①

ON AN ITERATIVE PROCEDURE FOR OBTAINING
THE PERRON ROOT OF A POSITIVE MATRIX

Richard Bellman

P–577

5 October 1954

The RAND Corporation
1700 MAIN ST. · SANTA MONICA · CALIFORNIA

SUMMARY

An iterative procedure for obtaining the characteristic root of largest absolute value of a positive matrix, the Perron Root, is derived which yields geometric convergence.

ON AN ITERATIVE PROCEDURE FOR OBTAINING THE PERRON ROOT

OF A POSITIVE MATRIX

Richard Bellman

§1. Introduction

The purpose of this paper is to present a new iterative procedure for obtaining the characteristic root of largest absolute value of a positive matrix.

The origin of the method is as follows. There is a result of Von Neumann, [7], a generalization of his fundamental min-max theorem in the theory of games, [8], to the effect that

$$(1) \quad \underset{y}{\text{Min}}\ \underset{x}{\text{Max}}\ \frac{(x,Ay)}{(x,By)} = \underset{x}{\text{Max}}\ \underset{y}{\text{Min}}\ \frac{(x,Ay)}{(x,By)}$$

where the variation is over the region defined by

$$(2) \quad \text{(a)} \quad x_i \geqslant 0, \quad \sum_{i=1}^{n} x_i = 1,$$

$$R: $$

$$\text{(b)} \quad y_i \geqslant 0, \quad \sum_{i=1}^{n} y_i = 1,$$

and it is assumed that B has the property that $(x,By) \geqslant b > 0$ for all $(x,y) \in R$.

It was observed by Shapley, [6], that this result can be obtained as a by-product of the theory of "games of survival", cf [1], [5], [6], which requires only the fundamental min-max theorem, by considering the equation for λ,

(3) $\lambda = \underset{y}{\text{Min}} \ \underset{x}{\text{Max}} \ \left[\ (x,Ay) + \lambda (1-(x,By)) \ \right]$,

$= \underset{x}{\text{Max}} \ \underset{y}{\text{Min}} \ \left[\ (x,Ay) + \lambda (1-(x,By)) \ \right]$,

where we impose the additional assumption the $1 > (x,By)$ for all $(x,y) \in R$. This restriction is of no importance because of the homogeneity of the ratio in (1).

It is then easy to prove that there is a unique solution of (3) which may be obtained as the limit of the sequence $\left\{ \lambda_n \right\}$ defined by

(4) $\lambda_0 = \underset{y}{\text{Min}} \ \underset{x}{\text{Max}} \ \left[\ (x,Ay) \ \right] = \underset{x}{\text{Max}} \ \underset{y}{\text{Min}} \ \left[\ (x,Ay) \ \right]$,

$\lambda_{n+1} = \underset{y}{\text{Min}} \ \underset{x}{\text{Max}} \ \left[\ (x,Ay) + \lambda_n (1-(x,By)) \ \right]$,

$= \underset{x}{\text{Max}} \ \underset{y}{\text{Min}} \ \left[\ (x,Ay) + \lambda_n (1-(x,By)) \ \right]$,

and that this solution is the given also by the common value of the ratio in (1).

This procedure yields a theoretical and computational hold on λ, which is quite useful. Furthermore, by means of this ingenious device we have a means of linearizing a number of problems relating to ratios. In this paper we shall apply this idea to the problem of determining the root of a positive matrix of largest absolute value, using a variational representation for this root involving a ratio.

§2. The Perron Root

Let A be a square matrix (a_{ij}). It is called <u>positive</u>[*] if $a_{ij} > 0$ for all i and j. The basic result concerning positive matrices is due to Perron and is the following.

<u>Lemma 1.</u> <u>If A is a positive matrix, there is a unique characteristic root of largest absolute value. This root is positive and its associated characteristic vector may be taken to be positive.</u>

<u>Notation.</u> We denote this root by p(A), the Perron root of A.

An alternative definition of this root, possessing the great merit of involving a variation, is

<u>Lemma 2.</u>

$$(1) \quad p(A) = \underset{x}{\text{Max}} \ \underset{i}{\text{Min}} \ \sum_{j=1}^{n} a_{ij} x_j / x_i$$

$$= \underset{x}{\text{Min}} \ \underset{i}{\text{Max}} \ \sum_{j=1}^{n} a_{ij} x_j / x_i$$

This result has been used by several authors independently, and does not seem to have any particular known origin. It was communicated to the author several years ago by H. Bohnenblust in connection with a different problem, see [2].

