Cambridge University Press
978-0-521-57167-8 - Nonnegative Matrices and Applications
R. B. Bapat and T. E. S. Raghavan
Excerpt
More information

1

Perron-Frobenius theory and matrix games

The Perron-Frobenius Theorem is central to the theory of nonnegative matrices. An irreducible nonnegative matrix can be viewed as the payoff matrix of a zero-sum, two-person game with positive value. A matrix game is said to be completely mixed if no row or column is dispensable for optimal play. In this chapter we first exploit the properties of completely mixed matrix games to prove the Perron-Frobenius Theorem. The next few sections deal with certain related topics such as M-matrices, the structure of reducible nonnegative matrices, primitive matrices, and polyhedral sets with a least element. We then describe the basic aspects of finite Markov chains. In the final section we prove the Perron-Frobenius Theorem for operators that leave the Lorentz cone invariant.

1.1. Irreducible nonnegative matrices

We work with real matrices throughout, unless stated otherwise. Let $A = (a_{ij})$ be an $m \times n$ matrix. We say that the matrix A is *nonnegative* and write $A \geq 0$, if $a_{ij} \geq 0$ for all i, j. If $a_{ij} > 0$ for all i, j, then the matrix A is called *positive* and we write $A > 0$. For matrices A, B, we say $A \geq B$ if $A - B \geq 0$. Similar definitions and notation apply for vectors. The Euclidean n-space is denoted by R^n. The identity matrix of the appropriate order is denoted by I. The transpose of the matrix A is denoted by A^T.

An $n \times n$ matrix P is called a *permutation matrix* of order n if P can be obtained from the $n \times n$ identity matrix by permuting its rows and columns. Suppose we permute the rows of a matrix A to get the new matrix B. We can write the matrix B as $B = PA$, where P is the permutation matrix obtained by permuting the rows of the identity matrix, in the same way as B is obtained from A. Similarly, any column permutation of A corresponds to a matrix $C = AP$, where P is a permutation matrix.

A matrix A of order $n \times n$ is said to be *reducible*, either if A is the 1×1 zero matrix or if $n \geq 2$ and there exists a permutation matrix P such that

$$PAP^T = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix},$$

where B and D are square matrices and 0 is a zero matrix. The matrix A is *irreducible* if it is not reducible.

The following lemma is useful in identifying reducible matrices.

Lemma 1.1.1. *Let A be an $n \times n$ matrix with $n \geq 2$. Let $a_{ij} = 0$ for $i \in S$, $j \notin S$ for some nonempty, proper subset S of $\{1, 2, \ldots, n\}$. Then A is reducible.*

Proof. Let $S = \{i_1, i_2, \ldots, i_k\}$, where we assume, without loss of generality, that $i_1 < i_2 < \cdots < i_{k-1} < i_k$. Let $S^c = T$ be the complement of S consisting of the ordered set of elements $j_1 < j_2 < \cdots < j_{n-k}$. Consider the permutation σ of $\{1, 2, \ldots, n\}$ given by

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & k & k+1 & k+2 & \cdots & n \\ i_1 & i_2 & \cdots & i_k & j_1 & j_2 & \cdots & j_{n-k} \end{pmatrix}.$$

Note that σ can be represented by the permutation matrix $P = (p_{ij})$, where $p_{rs} = 1$ if $\sigma(r) = s$. We prove that

$$PAP^T = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix},$$

where B and D are square matrices and 0 is a $k \times (n-k)$ zero matrix. Consider row α and column β, where $1 \leq \alpha \leq k$ and $k+1 \leq \beta \leq n$. Now

$$(PAP^T)_{\alpha\beta} = \sum_i \sum_j p_{\alpha i} a_{ij} p_{\beta j}.$$

It is enough to show that each term in the summation is zero. Suppose $p_{\alpha i} = p_{\beta j} = 1$. Thus $\sigma(\alpha) = i$ and $\sigma(\beta) = j$. Since $1 \leq \alpha \leq k$, then $i \in \{i_1, i_2, \ldots, i_k\}$; similarly, since $k+1 \leq \beta \leq n$, we have $j \in \{j_1, j_2, \ldots, j_{n-k}\}$. By assumption, for such a pair i, j, we have $a_{ij} = 0$. That completes the proof. ■

Some important characterizations of irreducible matrices are given in the next result.

