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

Richard Bellman

P-577

5 October 1954

Approved for OTS release

12 p

COPY OF

HARD COPY $1.00
MICROFICHE $0.50

DDC
AUG 27 1964
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

\[
\operatorname{Min} \operatorname{Max} \left( \frac{y}{x} \right) = \operatorname{Max} \operatorname{Min} \left( \frac{x}{y} \right)
\]

where the variation is over the region defined by

\[
\begin{align*}
(a) \quad x_i & \geq 0, \quad \sum_{i=1}^{n} x_i = 1, \\
(b) \quad y_i & \geq 0, \quad \sum_{i=1}^{n} y_i = 1,
\end{align*}
\]

and it is assumed that \( B \) has the property that \( (x, By) \geq b > 0 \) 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", of [1], [5], [6], which requires only the fundamental min-max theorem, by considering the equation for \( \lambda \).
(3) \[ \lambda = \min_y \max_x \left[ (x, Ay) + \lambda (1-(x, By)) \right], \]
\[ = \max_x \min_y \left[ (x, Ay) + \lambda (1-(x, By)) \right], \]

where we impose the additional assumption the \( 1 > (x, By) \) for all \( (x,y) \in \mathbb{R} \). 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 \( \{ \lambda_n \} \) defined by

\[ \lambda_0 = \min_y \max_x \left[ (x, Ay) \right] = \max_x \min_y \left[ (x, Ay) \right], \]
\[ \lambda_{n+1} = \min_y \max_x \left[ (x, Ay) + \lambda_n (1-(x, By)) \right], \]
\[ = \max_x \min_y \left[ (x, Ay) + \lambda_n (1-(x, By)) \right], \]

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 \( \lambda \), 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 positive* if \( a_{ij} > 0 \) for all \( i \) and \( j \). The basic result concerning positive matrices is due to Perron and is the following.

**Lemma 1.** 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.

**Notation.** 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

**Lemma 2.**

\[
(1) \quad p(A) = \max \min_{x} \sum_{j=1}^{n} a_{ij}x_{j}/x_{1} \\
= \min \max_{x} \sum_{j=1}^{n} a_{ij}x_{j}/x_{1}
\]

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

\[
(2) \quad x_{i} \geq 0, \quad \sum_{i} x_{i} = 1.
\]

* not to be confused with positive definite
§3. A Refinement of Lemma 2

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

Lemma 3. We have

\[
(1) \quad p(A) = \max_{R'} \min_{1} \sum_{j=1}^{n} a_{ij}x_{j}/x_{i},
\]

\[
= \min_{R'} \max_{1} \sum_{j=1}^{n} a_{ij}x_{j}/x_{i},
\]

where \( R' \) is defined by

\[
(2) \quad x_{i} \geq d, \sum_{i} x_{i} = 1,
\]

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

\[
(3) \quad d = \min_{i,j} a_{ij}/\max_{i} \left( \sum_{j=1}^{n} a_{ij} \right).
\]

Proof: The minimizing \( x_{i} \) constitute the characteristic vector associated with \( p(A) \), normalized by the condition that \( \sum_{i} x_{i} = 1 \). Hence

\[
(4) \quad p(A)x_{i} = \sum_{j=1}^{n} a_{ij}x_{j}, \quad i = 1, 2, \ldots, n.
\]

Thus

\[
(5) \quad p(A) \min_{i} x_{i} \geq (\min_{i,j} a_{ij}) \sum_{j=1}^{n} x_{j} = \min_{i,j} a_{ij}
\]
On the other hand,

\[(6) \quad p(A) \max_{1} x_{1} \leq \max_{1} x_{1} \left( \max_{j=1}^{n} a_{1j} \right), \]

whence

\[(7) \quad p(A) \leq \max_{1} \left( \sum_{j=1}^{n} a_{1j} \right). \]

Combining (5) and (7) we have

\[(8) \quad \min_{1} x_{1} \geq \min_{1,j} a_{1j} / \max_{1} \left( \sum_{j=1}^{n} a_{1j} \right). \]

\[\text{§4. An Alternative Definition of } p(A) \]

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

**Lemma 4.** \( p(A) \) is the unique solution of

\[(1) \quad \lambda = \max_{R'} \min_{1} \left[ \sum_{j=1}^{n} a_{1j} x_{j} + \lambda (1-x_{1}) \right] \]

or of

\[(2) \quad \lambda = \min_{R'} \max_{1} \left[ \sum_{j=1}^{n} a_{1j} x_{j} + \lambda (1-x_{1}) \right], \]

where \( R' \) is as defined by (3.2).