Here the variation is over the region defined by

$$(2) \quad x_i \geqslant 0, \ \sum_{i} x_i = 1.$$

---

[*] not to be confused with <u>positive definite</u>

§3.   A Refinement of Lemma 2

Let us show that Lemma 2 may be replaced by the stronger result

Lemma 3.   We have

(1) $p(A) = \underset{R'}{\text{Max}} \; \underset{i}{\text{Min}} \; \sum_{j=1}^{n} a_{ij} x_j / x_i$,

$= \underset{R'}{\text{Min}} \; \underset{i}{\text{Max}} \; \sum_{j=1}^{n} a_{ij} x_j / x_i$,

where R' is defined by

(2) $x_i \geqslant d, \quad \sum_i x_i = 1$,

and d is some parameter depending only upon A.   Specifically, we may take

(3) $d = \underset{i,j}{\text{Min}} \; a_{ij} / \underset{i}{\text{Max}} \; \left(\sum_{j=1}^{n} a_{ij} \right).$

Proof: The minimizing x_i constitute the characteristic vector associated with $p(A)$, normalized by the condition that $\sum_i x_i = 1$. Hence

(4) $p(A) x_i = \sum_{j=1}^{n} a_{ij} x_j$, $i = 1, 2, \ldots, n.$

Thus

(5) $p(A) \; \underset{i}{\text{Min}} \; x_i \geqslant \left(\underset{i,j}{\text{Min}} \; a_{ij} \right) \sum_{j=1}^{n} x_j = \underset{i,j}{\text{Min}} \; a_{ij}$

On the other hand,

$$(6) \quad p(A) \underset{1}{\text{Max}} \; x_1 \leqslant \underset{1}{\text{Max}} \; x_1 \; (\underset{1}{\text{Max}} \sum_{j=1}^{n} a_{ij}) \; ,$$

whence

$$(7) \quad p(A) \leq \underset{1}{\text{Max}} \; (\sum_{j=1}^{n} a_{ij}).$$

Combining (5) and (7) we have

$$(8) \quad \underset{1}{\text{Min}} \; x_1 \geq \underset{1,j}{\text{Min}} \; a_{ij} / \underset{1}{\text{Max}} \; (\sum_{j=1}^{n} a_{ij}).$$

§4. <u>An Alternative Definition of p(A)</u>

Let us now show, following the lead of **Shapley**, that we may define p(A) as follows

<u>Lemma</u> 4. p(A) <u>is the unique solution of</u>

$$(1) \quad \lambda = \underset{R'}{\text{Max}} \; \underset{1}{\text{Min}} \; [\; \sum_{j=1}^{n} a_{ij}x_j + \lambda(1-x_1) \;]$$

<u>or of</u>

$$(2) \quad \lambda = \underset{R'}{\text{Min}} \; \underset{1}{\text{Max}} \; [\; \sum_{j=1}^{n} a_{ij}x_j + \lambda(1-x_1) \;] \; ,$$

<u>where</u> R' <u>is as defined by</u> (3.2).

<u>Proof</u>: It is sufficient to prove that p(A) satisfies (1). The proof of the other statement is similar. We have, for all x in R',

(3) $\lambda \geq \underset{1}{\text{Min}} \left[\sum_{j=1}^{n} a_{1j}x_j + \lambda(1-x_1) \right]$,

for any solution λ, with equality for at least one x. We shall prove below that there is exactly one solution which may be obtained iteratively.

Hence, for all $x \in R'$,

(4) $0 \geq \underset{1}{\text{Min}} \left[\sum_{j=1}^{n} a_{1j}x_j - \lambda x_1 \right]$,

or

(5) $0 \geq \underset{1}{\text{Min}} \left[x_1 \left\{ \sum_{j=1}^{n} a_{1j}x_j/x_1 - \lambda \right\} \right]$,

for all $x \in R'$. Since $x_1 > 0$, it follows that

(6) $\lambda \geq \underset{1}{\text{Min}} \left(\sum_{j=1}^{n} a_{1j}x_j/x_1 \right)$

for all x, with equality for one x, at least. Hence

(7) $\lambda = \underset{R'}{\text{Max}} \underset{1}{\text{Min}} \left(\sum_{j=1}^{n} a_{1j}x_j/x_1 \right) = p(A)$,

which shows uniqueness provided we assume existence.

Similarly we may demonstrate the result in (2)

§5. A Non-Linear Recurrence Relation

Let us now consider the non-linear recurrence relation

(1) $\quad u_{n+1} = \underset{R'}{Min} \underset{i}{Max} \left[\sum_{j=1}^{n} a_{ij}x_j + u_n(1-x_i) \right]$,

where R' is as above and u_0 is arbitrary. We shall prove

Theorem

(2) $\quad p(A) = \underset{n \to \infty}{lim} u_n$

A similar result holds for the recurrence relation based upon (4.2)

Proof: We have

(3) $\quad u_{n+1} = \underset{R'}{Min} \underset{i}{Max} \left[\sum_{j=1}^{n} a_{ij}x_j + u_n(1-x_i) \right]$

$\qquad = \underset{R'}{Min} \underset{y}{Max} \left[\sum_{i=1}^{n} y_i \left[\sum_{j=1}^{n} a_{ij}x_j + u_n(1-x_i) \right] \right]$,

where the maximum in y is over the region $y_i \geq 0$, $\sum_{i=1}^{n} y_i = 1$.