Theorem 1.1.2. *Let $A \geq 0$ be an $n \times n$ matrix. Then the following conditions are equivalent:*

(1) A is irreducible.
(2) $(I + A)^{n-1} > 0$.
(3) For any pair (i, j), $1 \le i, j \le n$, *there is a positive integer* $t = t(i, j) \le n$ *such that* $(A^t)_{ij} = a_{ij}^{(t)} > 0$.

Proof. (1) ⇒ (2): Let $y \ge 0$, $y \ne 0$ be an arbitrary vector in R^n. If a coordinate of y is positive, the same coordinate is positive in $y + Ay = (I + A)y$ as well. We claim that $(I + A)y$ has fewer zero coordinates than y as long as y has a zero coordinate. If the claim is not true, then $y_j = 0 \Rightarrow y_j + (Ay)_j = 0$ for any coordinate j. Let $J = \{j : y_j > 0\}$. For any $j \notin J, r \in J$, we have $(Ay)_j = \sum_k a_{jk} y_k = 0$ and $y_r > 0$. Thus, $a_{jr} = 0$. It follows by Lemma 1.1.1 that A is reducible, which is a contradiction and the claim is proved. Thus $(I + A)y$ has at most $n - 2$ zero coordinates. Continuing in this manner we conclude that $(I + A)^{n-1} y > 0$. We now set y as a column of the identity matrix, so the corresponding column of $(I + A)^{n-1}$ must be positive. Thus (2) holds.

(2) ⇒ (3): Since $(I + A)^{n-1} > 0$, $A \ge 0$, then $A \ne 0$ and we have

$$A(I + A)^{n-1} = \sum_{k=1}^{n} \binom{n-1}{k-1} A^k > 0.$$

Thus for any i, j, at least one of the matrices $A, A^2, \ldots, A^n$ has its (i, j)-th coordinate positive.

(3) ⇒ (1): Suppose A is reducible. Then for some permutation matrix P,

$$PAP^T = \begin{bmatrix} B_1 & 0 \\ C_1 & D_1 \end{bmatrix},$$

where B_1 and D_1 are square matrices. Furthermore, $PAP^T PAP^T = PA^2P^T$, whence for some square matrices B_2, C_2 we have

$$PA^2P^T = \begin{bmatrix} B_2 & 0 \\ C_2 & D_2 \end{bmatrix}.$$

More generally, for some matrix C_t and square matrices B_t and D_t,

$$PA^tP^T = \begin{bmatrix} B_t & 0 \\ C_t & D_t \end{bmatrix}.$$

Thus $(PA^tP^T)_{\alpha\beta} = 0$ for $t = 1, 2, \ldots$ and for any α, β corresponding to an entry of the zero submatrix in PAP^T.
Now

$$0 = (PA^tP^T)_{\alpha\beta} = \sum_k \sum_l p_{\alpha k} a_{kl}^{(t)} p_{\beta l} \quad \text{for } t = 1, \ldots, n.$$

Choose k, l so that $p_{\alpha k} = p_{\beta l} = 1$. Then $a_{kl}^{(t)} = 0$ for all t, contradicting the hypothesis and thereby completing the proof. ■

It follows from Theorem 1.1.2 that A^T is irreducible whenever A is irreducible.

We will make use of elementary concepts from graph theory without defining them explicitly. We refer the reader to Lovász (1979) and Bondy and Murty (1976) for these concepts.

If A is a nonnegative $n \times n$ matrix then we may associate a directed graph with A as follows: The graph has n vertices, which we denote by $1, 2, \ldots, n$. There is an edge from vertex i to vertex j if and only if a_{ij} is positive. Denote this graph by $G(A)$. A directed graph is said to be *strongly connected* if there is a path from any vertex to any other vertex. (In contrast with the standard terminology, we will not make any distinction between a walk and a path; thus a path may have a vertex appearing more than once.) Observe that $a_{ij}^{(t)} > 0$ if and only if there is a path of length t from vertex i to vertex j in $G(A)$. It follows from the equivalence of (1) and (3) in Theorem 1.1.2 that A is irreducible if and only if $G(A)$ is strongly connected.

