ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS 64

NONNEGATIVE MATRICES AND APPLICATIONS

R. B. BAPAT

T. E. S. RAGHAVAN

CAMBRIDGE

more information – www.cambridge.org/9780521571678

This page intentionally left blank

This book presents an integrated treatment of the theory of nonnegative matrices, emphasizing connections with the themes of game theory, combinatorics, inequalities, optimization, and mathematical economics. Some related classes of positive matrices such as positive semidefinite matrices, M-matrices, P-matrices, and distance matrices are also discussed, but the main emphasis is on entrywise nonnegative matrices.

The book begins with the basics of the subject, such as the Perron-Frobenius Theorem. Only a minimal background in linear algebra is assumed, although familiarity with linear programming and statistics will be helpful in following some sections. Each of the later chapters is devoted to an area of applications, including doubly stochastic matrices (price fixing, scheduling, and the fair division problem), combinatorial matroids, and economics. These applications have been carefully chosen both for their elegant mathematical content and for their accessibility. The treatment is rigorous and almost all results are proved completely.

About half of the material in the book presents standard topics in a novel fashion, the remaining portion reports many new results in matrix theory for the first time in a book form.

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

EDITED BY G.-C. ROTA

Editorial Board

R. S. Doran, M. Ismail, T.-Y. Lam, E. Lutwak, R. Spigler

Volume 64

Nonnegative Matrices and Applications

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

4 W. Miller, Jr. *Symmetry and separation of variables*
6 H. Minc *Permanents*
11 W. B. Jones and W. J. Thron *Continued fractions*
12 N. F. G. Martin and J. W. England *Mathematical theory of entropy*
18 H. O. Fattorini *The Cauchy problem*
19 G. G. Lorentz, K. Jetter, and S. D. Riemenschneider *Birkhoff interpolation*
21 W. T. Tutte *Graph theory*
22 J. R. Bastida *Field extensions and Galois theory*
23 J. R. Cannon *The one-dimensional heat equation*
25 A. Salomaa *Computation and automata*
26 N. White (ed.) *Theory of matroids*
27 N. H. Bingham, C. M. Goldie, and J. L. Teugels *Regular variation*
28 P. P. Petrushev and V. P. Popov *Rational approximation of real functions*
29 N. White (ed.) *Combinatorial geometries*
30 M. Pohst and H. Zassenhaus *Algorithmic algebraic number theory*
31 J. Aczel and J. Dhombres *Functional equations in several variables*
32 M. Kuczma, B. Chozewski, and R. Ger *Iterative functional equations*
33 R. V. Ambartzumian *Factorization calculus and geometric probability*
34 G. Gripenberg, S.-O. Londen, and O. Staffans *Volterra integral and functional equations*
35 G. Gasper and M. Rahman *Basic hypergeometric series*
36 E. Torgersen *Comparison of statistical experiments*
37 A. Neumaier *Interval methods for systems of equations*
38 N. Korneichuk *Exact constants in approximation theory*
39 R. A. Brualdi and H. J. Ryser *Combinatorial matrix theory*
40 N. White (ed.) *Matroid applications*
41 S. Sakai *Operator algebras in dynamical systems*
42 W. Hodges *Basic model theory*
43 H. Stahl and V. Totik *General orthogonal polynomials*
44 R. Schneider *Convex bodies*
45 G. Da Prato and J. Zabczyk *Stochastic equations in infinite dimensions*
46 A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. Ziegler *Oriented matroids*
47 G. A. Edgar and L. Sucheston *Stopping times and directed processes*
48 C. Sims *Computation with finitely presented groups*
49 T. Palmer *Banach algebras and the general theory of *-algebras*
50 F. Borceux *Handbook of categorical algebra I*
51 F. Borceux *Handbook of categorical algebra II*
52 F. Borceux *Handbook of categorical algebra III*
54 A. Katok and B. Hasselblatt *Introduction to the modern theory of dynamical systems*
55 V. N. Sachkov *Probabilistic methods of discrete mathematics*
56 V. N. Sachkov *Combinatorial methods of discrete mathematics*
57 P. M. Cohn *Skew fields*
58 Richard Gardner *Geometric tomography*
59 George A. Baker, Jr. and P. Graves-Morris *Padé approximants*
60 J. Krajíček *Bounded arithmetic, propositional logic, and complexity theory*
61 H. Groemer *Geometric applications of Fourier series and spherical harmonics*
62 H. O. Fattorini *Infinite dimensional optimization and control theory*
63 A. C. Thompson *Minkowski geometry*

