$\rho^*(y) > R = \rho^*(z) \geqslant \rho, \qquad \text{.................................(8)}$$

if y belongs to P_1 but not to P_2.

5. In this section we shall show that ρ is a characteristic root of A, and that the positive vector x is a characteristic vector belonging to ρ, viz. that $Ax = \rho x$. We shall prove the equivalent proposition : *If* $z > 0$ *and* $\rho_*(z) < \rho^*(z)$, *then* $\rho < \rho^*(z)$. Our proof is entirely due to Householder (**2**). Suppose that

$$\frac{(Az)_i}{z_i} = \rho^*(z), \quad i = 1, \ldots, k < n \qquad \text{..........................(9)}$$

$$\frac{(Az)_i}{z_i} < \rho^*(z), \quad i = k+1, \ldots, n. \qquad \text{.......................(10)}$$

Since A is irreducible there exists a positive element a_{pq} with $1 \leqslant p \leqslant k$ and $k+1 \leqslant q \leqslant n$. Define the vector z' by setting $z_i' = z_i$ if $i \neq q$ and $0 < z_q' < z_q$, where z_q' is chosen sufficiently close to z_q to ensure that

$$\frac{(Az')_q}{z_q'} < \rho^*(z). \qquad \text{..................................(11)}$$

But for all $i \neq q$

$$\frac{(Az')_i}{z_i'} \leqslant \frac{(Az)_i}{z_i}. \qquad \text{..................................(12)}$$

It follows from (9), (10), (11) and (12) that

$$\rho^*(z') \leqslant \rho^*(z).$$

Since

$$\frac{(Az')_p}{z_p'} < \frac{(Az)_p}{z_p}$$

the equality

$$\frac{(Az')_i}{z_i'} = \rho^*(z)$$

can hold for at most $k-1$ indices i. Thus, by repetition of this process we may construct a vector z'' satisfying $\rho^*(z'') < \rho^*(x)$. We deduce that $\rho < \rho^*(z)$.

6. If ρ is not a simple characteristic root of A, suppose first that ρ has the linearly independent characteristic vectors $x > 0$ and z belonging to it. By choosing the real numbers α and β suitably, we may obtain a positive characteristic vector $\alpha x + \beta z$ on P_i belonging to ρ, which has some element less than δ. But this is impossible, by (6) and (8). Next, suppose that ρ is not simple, but has only one linearly independent characteristic vector belonging to it.

Then there exists a column vector y satisfying

$$Ay=\rho y-x, \qquad (13)$$

and replacing y by $y+\gamma x$, it follows there exists a positive y satisfying (13). For this y, $\rho^*(y)<\rho$, which is impossible. We have proved that ρ is a simple characteristic root.

7. For the sake of completeness, we add a standard proof that ρ is a maximal characteristic root. Let σ be any characteristic root of A and u a row vector satisfying $uA=\sigma u$. Then, for all j,

$$|\sigma|\,|u_j|=|\Sigma_{i=1}^n u_i a_{ij}|\leqslant \Sigma_{i=1}^n |u_i|\,a_{ij}$$

whence

$$|\sigma|\,\Sigma_{j=1}^n |u_j|\,x_j \leqslant \Sigma_{i,j=1}^n |u_i|\,a_{ij}x_j=\rho\Sigma_{i=1}^n |u_i|\,x_i,$$

and so

$$|\sigma|\leqslant\rho,$$

since

$$\Sigma_{i=1}^n |u_i|\,x_i>0.$$

8. The inequality (3) leads to a positive lower bound for the ratios of the elements of the characteristic vector x. In view of $\rho^*(x)=\rho$ and (8) it follows that

$$\frac{\min_i x_i}{\max_i x_i}\geqslant\left(\frac{\lambda}{\rho-\kappa}\right)^{n-1}\geqslant\left(\frac{\lambda}{R-\kappa}\right)^{n-1}.$$

For the case $A>0$, lower bounds for the ratios x_j/x_i have already been found by Ostrowski (3) and A. Brauer (1). These bounds, however, reduce to 0 when A has zero elements.

REFERENCES

(1) A. Brauer, The theorems of Ledermann and Ostrowski on positive matrices, *Duke Math. J.*, **24** (1957), 265-274.

(2) A. S. Householder, *On the convergence of matrix iterations*, Oak Ridge National Laboratory, Physics No. 1883 (1955).

(3) A. Ostrowski, Bounds for the greatest latent root of a positive matrix, *J. London Math. Soc.*, **27** (1952), 253-256.

(4) H. Wielandt, Unzerlegbare, nicht negative Matrizen, *Math. Zeitschrift*, **52** (1950), 642-8.

The Queen's University
Belfast