**Proof:** It is sufficient to prove that \( p(A) \) satisfies (1). The proof of the other statement is similar. We have, for all \( x \) in \( R' \),
\( \lambda \geq \min_{i} \left[ \sum_{j=1}^{n} a_{ij}x_{j} + \lambda(1-x_{i}) \right] \),

for any solution \( \lambda \), 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 \( x \in \mathbb{R}^{n} \),

\( 0 \geq \min_{i} \left[ \sum_{j=1}^{n} a_{ij}x_{j} - \lambda x_{i} \right] \),

or

\( 0 \geq \min_{i} \left[ x_{i} \left( \sum_{j=1}^{n} a_{ij}x_{j}/x_{i} - \lambda \right) \right] \),

for all \( x \in \mathbb{R}^{n} \). Since \( x_{i} > 0 \), it follows that

\( \lambda \geq \min_{i} \left( \sum_{j=1}^{n} a_{ij}x_{j}/x_{i} \right) \)

for all \( x \), with equality for one \( x \), at least. Hence

\( \lambda = \max_{\mathbb{R}^{n}} \min_{i} \left( \sum_{j=1}^{n} a_{ij}x_{j}/x_{i} \right) = p(A) \),

which shows uniqueness provided we assume existence.

Similarly we may demonstrate the result in (2)

\[ \frac{55}{5} \text{ A Non-Linear Recurrence Relation} \]

Let us now consider the non-linear recurrence relation
(1) \( u_{n+1} = \min_{R'} \max_R \left[ \sum_{j=1}^{n} a_{ij} x_j + u_n(1-x_1) \right] \),

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

Theorem

(2) \( p(A) = \lim_{n \to \infty} u_n \)

A similar result holds for the recurrence relation based upon

(4.2)

Proof: We have

(3) \( u_{n+1} = \min_{R'} \max_R \left[ \sum_{j=1}^{n} a_{ij} x_j + u_n(1-x_1) \right] \)

\[ = \min_{R'} \max_R \left[ \sum_{i=1}^{n} y_1 \left[ \sum_{j=1}^{n} a_{ij} x_j + u_n(1-x_1) \right] \right], \]

where the maximum in \( y \) is over the region \( y_1 \geq 0, \sum_{i=1}^{n} y_1 = 1 \).

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

(4) \( u_{n+1} = \max_{y} \min_{R'} \left\{ \sum_{i=1}^{n} y_1 \left[ \sum_{j=1}^{n} a_{ij} x_j + u_n(1-x_1) \right] \right\}. \)

Let us write this recurrence relation in the form

(5) \( u_{n+1} = \min_{R'} \max_y T(u_n, x, y) = \max_y \min_{R'} T(u_n, x, y) \)

Then using a device we have employed frequently in the theory of
dynamic programming, cf [2], [3], we have
(6) \[ u_{n+1} = T(u_n, \bar{x}, \bar{y}) , \]
\[ u_n = T(u_{n-1}, \bar{x}, \bar{y}) , \]

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

Hence, by virtue of (5)

(7) \[ u_{n+1} = T(u_n, \bar{x}, \bar{y}) \geq T(u_n, \bar{x}, \bar{y}) \]
\[ \leq T(u_n, \bar{x}, \bar{y}) , \]

and

(8) \[ u_n = T(u_{n-1}, \bar{x}, \bar{y}) \geq T(u_{n-1}, \bar{x}, \bar{y}) \]
\[ \leq T(u_{n-1}, \bar{x}, \bar{y}) . \]

From this we obtain

(9) \[ u_{n+1} - u_n \geq T(u_n, \bar{x}, \bar{y}) - T(u_{n-1}, \bar{x}, \bar{y}) \]
\[ \leq T(u_n, \bar{x}, \bar{y}) - T(u_{n-1}, \bar{x}, \bar{y}) , \]

which yield:

(10) \[ u_{n+1} - u_n \geq (u_n - u_{n-1}) \sum_{i=1}^{n} \bar{y}_i (1 - \bar{x}_1) \]
\[ \leq (u_n - u_{n-1}) \sum_{i=1}^{n} \bar{y}_i (1 - \bar{x}_1) . \]

Hence
Since $x_1$, $x_1 > d > 0$, and $y_1 > 0$, we have

$$\sum_{i=1}^{n} \overline{y}_i(1-\overline{x}_i), \sum_{i=1}^{n} \overline{y}_i(1-\overline{x}_i)$$

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 value 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. **Monotone Convergence**

If we set

$$u_0 = \text{Min}_{H} \text{Max}_{i} \sum_{j} a_{ij}x_j$$

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

Similarly, if we have

$$a_1 \leq a_{ij} \leq a_2$$

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