1.2. Perron's Theorem on positive matrices

A set $S \subset R^n$ is said to be *convex* if for any $x, y \in S$ and for any $0 \leq \lambda \leq 1$, $\lambda x + (1 - \lambda)y \in S$. The empty set and any set with exactly one element are convex. Geometrically, a set S is convex if for any pair of points in S, the line segment joining the pair of points completely lies in S.

Here are some examples of convex sets:

(i) $S_1 = \{(x_1, x_2, \ldots, x_n) : \sum_j x_j^2 \leq 1\}$.
(ii) $S_2 = \{(x_1, x_2, \ldots, x_n) : \sum_j a_{ij}x_j \geq b_i, i = 1, 2, \ldots, m\}$, where a_{ij}, b_i are given real numbers.
(iii) $S_3 = \{(x_1, x_2, \ldots, x_n) : \sum_j x_j = 1; x_j \geq 0 \text{ for all } j\}$.

We may think of a nonnegative $n \times n$ matrix A as a linear transformation with respect to a fixed basis. Notice that if $x \geq 0$ in R^n, then $Ax \geq 0$. Thus the set of all nonnegative vectors in R^n is mapped into itself by the matrix A. A set $K \subset R^n$ is called a *cone* if $x, y \in K \Rightarrow x + y \in K$ and $x \in K \Rightarrow \lambda x \in K$ for any $\lambda \geq 0$. A set K in R^n is called a *convex cone* if it is a convex set and if for any $x \in K$ and $\lambda \geq 0$, $\lambda x \in K$. The set of all nonnegative vectors is a convex cone, and a nonnegative matrix leaves this cone invariant.

The Perron-Frobenius Theorem was originally proved by Perron for positive matrices. In this section we prove the main aspects of Perron's Theorem.

The technique is elementary, except for the fact that we will use Brouwer's Fixed Point Theorem, which we now state without proof. For a proof using combinatorial ideas, see Bondy and Murty (1976).

Theorem 1.2.1 (Brouwer's Fixed Point Theorem). *Let S be a nonempty, closed, bounded, convex set in R^n. Let $f : S \to S$ be a continuous map. Then there exists an $x \in S$ such that $f(x) = x$.*

We now recall some elementary facts about multiplicities of eigenvalues, which will be needed in subsequent sections. For any square matrix of order n with real or complex entries, the characteristic polynomial of the matrix can be written as $p(\lambda) = c\prod_{i=1}^{k}(\lambda - \lambda_i)^{m_i}$, where c is a constant and $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct. Here m_i is called the *algebraic multiplicity* of the characteristic root $\lambda_i, i = 1, 2, \ldots, k$. We call a root, say λ_1, a *simple root,* if $m_1 = 1$. If $m_1 = 1$, then $\frac{d}{d\lambda}p(\lambda)|_{\lambda=\lambda_1} \neq 0$. If $m_1 > 1$, then $\frac{d}{d\lambda}p(\lambda)|_{\lambda=\lambda_1} = 0$. Thus, a characteristic root λ_1 is simple if and only if the derivative $p'(\lambda_1)$ is nonzero. Another notion of multiplicity is that of the geometric multiplicity of the characteristic root. For any characteristic root λ, let $S_\lambda = \{u : Au = \lambda u\}$. Here S_λ is a vector space in its own right and A, viewed as a linear transformation, leaves the subspace invariant. The dimension of S_λ is called the *geometric multiplicity* of the characteristic root λ. In general the two multiplicities need not be the same. For example, the matrix

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

has $p(\lambda) = |A - \lambda I| = -\lambda^3$, where $|\cdot|$ denotes determinant. Thus, 0 is a root of A with algebraic multiplicity 3. However, the only characteristic vector for the characteristic root 0 is $(0, 0, 1)^T$, up to a scalar multiple. Hence the geometric multiplicity of λ is 1. In general the algebraic multiplicity is not less than the geometric multiplicity. The following argument can be made precise to show this claim. Fix an eigenvalue λ_0, and think of A as a linear transformation on the vector space S_{λ_0} into itself. This restriction of A to S_{λ_0} can be thought of as another linear transformation A_0 with $|A_0 - \lambda I| = (\lambda - \lambda_0)^m$, where m is the dimension of S_{λ_0}. Clearly, we have $m \leq m_{\lambda_0}$ where m_{λ_0} is the algebraic multiplicity of λ_0 as an eigenvalue of A.