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

Nonnegative Matrices and Applications

R. B. BAPAT
Indian Statistical Institute

T. E. S. RAGHAVAN
University of Illinois at Chicago

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521571678

First published 1997

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data

R. B. Bapat.
Nonnegative matrices and applications / R. B. Bapat, T. E. S. Raghavan.
p. cm. - - (Encyclopedia of mathematics and its applications; v. 64)
ISBN 0-521-57167-7 hardback
1. Non-negative matrices. I. Raghavan, T. E. S. II. Title.
III. Series.
QA188.B355 1996
512.9′434--dc20 96-12844
CIP

ISBN 978-0-521-57167-8 hardback

Transferred to digital printing 2007

To the memory of my father,

Bhalachandra S. Bapat

- R. B. Bapat

To the memory of my father,

Eachambadi S. Narasimhachari

- T. E. S. Raghavan

CONTENTS

PREFACE

This book is aimed at first year graduate students as well as research workers with a background in linear algebra. The theory of nonnegative matrices is unfolded in the book using tools from optimization, inequalities and combinatorics. The topics and applications are carefully chosen to convey the excitement and variety that nonnegative matrices have to offer. Some of the applications also illustrate the depth and the mathematical elegance of the theory of nonnegative matrices. The treatment is rigorous and almost all the results are completely proved. While about half of the material in the book presents many topics in a novel fashion, the remaining portion reports many new results in matrix theory for the first time in a book form. Although the only prerequisite is a first course in linear algebra and advanced calculus, familiarity with linear programming and statistics will be helpful in appreciating some sections.

To give some examples, the Perron-Frobenius Theorem and many of its consequences are derived using the theory of matrix games where all rows and columns are essential for optimal play. The chapter on conditionally positive definite matrices and distance matrices has several new results appearing for the first time in a book. A transparent proof of the Alexandroff inequality for mixed discriminants is presented and a characterization of graphs giving rise to a finite Coxeter group is given in the chapter on combinatorial theory. The importance of P-matrices and M-matrices to several areas besides linear economic models is stressed and many of these results are seen via Game theory.

The application topics include, among other things, areas like game theory, Markov chains, probabilistic algorithms, numerical analysis, discrete distributions, categorical data, group theory, matrix scaling and economics. The chapter on doubly stochastic matrices contains applications to the pricing of houses in markets with known expectations, the problem of arriving at a fair division that respects peoples' individual preferences, scheduling jobs in a preemptive

environment, and assigning people to jobs to minimize total job completion time. As described in the chapter on matrix scalings, while scaling is a powerful tool to speed up convergence, it is also a useful tool to estimate cell entries of an unknown matrix of the future parametric values from the known entries of the past. Such a procedure could be used to estimate the size of a population in a city or the growth of tumor after an operation and to reconstruct an image based on partial information.

The chapter on topics in economics describes the economic implications of properties of many special classes of matrices. The Perron-Frobenius Theorem can be used to explain both the Leontief and the Sraffa system. In the Leontief system, workers slave with fixed consumption and fixed wage to achieve targeted social output for the future generations. In the Sraffa system, all are entrepreneurs of identical skill who get equal rate of return. In international trade, the famous Hecksher-Ohlin Theorem shows that free trade could very well be a substitute for immigration. Another application for such matrices occurs in the price stability of an economy where any two goods are gross substitutes. In such economies, consumers switch from one brand to another if there is a steep price increase for a particular brand. An increase in price for a brand at a price equilibrium triggers a price increase for other substitute brands due to increased demand. However a dynamic price stability for the brand can be attained, allowing for varying speeds of price adjustments for other brands that are gross substitutes. It may be noted that these topics and results have been cited as major contributions to economic theory by P. Samuelson, Sir John Hicks and B. Ohlin (all Noble prize laureates) by the Noble prize committee.

This book was being written over a period of several years and it has benefitted from the comments, suggestions, and works of a large number of people. We only mention here some names that immediately come to mind. At the outset we wish to acknowledge the influence of the magnificent contributions of John von Neumann and Issai Schur on our approach to the subject at many places. We would also like to mention that the short note by Blackwell on minimax and irreducible matrices and Kaplansky's notion of completely mixed games have given the necessary game theoretic armory for many matrix problems.