Using the min–max theorem of Von Neumann, this may also be written

(4) $\quad u_{n+1} = \underset{y}{Max} \underset{R'}{Min} \left\{ \sum_{i=1}^{n} y_i \left[\sum_{j=1}^{n} a_{ij}x_j + u_n(1-x_i) \right] \right\}$.

Let us write this recurrence relation in the form

(5) $\quad u_{n+1} = \underset{R'}{Min} \underset{y}{Max} \ T(u_n,x,y) = \underset{y}{Max} \underset{R'}{Min} \ T(u_n,x,y)$

Then using a device we have employed frequently in the theory of dynamic programming, cf [2], [3], we have

(6) $\quad u_{n+1} = T(u_n, \bar{x}, \bar{y})$,

$\quad\quad u_n \quad = T(u_{n-1}, \bar{\bar{x}}, \bar{\bar{y}})$,

where $(\bar{x}, \bar{y})$ and $(\bar{\bar{x}}, \bar{\bar{y}})$ are respectively values where the min—max and max—min are assumed for n and n-1 respectively.

Hence, by virtue of (5)

(7) $\quad u_{n+1} = T(u_n, \bar{x}, \bar{y}) \geq T(u_n, \bar{x}, \bar{\bar{y}})$

$\quad\quad\quad\quad\quad\quad \leq T(u_n, \bar{\bar{x}}, \bar{y})$,

and

(8) $\quad u_n = T(u_{n-1}, \bar{\bar{x}}, \bar{\bar{y}}) \geq T(u_{n-1}, \bar{\bar{x}}, \bar{y})$

$\quad\quad\quad\quad\quad\quad \leq T(u_{n-1}, \bar{x}, \bar{\bar{y}})$.

From this we obtain

(9) $\quad u_{n+1} - u_n \geq T(u_n, \bar{x}, \bar{\bar{y}}) - T(u_{n-1}, \bar{x}, \bar{\bar{y}})$

$\quad\quad\quad\quad\quad \leq T(u_n, \bar{\bar{x}}, \bar{y}) - T(u_{n-1}, \bar{\bar{x}}, \bar{y})$,

which yields

(10) $\quad u_{n+1} - u_n \geq (u_n - u_{n-1}) \sum_{i=1}^{n} \bar{\bar{y}}_i (1 - \bar{x}_i)$

$\quad\quad\quad\quad\quad \leq (u_n - u_{n-1}) \sum_{i=1}^{n} \bar{y}_i (1 - \bar{\bar{x}}_i)$.

Hence

(11) $|u_{n+1} - u_n| \leq |u_n - u_{n-1}|$ Max $\left[\sum_{i=1}^{n} \bar{y}_i (1 - \bar{x}_i), \sum_{i=1}^{n} \bar{y}_i (1 - \bar{\bar{x}}_i) \right]$.

Since $\bar{x}_i$, $\bar{\bar{x}}_i \geq d > 0$, and $y_i \geq 0$, $\sum_i y_i = 1$, we have

(12) $|u_{n+1} - u_n| \leq (1-d) |u_n - u_{n-1}|$,

and hence geometric convergence of $\sum_{w=0}^{\infty} (u_{n+1} - u_n)$. The limit of u_n exists, and must equal $p(A)$. Observe that this is a situation where only the <u>value</u> of a game is of interest, if we wish only to determine $p(A)$. Consequently, the iterative procedure of [4], may be of some merit here.

§6. <u>Monotone Convergence</u>

If we set

(1) $u_0 = \underset{H'}{\text{Min}} \underset{i}{\text{Max}} \sum_{j} a_{ij} x_j$,

we see that $u_1 \geq u_0$ and hence $u_{n+1} \geq u_n$, which ensures monotone convergence.

Similarly, if we have

(2) $a_1 \leq a_{ij} \leq a_2$,

and use the Perron roots of the associated matrices as initial approximations, we obtain monotone increasing and monotone decreasing sequences respectively.

BIBLIOGRAPHY .

1. Bellman, R.,   "An Introduction to the Theory of Dynamic
                   Programming", RAND Report No. R-245, 1953.

2. ——————,   "A Survey of the Theory of Time-Lag, Retarded
                   Control and Hereditary Processes", RAND Report
                   No. R-271, 1954.

3. ——————,   "Some Functional Equations in the Theory of
                   Dynamic Programming – I", (to appear).

4. ——————,   "On a New Iterative Algorithm for Finding the
                   Solutions of Games and Linear Programming Problems",
                   Symposium on Monte Carlo Methods, U. of Florida,
                   March 1954.

5. Bellman, R. and LaSalle, J., "Non-Zero Sum Games and Stochastic
                   Processes", RAND Memorandum No. RM-271, 1949.

6. Shapley, L.,   "Stochastic Games", Proc. Nat. Acad. Sci., Vol. 39,
                   1953, pp 1095-1100.

7. Von Neumann, J., Ergebnisse eines Math. Kolloquiums, 8(1937),
                   pp 73-78.

8. Von Neumann, J., and Morgenstern, "Theory of Games and Economic
                   Behavior", Princeton U. Press, 1944.