Theorem 1.2.2 (Perron's Theorem). *Let $A > 0$ be an $n \times n$ matrix. Then*

(i) $Ay = \lambda_0 y$ for some $\lambda_0 > 0$, $y > 0$.
(ii) The eigenvalue λ_0 is maximal in modulus among all the eigenvalues of A. That is, for any eigenvalue μ of A, $|\mu| \leq \lambda_0$.

(iii) The eigenvalue λ_0 is geometrically simple. That is, any two eigenvectors corresponding to λ_0 are linearly dependent.

(iv) Any positive eigenvector of A (corresponding to any eigenvalue) is a scalar multiple of y.

Proof. Let

$$S = \left\{ (x_1, x_2, \dots, x_n)^T : \sum_i x_i = 1 \text{ and } x_i \geq 0 \text{ for all } i \right\}.$$

Define the map $f : S \to S$ as follows:

$$f(x) = \left\{ \sum_i (Ax)_i \right\}^{-1} Ax.$$

Here $(Ax)_i$ is the i-th coordinate of Ax. If $x \in S$ then, because $A > 0$, the vector Ax is nonzero and the map f is well defined. It is easily checked that f is continuous and maps S into S. By Brouwer's Fixed Point Theorem, $f(y) = y$ for some $y \in S$. Thus

$$\left\{ \sum_i (Ay)_i \right\}^{-1} Ay = y.$$

If we set $\sum_i (Ay)_i = \lambda_0$, then $\lambda_0 > 0$ and $Ay = \lambda_0 y$. Since $A > 0$, it follows that $y > 0$. Hence (i) is proved.

If we apply (i) to A^T, then we conclude that $A^T z = \lambda_1 z$ for some $\lambda_1 > 0$, $z > 0$. Now

$$\lambda_1 y^T z = y^T A^T z = \lambda_0 y^T z,$$

and since $y^T z > 0$, we have $\lambda_0 = \lambda_1$. Thus $A^T z = \lambda_0 z$. This fact will be used in the rest of the proof. Let $Au = \mu u$ for some real or complex eigenvalue μ. Let u^+ be defined by $u^+ = (|u_1|, |u_2|, \dots, |u_n|)^T$, where $u = (u_1, u_2, \dots, u_n)^T$. Without loss of generality, let u^+ be a probability vector. We have

$$\sum_j a_{ij} |u_j| \geq \left| \sum_j a_{ij} u_j \right| = |\mu u_i| = |\mu| |u_i|.$$

Thus $Au^+ \geq |\mu| u^+$. Premultiply this last inequality by z^T to conclude that $|\mu| \leq \lambda_0$. This completes the proof of (ii).

We now prove (iii). Suppose $Av = \lambda_0 v$ for some real, nonzero vector v. We must show that v is a scalar multiple of y. If v and y are linearly independent, then there exists a real number α such that $y - \alpha v$ is a nonnegative, nonzero vector with at least one zero coordinate. Since

$$A(y - \alpha v) = \lambda_0 (y - \alpha v),$$

$y - \alpha v$ is an eigenvector of A. However, since $A > 0$, any nonnegative eigenvector of A must in fact be positive and we get a contradiction. Thus v is a scalar multiple of y. By considering the real and the imaginary parts separately, we can show that any complex eigenvector of A corresponding to λ_0 is a scalar multiple of y.

To prove (*iv*), suppose $Au = \mu u$ for $u > 0$. We have $\mu z^T u = z^T Au = \lambda_0 z^T u$. Since $z^T u > 0$, we have $\mu = \lambda_0$. The result now follows by (*iii*). ∎

1.3. Completely mixed games

The theory of games is concerned with problems of conflict. The simplest form of such games are the so-called *matrix games* played as follows: Players I and II secretly choose a column j and a row i, respectively, of a matrix $A = (a_{ij})$, $1 \le i \le m$, $1 \le j \le n$. Their choices are revealed to a referee who then calls Player II to pay Player I the amount a_{ij}. If $a_{ij} < 0$, it is an income to Player II from Player I. (We have slightly deviated from the standard conventions of matrix games in which the rows are usually chosen by Player I.)