We gratefully acknowledge comments and corrections, pertaining to various portions of the manuscript, due to Adi Ben-Israel, Charles Broyden, Gregory M. Constantine, S. Chandrasekaran, John Copas, M. V. Menon, S. R. Mohan, Dale Olesky, S. Panchapakesan, Ashok Ramu, Arunava Sen, Debapriya Sengupta, R. Sridhar, Pauline van den Driessche, and James Weber.

Evangelista Fe, Tamas Solymosi, Murali Srinivasan, Zamir Syed, and Julin Wu took a course based on the first three chapters from Raghavan and made helpful remarks. In particular, Fe made several corrections in spelling, grammer

etc., Solymosi suggested improvements on the section on cooperative games and assignment problems and corrected some mathematical typos, while Murali carefully went through some topics on doubly stochastic matrices.

We owe a lot to S. Sankaran who made extensive corrections, particularly in the usage of syntax and grammar.

Dipankar Dasgupta convinced Raghavan of the Sraffa system as a much refined application of the Perron-Frobenius Theorem to Economics, whereas V. K. Chetty spent hours explaining to him the contributions and the depth of the works of Ricardo and Sraffa's formulation of the theory of production of commodities by means of commodities as a clear solution to the problem that Ricardo attempted to solve without success.

We warmly acknowledge the support given by our family members, Ragini, Sudeep, Usha, Deepa, Sampath, Santanu, Tara, and Manu, during the course of writing this book. The theory of Nonnegative matrices is indeed a fascinating and rewarding area. The Perron-Frobenius Theorem, the central result of the theory, was formulated at the beginning of the twentieth century and many significant developments have taken place during the past ninety years. In this book we have tried to touch upon some of these developments and the choice of topics clearly reflects our personal interests. We would consider our efforts amply rewarded even if a single curious and youthful mind is drawn to the subject and happens to share the same fascination after reading this book.

RBB
TESR

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) $\Rightarrow$ (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) $\Rightarrow$ (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) $\Rightarrow$ (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

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

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 $u^+ = (|u_1|, |u_2|, \ldots, |u_n|)^T$, where $u = (u_1, u_2, \ldots, u_n)^T$. 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 $A = (a_{ij}), 1 \leq i \leq m, 1 \leq j \leq n$. 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

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

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 Cs 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

is independent of an optimal strategy x of Player II. Let π be such a solution. Without loss of generality, either $\sum_i \pi_i = 0$ or $\sum_i \pi_i = 1$.

Case (i): $\sum \pi_i = 0$.
Since $\pi \neq 0$ and $\sum \pi_i = 0$, then $\pi_i > 0$ for some i. Let $\frac{1}{\theta} = \max_i(\frac{\pi_i}{x_i}) > 0$. The vector $x - \theta\pi$ (≥ 0) has at least one of its coordinates equal to zero. (It is that coordinate i for which the above maximum is attained.) Further, since $\sum \pi_i = 0$, $x - \theta\pi$ is a probability vector. Also, we have $A^T(x - \theta\pi) = A^T x \leq 0$. This shows that the vector $x - \theta\pi$ is optimal for Player II, but is not completely mixed. This is a contradiction to our assumption. Thus rank $A \geq m - 1$.

Case (ii): $\sum \pi_i = 1$.
Let $x > 0$ be any optimal strategy for Player II. Consider the vector $z = (1 + \theta)x - \theta\pi$. For $\theta > 0$ and sufficiently small, the vector z is positive. If θ is chosen such that $\frac{1+\theta}{\theta} = \max_i \frac{\pi_i}{x_i}$, then z is a nonnegative vector with at least one zero coordinate. Since $\sum \pi_i = 1$, z is a probability vector. Furthermore, $A^T z = (1 + \theta)A^T x - \theta A^T \pi = (1 + \theta)A^T x \leq 0$. Thus, z is an optimal strategy for Player II that is not completely mixed. This contradicts our assumption and therefore rank $A \geq m - 1$.

Lastly, let rank $A = m - 1$. This means that the equation $A^T\pi = 0$ has precisely one solution up to a scalar multiple. Suppose Player II has two distinct optimal strategies. They are necessarily linearly independent, and one of them will be linearly independent of π. We can repeat the above proof verbatim with π and the chosen optimal strategy of Player II, which is independent of π. As proved above, we can again contradict our assumption that the optimal strategies of Player II are completely mixed. Hence, when rank $A = m - 1$, Player II has a unique optimal strategy. ■

We now wish to obtain an important characterization of completely mixed games due to Kaplansky (1945). We first prove some preliminary results.

