COWLES COMMISSION DISCUSSION PAPER: MATHEMATICS NO. 410

Positive and Nonnegative Matrices I.

I. N. Herstein

October 2, 1951

In a series of three papers in the Sitzungsberichte der Berliner Akademie, 1908, 1909 and 1912 [1] Frobenius proved some interesting and very general results about the nature of the characteristic roots of matrices all of whose elements are positive, and some extensions of these results to the case of matrices whose entries are all nonnegative. These results seem to be of interest to economists, for special cases of Frobenius' theorems have been used or proved recently by several authors [2, 3, 4]. One of the greatest motivations for writing this paper has been the feeling on the author's part that these widesweeping results of Frobenius should be more widely known. We should like to stress that the proofs presented here, while not exactly those of Frobenius', use ideas that can be found scattered through the three papers cited; and in a good many places are merely translations, from the German, of Frobenius' work. In this paper we consider the case of positive matrices, and in a future paper we hope to discuss the case of nonnegative matrices.

Let C be an nxn matrix of rank n-1. Then as is well known [5. P. 47, theorem 4].

Lemma 1. *If* $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *is a solution of* $Cx = 0$, *the* x_i*'s are proportional to the cofactors of the elements of any row of* C.

Suppose now that $A = (a_{ij})$ is an nxn matrix such that all the $a_{ij} > 0$. We then call A a positive matrix and denote this by writing $A > 0$.

The "Fundamental Theorem" proved by Frobenius in his 1908 paper is

Theorem 2. *Suppose* $A > 0$. *Then*

1. *A has a positive characteristic root*

2. *if r is its largest positive characteristic root, r is a simple root and for any other characteristic root, p, of A*

$$r > |p|$$

3. *if* $s \geq r$ *the cofactors of the elements of* $sI-A$ *are all positive, where I is the identity matrix.*

Proof. Let $\varphi(s) = A(s) = \det(sI-A)$.

$$(1) \qquad A(s) = \begin{vmatrix} s-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & s-a_{22} & \cdots & -a_{2n} \\ \vdots & & & \\ -a_{n1} & \cdots & & s-a_{nn} \end{vmatrix}$$

In $A(s)$ let $A_{ij}(s)$ denote the cofactor of the (i,j) element. Thus $B(s) = A_{11}(s)$ is the part of $A(s)$ enclosed by $\ulcorner$. In $B(s)$ let the cofactor of the element $s\,\delta_{ij} - a_{ij}$, $(i, j > 1$, δ_{ij} = Kronecker delta) be $B_{ij}(s)$.

We proceed by induction, assuming the theorem true for all $B > 0$ of order $n-1$. Thus $B(s) = 0$ has a positive, simple root q which is the largest root in absolute value, and is such that if $s \geq q$, $B_{ij}(s) > 0$.

Expanding $A(s)$ we obtain

$$(2) \quad A(s) = (s-a_{11})\,B(s) - \sum_{i=2,j=2}^{i=n,j=n} a_{1i}\, a_{j1}\, B_{ji}(s)\,.$$

Putting $s = q$ in (2) and using $B(q) = 0$, $B_{k\ell}(q) > 0$, we obtain, since the $a_{ij} \geq 0$, $A(q) < 0$. However $A(s)$ is a polynomial in with leading coefficient 1 so as $s \longrightarrow +\infty$, $A(s) \longrightarrow +\infty$. Thus $A(s) = 0$ must have a real root t, $t > q > 0$. This is the first assertion of the theorem. Let r be the largest positive characteristic root of A.

Differentiating $\varphi(s) = A(s)$ [see (1)] with respect to s we obtain

(3) $$\varphi'(s) = \sum_i A_{ii}(s).$$

Just as for $B(s) = A_{11}(s)$ we had $r > q$, $r >$ largest positive root of $A_{ii}(s)$. Thus for all $s \geq r$, $A_{ii}(s) > 0$, so for all $s \geq r$

(4) $$\varphi'(s) > 0.$$

In particular $\varphi'(r) > 0$ and so r is a simple root of $A(s) = 0$. (This is the first half of the second assertion of the theorem.)