Example 1.1. Player I shows 1, 2, or 3 fingers, and simultaneously Player II shows 1 or 2 fingers. The payoff to I from II is the total number of fingers shown by both. The matrix of this game is

$$\begin{array}{lcc} & \text{I's actions} = & 1f \quad 2f \quad 3f \\ \text{II's actions} = & \begin{array}{c} 1f \\ 2f \end{array} & \begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}. \end{array}$$

Obviously, if they play the game several times, Player I will show 3 fingers and Player II will show 1 finger every time.

Example 1.2. The game is the same as above but the payoff to Player I is 1 if the total number of fingers shown is odd and -1 if the number is even. The payoff matrix is

$$A = \begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}.$$

In repeated play it is not good for Player II to always show 1 finger, for in that case, Player I will show 2 fingers every time and collect 1 unit from II. Similarly, it is not desirable for I to always show the same number of fingers. It is clear, however, that for Player I showing 1 finger is the same as showing 3 fingers against any choice of the opponent and Player I might as well not bother about showing 3 fingers. Suppose Player I shows 1 or 2 fingers based on the outcome of the toss of a fair coin. If Player II chooses 1 finger, then I loses 1 unit half the time and gains 1 unit half the time. Irrespective of the choice of Player II, the average gain for I is zero. Similarly, Player II's loss on the average is zero if he also uses a fair coin to show 1 or 2 fingers. Thus, tossing a coin to select 1 or 2 fingers is a good strategy for both Players.

Example 1.3. The game is the same as above. We have the following modified payoff to Player I: Player I receives from Player II the total of the number of fingers shown if the total is odd; otherwise he pays Player II the total of the number of fingers shown. The payoff matrix to Player I is

$$A = \begin{bmatrix} -2 & 3 & -4 \\ 3 & -4 & 5 \end{bmatrix}.$$

Suppose, as in the previous example, Player II tosses a fair coin and decides to show 1 or 2 fingers depending on the outcome. By showing 1 finger all the time, Player I can gain on the average at most half a unit. Player II can do better. Suppose he shows 1 finger with chance $7/12$ and 2 fingers with chance $5/12$. The average gain to Player I would be $1/12$ if he shows 1 or 2 fingers and $-3/12$ if he shows 3 fingers. Thus, Player II loses no more than $1/12$ on the average—no matter what Player I does. This is certainly better than tossing a coin. However, it is not clear whether Player II can do better than this. From the point of view of Player I, the strategy that selects 1 finger with chance $7/12$ and 2 fingers with chance $5/12$ guarantees, on the average, $1/12$ to player I, no matter how many fingers Player II shows. Therefore, we say that the mixed strategy $(7/12, 5/12)$ is *optimal* for Player II and the mixed strategy $(7/12, 5/12, 0)$ is optimal for Player I. Further, we say that the *value* of the game is $1/12$.

Example 1.4. The following dialogue takes place between a teacher and a student:

Student: Professor, will you give us a take home final for the Game Theory course?
Teacher: I don't believe in them.
Student: I am nervous in the regular exam. I get only *C*s in those exams.
However, I can do much better with a take home.
Teacher: I know.
Student: I heard that you always recycle old questions!
Teacher: I would not contradict that. In fact, for the final exam, I have photocopied one among the ten questions that are on my desk.
Student: Can I pick up all of them to try at home?
Teacher: No, not really. I can let you pick up only one of them.

Student: I appreciate your help. I am curious. What do you expect from me if I take one of these questions home?

Teacher: It depends. If you pick up question i, and if it also happens to be the final exam question, you could probably score $q_i > 0$; otherwise you can hope for your usual score c.

How shall we compute the expected score for this pessimistic student? Without loss of generality, let $q_1 \geq q_2 \geq \cdots \geq q_{10} > c > 0$. The student keen on maximizing his average score and ensuring a score c on any exam, can use the following payoff matrix for the game between him and the teacher:

$$\begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ 10 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & \cdots & 10 \end{array} \\ \begin{pmatrix} q_1 - c & 0 & \cdots & 0 \\ 0 & q_2 - c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & q_{10} - c \end{pmatrix} \end{array}.$$

If x_i is the chance for the i-th question to be selected by the student, then the expected score for the student is at least $\min_i (q_i - c)x_i$. Temporarily pretending $(q_i - c)x_i = v$ for all i, we get $x_i = v/(q_i - c)$. Summing x_i, we get $v = 1/\sum_i 1/(q_i - c)$. The same strategy is seen to also work for the sadistic teacher.