Theorem 1.3.4. *Let A be an $m \times n$ payoff matrix. If $m > n$, then Player II can optimally skip a row. If $m < n$, then Player I can optimally skip a column.*

Proof. If v is the value of the matrix game $A = (a_{ij})$, then the matrix $B = (a_{ij} - v)$ has value zero and the optimal strategies of the players remain unchanged. Therefore, we may assume that $v = 0$. Let $m > n$. If every optimal strategy of Player II is completely mixed, then by Theorem 1.3.3, $m - 1 \leq n - 1$. Because this contradicts our assumption there should be an optimal strategy that selects a row with zero probability (that is, it skips a row). A similar argument can be given when $m < n$. ■

Theorem 1.3.5. *Let A be an $n \times n$ payoff matrix. If the game is not completely mixed, then both players have optimal strategies that are not completely mixed.*

Proof. We again assume, without loss of generality, that $v = 0$. Contrary to the assertion, let us suppose that every optimal strategy of Player II is completely mixed. By Theorem 1.3.3 the optimal strategy is unique for Player II. since the game is not completely mixed, Player I has an optimal strategy that is not completely mixed. Let y be such a strategy with, say, $y_1 = 0$. By Theorem 1.3.2, $Ay = 0$. Also, because A is singular,

$$\sum_j a_{ij} A_{kj} = 0, \quad i = 1, 2, \ldots, n \text{ for each fixed } k,$$

where A_{kj} denotes the cofactor of a_{kj}. By Theorem 1.3.3, the rank of A is $n-1$. Therefore,

$$(A_{k1}, A_{k2}, \ldots, A_{kn}) = \alpha_k (y_1, y_2, \ldots, y_n)$$

for some α_k, and $A_{k1} = \alpha_k y_1 = 0$, for each k. Thus, $A_{11} = A_{21} = \cdots = A_{n1} = 0$ and, therefore,

$$\begin{bmatrix} a_{12} & a_{13} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$$

has rank $\leq n-2$. Then the system of equations $\sum_i a_{ij} u_i = 0$, $j = 2, \ldots, n$ has at least two linearly independent solutions. The optimal strategy x for Player II comprises one solution; the other solution, call it π, necessarily satisfies $\sum_i a_{i1} \pi_i \neq 0$, for otherwise $\sum_i a_{ij} u_i = 0$, $j = 1, 2, \ldots, n$ will have x and π as two linearly independent solutions, which contradicts rank $A = n-1$. As in Theorem 1.3.3 we can assume $\sum_i \pi_i = 0$ or $\sum_i \pi_i = 1$.

Case (i): $\sum_i a_{i1} \pi_i > 0, \sum_i \pi_i = 0$.
As in the proof of Theorem 1.3.3 we can choose $\lambda > 0$ such that the vector $z = (z_1, z_2, \ldots, z_n)^T$ with $z_i = x_i - \lambda \pi_i$, $i = 1, 2, \ldots, n$ satisfies $z_i \geq 0$ for all i, and $z_{i_0} = 0$ for some i_0. However, $\sum_i a_{ij} z_i = \sum_i a_{ij} x_i - \lambda \sum_i a_{ij} \pi_i = -\lambda \sum_i a_{ij} \pi_i$ for all j. For $j = 1$, $\sum_i a_{i1} z_i = -\lambda \sum_i a_{i1} \pi_i < 0$ and, by assumption, $\sum_i a_{ij} z_i = -\lambda \sum_i a_{ij} \pi_i = 0$ for $j = 2, 3, \ldots, n$. Also, $\sum z_i = 1$. Thus, z is optimal for Player II with $z_{i_0} = 0$. This contradicts our assumption on the optimal strategies of Player II.

Case (ii): $\sum_i a_{i1} \pi_i < 0, \sum_i \pi_i = 0$.
The vector $(z_1, z_2, \ldots, z_n)^T$ with $z_i = x_i + \lambda \pi_i$, $i = 1, 2, \ldots, n$ can be used

to contradict our assumption on the optimal strategies of Player II.

Case (iii): $\sum_i a_{i1}\pi_i > 0, \sum_i \pi_i = 1$.
The vector $z = (z_1, z_2, \ldots, z_n)^T$ with $z_i = (1+\lambda)x_i - \lambda\pi_i$ for suitable $\lambda > 0$ can be used to contradict our assumption on the optimal strategies of Player II as in case (ii).