From (2) and the fact that $A(r) = 0$ we have

(5) $$0 = (r-a_{11})\, B(r) - \sum_{\substack{i=2\\ j=2}}^{n} a_{1i}\, a_{j1}\, B_{ji}(r).$$

Differentiating this partially with respect to a_{1K}, $K \neq 1$ we have

(6) $$0 = \frac{\partial r}{\partial a_{1K}} [B(r) + (r-a_{11})\, B'(r) - \sum_{i,j} a_{1i}\, a_{j1}\, B'_{ji}(r)] - \sum_j a_{j1}\, B_{jK}(r) .$$

The part of (6) in brackets is merely $\varphi'(r)$ and so is positive, and since $a_{j1} > 0$ $B_{jK}(r) > 0$, (since $r > q$) $\sum_j a_{j1}\, B_{jK}(r) > 0$. Thus we obtain

(7) $$\frac{\partial r}{\partial a_{1K}} > 0 \quad , \qquad K \neq 1 .$$

If, on the other hand we differentiate

$$A(r) = \begin{vmatrix} r-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & r-a_{22} & & \\ \vdots & & & \\ -a_{n1} & \cdots & \cdots & r-a_{nn} \end{vmatrix} = 0$$

with respect to a_{1K} we obtain

(8) $$\varphi'(r) \frac{\partial r}{\partial a_{1K}} = A_{1K}(r), \qquad K \neq 1.$$

Both quantities on the left hand side of (8) are positive, so $A_{1K}(r) > 0$. Since $r > q$, by the induction $A_{11}(r) > 0$. So for all K, $A_{1K}(r) > 0$. Had we expanded the determinant $A(s)$ in terms of the i^{th} row we would have obtained in the same way that $A_{iK}(r) > 0$. This is then the third assertion of the theorem.

All that remains to be proved is the second half of the second assertion.

Suppose $p \neq r$ is a characteristic root of A, $|p| = g$. Thus there exist numbers $x_1, \ldots, x_n$ so that

(9) $\sum_j a_{ij} x_j = p\, x_i$.

Let $y_j = |x_j|$. Then

(10) $\left| \sum_j a_{ij} x_j \right| < \sum a_{ij} y_j$.

The possibility of the equality in (10) is ruled out as follows: equality occurs only if for each j,

$x_j = y_j e^{i\theta}$, in which case

$\sum a_{ij} y_j = p\, y_i$, whence p would be positive, and $p \neq r$ implies $p < r$ since r is a simple root.

From (9) and (10) we have

(11) $\sum_j a_{ij} y_j > g\, y_i$ where $g = |p|$.

Now, since $rI-A$ is of rank $n-1$ and in $rI-A$ the cofactor of any element is positive, and since this is also true for the transpose of $rI-A$, by lemma 1 there exist z_i, all positive, so that

(12) $\sum_i a_{ij} z_i = r\, z_j$.

Now

$$r \sum_j y_j z_j = \sum_{i,j} a_{ij} z_i y_j = \sum_i z_i \left(\sum_j a_{ij} y_j\right) > g \sum_i y_i z_i$$

from (11) and (12). Since $\sum_i y_i z_i > 0$, we obtain $r > g$, completing the induction and proving the theorem.

In what follows $A > 0$ and r its largest positive characteristic root.

Theorem 3. $\frac{\partial r}{\partial a_{ij}} > 0$.

This was, in effect, proved in the course of proving theorem 2. It interprets as: an increase in any element of A yields an increase in the largest positive characteristic root of A.

Let $\alpha_i = \sum_j a_{ij}$. Then

Theorem 4. $\min_i \alpha_i \leq r \leq \max_i \alpha_i$.

Proof. There exists a vector $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $x_i > 0$, so that $Ax = rx$. Let $x_m = \min_i x_i$.

Now $\sum_j a_{mj} x_j = rx_m \geq (\sum_j a_{mj})x_m$,

whence $r \geq \sum_j a_{mj} \geq \min_i \alpha_i$. We proceed analagously for the other half of the inequality. Clearly similar bounds could have been obtained for the column sums, and from these, a combination with the row bounds might, in practice, yield tighter bounds on r.

Of side interest is the following result which we mention here without proof, which gives a means of approximating r. (Its proof is in the 1908 paper.) Let $a_{ij}^{(\lambda)}$ be the (i,j) element of A^{λ}.

Then $\lim_{\lambda \to \infty} \frac{a_{ij}^{(\lambda+1)}}{a_{ij}^{(\lambda)}} = r$.

REFERENCES

[1] Frobenius, G., "Über Matrizen aus positiven Elementen," Sitz. Berl. Akad., 1908, p. 471.

______________, "Über Matrizen aus positiven Elementen II," Sitz. Berl. Akad., 1909, p. 514.

______________, "Über Matrizen aus nicht negativen elementen," Sitz. Berl. Akad, 1912, p. 456.

[2] Metzler, L., "Stability of Multiple Markets: The Hick's Condition," Econometrica, 1945, p. 277-292, in particular p. 286-287.

[3] Chipman, J., "Stability of Systems with Non-Negative Coefficients," Cowles Commission Discussion Paper, Economics 2016 & Supplement.

[4] Solow, R., "On the Structure of Linear Models," (unpublished) via Chipman [3].

[5] Bôcher, M., Introduction to Higher Algebra, 1924 edition.