The existence of such optimal strategies and value is not just a coincidence in these examples. We have the following celebrated theorem of von Neumann. For a proof, see, for example, Parthasarathy and Raghavan (1971).

Theorem 1.3.1 (Minimax Theorem). *Let $A = (a_{ij})$ be an $m \times n$ payoff matrix. Then there exists a unique constant v (called the value) and mixed strategies $x = (x_1, x_2, \ldots, x_m)$ for Player II and $y = (y_1, y_2, \ldots, y_n)^T$ for Player I such that*

$$\sum_j a_{ij} y_j \geq v, \quad i = 1, 2, \ldots, m, \qquad \sum_i a_{ij} x_i \leq v, \quad j = 1, 2, \ldots, n. \tag{1.3.1}$$

(Here $x_i \geq 0$ for all i and $\sum_i x_i = 1$, $y_j \geq 0$ for all j and $\sum_j y_j = 1$.) The strategy x is called an optimal strategy for Player II. The strategy y is called an optimal strategy for Player I.

A player not knowing the exact choice of his opponent must be ready to take care of himself under every possible course of action by the opponent. The first of these inequalities expresses the fact that when Player I chooses column j in the payoff matrix with chance y_j, then for any choice i by Player II, the expected income $\sum_j a_{ij} y_j$ to Player I is at least v. Similarly, Player II, though not knowing the actual choice of Player I, can safeguard his expected losses by choosing action i (row i) with chance x_i. The expected loss of Player II when Player I chooses column j is $\sum_i a_{ij} x_i$. The second inequality says that it is at most v.

A mixed strategy x for Player II is called *completely mixed* if $x > 0$, that is, if all the rows of the payoff matrix are chosen with positive probability. A completely mixed strategy for Player I is defined similarly. A matrix game is called completely mixed if every optimal mixed strategy x for Player II and y for Player I are completely mixed.

Before we state the next result we recall the following facts on systems of equations. Let A be a real $m \times n$ matrix. The *null space* $\{x : Ax = 0\}$ is a vector space with dimension $n - r$, where r is the rank of A. The dimension of the range space $\{Aw : w \in R^n\}$ is r. We can choose a basis for the null space with $n - r$ elements. Thus, we have $Aw^{(1)} = Aw^{(2)} = \cdots = Aw^{(n-r)} = 0$, where $w^{(1)}, w^{(2)}, \ldots, w^{(n-r)}$ are linearly independent.

If u, v are vectors in R^n, then we denote their inner product by $\langle u, v\rangle$. We denote by $\mathbf{1}$ the column vector of appropriate size with each entry equal to 1.

Theorem 1.3.2. *Let v be the value of the matrix game A. Let some optimal strategy of Player II be completely mixed. Then, for any optimal strategy y of Player I, $Ay = v\mathbf{1}$.*

Proof. If x and y are optimal strategies for Player II and Player I respectively, then it follows from (1.3.1) that $x^T Ay = v$. Now suppose x is completely mixed. We have

$$0 = \langle x, Ay\rangle - v = \langle x, Ay - v\mathbf{1}\rangle. \tag{1.3.2}$$

Let $u_i = (Ay - v\mathbf{1})_i = \sum_j a_{ij} y_j - v$. Since y is an optimal strategy for Player I, $u_i = \sum_j a_{ij} y_j - v \geq 0$. From (1.3.2) we have $\sum_i x_i u_i = 0$. Since $x_i > 0$ for all i, $u_i = 0$. That completes the proof. ■

Theorem 1.3.3. *Let the value of the $m \times n$ matrix game $A = (a_{ij})$ be zero and suppose that every optimal strategy for Player II is completely mixed. Then $m - 1 \leq \operatorname{rank} A \leq n - 1$. If $\operatorname{rank} A = m - 1$, then the optimal strategy for Player II is unique.*

Proof. By Theorem 1.3.2 we have $Ay = 0$ for any optimal strategy y of Player I. Since $y \neq 0$, rank $A \leq n - 1$. In case rank $A \leq m - 2$ there exist at least two linearly independent solutions to $A^T u = 0$. We can assume that one of them