Case (iv): $\sum_i a_{i1}\pi_i < 0, \sum \pi_i = 1$.
Use $z_i = (1-\lambda)x_i + \lambda\pi_i$, $i = 1, 2, \ldots, n$. The proof is similar to the previous cases. Thus, in all cases there is a strategy for Player II that is not completely mixed, and the proof is complete. ■

Theorem 1.3.6. *An $m \times n$ matrix game with value zero is completely mixed if and only if*

(i) $m = n$ and the rank of the matrix is $n - 1$.
(ii) All cofactors are different from zero and are of the same sign.
(iii) The optimal strategies are unique for the two players.

Proof. (Necessity). By Theorem 1.3.4 and Theorem 1.3.5 we observe that condition (i) is necessary. Furthermore, since both players have unique completely mixed optimal strategies, (iii) also follows. Now $\sum_j a_{ij}A_{kj} = 0$, $i = 1, 2, \ldots, n$ and $(A_{k1}, A_{k2}, \ldots, A_{kn}) = \alpha(y_1, y_2, \ldots, y_n)$ for the unique optimal y of Player I for some α. Since the y_is are positive, the vector $(A_{k1}, A_{k2}, \ldots, A_{kn})$ is either positive or negative or zero. By using the unique completely mixed strategy of Player II we have a similar assertion for $(A_{1j}, A_{2j}, \ldots, A_{mj})$ for any j. Thus all the cofactors A_{ij} are positive or negative or zero. Since the rank of the matrix is $n - 1$, some cofactor A_{ij} is nonzero and the matrix $C = (A_{ij})$ is either positive or negative.
(Sufficiency). Conversely, let the matrix A be square with rank $n - 1$ and with all cofactors A_{ij} of the same sign. Let $\frac{1}{\alpha} = \sum_j A_{1j}$. Observe that $y = \alpha(A_{11}, A_{12}, \ldots, A_{1n})$ is a mixed strategy and that

$$\sum_j a_{ij}y_j = \alpha \sum_j a_{ij}A_{1j} = 0, \quad i = 1, 2, \ldots, n.$$

Thus y is optimal for Player I.

Similarly, for $\frac{1}{\beta} = \sum_i A_{i1}$, the vector $x = \beta(A_{11}, A_{21}, \ldots, A_{n1})^T$ is a mixed strategy and

$$\sum_i a_{ij}x_i = \beta \sum_i a_{ij}A_{ij} = 0, \quad i = 1, 2, \ldots, n.$$

Hence, x and y are optimal strategies for Players II and I, respectively. From Theorem 1.3.2 any optimal strategy u of Player I will satisfy $\sum_j a_{ij}u_j = 0$.

Because the rank of A is $n-1$, $u = y$ and the optimal strategy for Player I is unique. Similarly, the optimal strategy for Player II is unique. Thus, the optimal strategies are completely mixed and, by Theorem 1.3.5, the game is completely mixed. That completes the proof. ∎

A determinantal formula for the value of a completely mixed game is given in the next result.

Theorem 1.3.7. *Let A be a square matrix such that the game A is completely mixed. Then $\sum_{i,j} A_{ij}$ is nonzero and the value v of A is given by*

$$v = \frac{|A|}{\sum_{i,j} A_{ij}}.$$

Proof. If $v = 0$, Theorem 1.3.2 shows that $Ay = 0$ and $|A| = 0$. Furthermore, by Theorem 1.3.6, $\sum_{i,j} A_{ij}$ is nonzero. This proves the theorem for $v = 0$. If, however, $v \neq 0$, then A must be nonsingular, for otherwise there would exist $\pi \neq 0$ with $A\pi = 0$. From Theorem 1.3.2 we get $Ay = v\mathbf{1}$, which is nonzero for the optimal y of Player I. Obviously, π and y are independent. We can use them as in Theorem 1.3.3 to contradict the fact that the game is completely mixed. Thus A cannot be singular. Now, $Ay = v\mathbf{1}$ gives

$$y = A^{-1}Ay = vA^{-1}\mathbf{1} = v\frac{(\text{adj } A)\mathbf{1}}{|A|}.$$

[Here (adj A) is the transpose of the matrix $C = (A_{ij})$.] Thus,

$$\frac{1}{v} = \frac{\sum_i y_i}{v} = \frac{\sum_{i,j} A_{ij}}{|A|}.$$

This completes the proof of the theorem. ∎

Theorem 1.3.8. *Let A be an $n \times n$ payoff matrix that is not completely mixed. Then the value v of the game satisfies the equation*

$$v = \min_i \max_j v_{ij} = \max_j \min_i v_{ij},$$

where v_{ij} is the value of the $(n-1) \times (n-1)$ game obtained by deleting the i-th row and the j-th column in A.

Proof. Let $A_{\cdot j}$ be the payoff matrix obtained from A by deleting column j with value $v_{\cdot j}$. Since the square matrix A is not completely mixed, by Theorem 1.3.5, both players can skip one of their choices (row or column) optimally. Thus, $v = \max(v_{\cdot 1}, v_{\cdot 2}, \ldots, v_{\cdot n})$. (Player I will skip the column in such a

way that the value is not reduced by playing the subgame.) Similarly, if $A_{i.}$ is the matrix obtained from A by deleting row i and with value $v_{i.}$, we have $v = \min(v_{1.}, v_{2.}, \ldots, v_{n.})$. Let v_{ij} be the value of the matrix game $A(i, j)$ obtained from A by deleting row i and column j. Since $A_{.j}$ is a payoff with more rows than columns, by Theorem 1.3.4, Player I can skip a row and play optimally. Therefore, $v_{.j} = \min_i v_{ij}$, which shows that $v = \max_j v_{.j} = \max_j \min_i v_{ij}$. Similarly, we have $v = \min_i v_{i.} = \min_i \max_j v_{ij}$, and the proof is complete. ∎

In a seminal paper Kaplansky (1945) first introduced the notion of completely mixed strategies. The theorems in this section are essentially from this paper.

1.4. The Perron-Frobenius theorem

The results on completely mixed matrix games proved in the previous section will now be used to establish the Perron-Frobenius Theorem on nonnegative matrices. The next result provides the crucial link for such a proof. For each real λ, let $v(\lambda)$ be the value of the matrix game with payoff $A - \lambda I$.

Theorem 1.4.1. *Let A be an $n \times n$ payoff matrix. Then*

(i) $v(\lambda) \to \pm\infty$ *as* $\lambda \to \mp\infty$.
(ii) $v(\lambda)$ *is nonincreasing in* λ.
(iii) $v(\lambda)$ *is Lipschitz continuous.*

Proof. *(i)* Let Player II choose all rows with equal chance $1/n$. Then Player I can expect on the average at most $\max_j \frac{1}{n}\sum_i a_{ij} - (\lambda/n)$ for the payoff $A - \lambda I$. Thus, no matter what Player I does, we have

$$v(\lambda) \le \max_j \frac{1}{n}\sum_i a_{ij} - \frac{\lambda}{n} \to -\infty \quad \text{as } \lambda \to \infty.$$

(ii) Since $A - \lambda I \ge A - \mu I$ when $\lambda \le \mu$, Player I gains as much by playing $A - \lambda I$ as by playing $A - \mu I$. Thus $v(\lambda) \ge v(\mu)$ and $v(\lambda)$ is nonincreasing in λ.
(iii) Let x° and y° be optimal strategies for Player II and Player I in $A - \lambda I$. Let x^*, y^* be optimal strategies for Player II and I in $A - \mu I$. We have for the inner product

$$\langle x^*, (A - \lambda I)y^\circ \rangle \ge v(\lambda) \tag{1.4.1}$$

and

$$\langle (A - \mu I)^T x^*, y^\circ \rangle = \langle x^*, (A - \mu I)y^\circ \rangle \le v(\mu). \tag{1.4.2}$$

Using (1.4.1) and (1.4.2) we get

$$\begin{aligned} v(\lambda) - v(\mu) &\leq \langle x^*, (A - \lambda I)y^\circ \rangle - \langle x^*, (A - \mu I)y^\circ \rangle \\ &\leq (\mu - \lambda)\langle x^*, y^\circ \rangle \\ &\leq |\mu - \lambda|, \end{aligned}$$

since $\langle x^*, y^\circ \rangle \leq \sum_i x_i^* \sum_i y_i^\circ = 1$. Similarly,

$$-|\mu - \lambda| \leq (\lambda - \mu)\langle x^\circ, y^* \rangle \leq v(\lambda) - v(\mu)$$

and therefore

$$|v(\lambda) - v(\mu)| \leq |\lambda - \mu|.$$

Thus v is Lipschitz continuous. ∎

Lemma 1.4.2. *Let $A \geq 0$ be an irreducible matrix of order n. Then there exists $\lambda_0 > 0$ with $v(\lambda_0) = 0$.*

Proof. Suppose $v(0) = \text{value}\,(A) \leq 0$. Thus, for some optimal x of Player II, $A^T x \leq 0$. Since $A \geq 0$ we have $A^T x = 0$. If $x_p > 0$ for some coordinate p of x, then $a_{pj} = 0$ for $j = 1, 2, \ldots, n$. Permuting the first row with the p-th row and the first column with the p-th column we find that the matrix A can be reduced to the form

$$\begin{bmatrix} B & 0 \\ C & D \end{bmatrix},$$

where B, D are square matrices of order 1 and $n - 1$, respectively. This contradicts the irreducibility of A. By Theorem 1.4.1, $v(\lambda) \to -\infty$ as $\lambda \to \infty$ and $v(\lambda)$ is Lipschitz continuous. Since $v(0) > 0$ we have $v(\lambda_0) = 0$ for some $\lambda_0 > 0$. ∎

Lemma 1.4.3. *Let $A \geq 0$ be irreducible. Let λ_0 be defined as in Lemma 1.4.2. Then the matrix game $A - \lambda_0 I$ is completely mixed.*

Proof. By Theorem 1.3.5 it is enough to prove that every optimal strategy of Player II in the game $A - \lambda_0 I$ is completely mixed. Suppose to the contrary that x is an optimal strategy of Player II with $x_p = 0$ (i.e., row p is skipped optimally by Player II). We have $(A - \lambda_0 I)^T x \leq 0$. This gives $A^T x \leq \lambda_0 x$. Since $A \geq 0$, $(A^T)^k x \leq \lambda_0^k x$ for all $k = 1, 2, \ldots, n$. In particular, $\sum_i a_{ip}^{(k)} x_i \leq \lambda_0^k x_p = 0$. However, at least some $x_q > 0$ and therefore $a_{qp}^{(k)} = 0$ for all k. This contradicts the irreducibility of A by Theorem 1.1.2 and completes the proof. ∎

We are now ready to prove the Perron-Frobenius Theorem.

Theorem 1.4.4 (Perron-Frobenius Theorem). *Let $A \geq 0$ be an $n \times n$ irreducible matrix. Then*

(i) $Ay = \lambda_0 y$ *for some* $\lambda_0 > 0$, $y > 0$.
(ii) The eigenvalue λ_0 *is geometrically simple.*
(iii) The eigenvalue λ_0 *is maximal in modulus among all the eigenvalues of A. That is, for any eigenvalue* μ *of* A, $|\mu| \leq \lambda_0$.
(iv) The only nonnegative, nonzero eigenvectors of A are just the positive scalar multiples of y.
(v) The eigenvalue λ_0 *is algebraically simple.*
(vi) Let $\lambda_0, \lambda_1, \ldots, \lambda_{(k-1)}$ *be the distinct eigenvalues of A with* $|\lambda_i| = \lambda_0$, $i = 1, 2, \ldots, k-1$. *Then they are precisely the solutions of the equation* $\lambda^k - \lambda_0^k = 0$.

Proof. Let λ_0 be defined as in Lemma 1.4.2. Then $v(\lambda_0) = 0$ and, by Lemma 1.4.3, the game $A - \lambda_0 I$ is completely mixed. If y is an optimal strategy for Player I and x is an optimal strategy for Player II, we know that $(A - \lambda_0 I)y = 0$. This proves *(i)*. We know from Theorem 1.3.6 that rank $(A - \lambda_0 I) = n - 1$. Thus, any solution u to $(A - \lambda_0 I)u = 0$ is a scalar multiple of y, which gives *(ii)*.

Let $Au = \mu u$ for some real or complex eigenvalue μ. Recall that u^+ is defined as

$$u^+ = (|u_1|, |u_2|, \ldots, |u_n|)^T,$$

where $u = (u_1, u_2, \ldots, u_n)^T$. Without loss of generality, let u^+ be a probability vector. Now, for $A = (a_{ij}) \geq 0$, we have

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

Thus $Au^+ \geq |\mu| u^+$ and $(A - |\mu| I)u^+ \geq 0$, showing that $v(|\mu|) \geq 0$. In the case where $|\mu| > \lambda_0$, we can use the completely mixed optimal strategy x of Player II in $A - \lambda_0 I$ to get

$$(A - |\mu| I)^T x = (A - \lambda_0 I)^T x + (\lambda_0 - |\mu|)x = 0 + (\lambda_0 - |\mu|)x < 0,$$

which contradicts $v(|\mu|) \geq 0$. Thus $|\mu| \leq \lambda_0$.

Let $Aw = \alpha w$, where $w \geq 0$. We know $A^T x = \lambda_0 x$ for the optimal strategy x of Player II in $A - \lambda_0 I$. We also have $x > 0$. Thus $\alpha \langle w, x \rangle = \langle \alpha w, x \rangle = \langle Aw, x \rangle = \langle w, A^T x \rangle = \langle w, \lambda_0 x \rangle = \lambda_0 \langle w, x \rangle$. Since $\langle w, x \rangle > 0$, $\alpha = \lambda_0$ and, from *(ii)*, w is a scalar multiple of y. This proves *(iv)*. ∎

Before proving (v) we recall the following:

Lemma 1.4.5. *Let $C(\lambda) = A - \lambda I$. Let $C_{ij}(\lambda)$ be the cofactor of the (i, j)-th entry in $C(\lambda)$. Then*

$$\frac{d}{d\lambda}|C(\lambda)| = -\sum_i C_{ii}(\lambda).$$

Proof. For any continuously differentiable function $F = F(u_1, u_2, \ldots, u_m)$ in $(u_1, \ldots, u_m)$ we have, by the chain rule,

$$dF = \sum_j \frac{\partial F}{\partial u_j} du_j.$$

If the u'_j are functions of a real parameter t, then

$$\frac{dF}{dt} = \sum_j \frac{\partial F}{\partial u_j} \frac{du_j}{dt}.$$

Applying this formula to $F = |C| = F(c_{11}, c_{12}, \ldots, c_{nn})$ we get

$$dF = \sum_i \sum_j \frac{\partial F}{\partial c_{ij}} dc_{ij},$$

where $c_{ij} = c_{ij}(\lambda)$. Thus,

$$\frac{d}{d\lambda}|C(\lambda)| = \sum_i \sum_j \frac{\partial}{\partial c_{ij}}(|C|)\frac{d}{d\lambda}c_{ij}(\lambda).$$

Since $|C| = \sum_j c_{ij}C_{ij}$ we have

$$\frac{\partial}{\partial c_{ij}}|C| = C_{ij} \quad \text{for all } i, j.$$

Therefore,

$$\frac{d}{d\lambda}|C(\lambda)| = \sum_i \sum_j C_{ij}(\lambda)\frac{d}{d\lambda}c_{ij}(\lambda).$$

However, $c_{ij}(\lambda) = a_{ij} - \lambda\delta_{ij}$ (where δ_{ij} denotes the Kronecker δ) and

$$\frac{d}{d\lambda}c_{ij}(\lambda) = -\delta_{ij}.$$

Thus $\frac{d}{d\lambda}|C(\lambda)| = -\sum_i C_{ii}(\lambda)$.

Now continuing with our proof of Theorem 1.4.4, for the characteristic polynomial $p(\lambda) = |A - \lambda I|$,

$$\frac{d}{d\lambda}p(\lambda_0) = \frac{d}{d\lambda}p(\lambda)|_{\lambda=\lambda_0} = -\sum_i C_{ii}(\lambda_0).$$

Because the game $A - \lambda_0 I$ is completely mixed with value zero, by Theorem 1.3.6, all cofactors $C_{ii}(\lambda_0)$ are different from zero and are of the same sign. Thus $\frac{d}{d\lambda} p(\lambda_0) \neq 0$ and, therefore, λ_0 is a simple root of the characteristic polynomial $p(\lambda)$. This proves (v). ■

Before we prove the last assertion of Theorem 1.4.4 we need the following lemma.

Lemma 1.4.6. *Let $x = (x_1, x_2, \ldots, x_n)$ be a vector with complex coordinates such that $|\sum_i x_i| = \sum_i |x_i|$. Then $x_i = \theta |x_i|$, $i = 1, 2, \ldots, n$ for some complex number θ with $|\theta| = 1$.*

Proof. The proof is easily seen by induction on n, and we verify only the case $n = 2$. Let x_1 and x_2 be two complex numbers. If $|x_1 + x_2| = |x_1| + |x_2|$ and $|x_1 + x_2| = 0$, then in that case $x_1 = x_2 = 0$ and the lemma follows. Let $x_1 + x_2 \neq 0$. We can represent $x_1, x_2, x_1 + x_2$ by points P, Q, R in the complex plane. Our hypothesis says that $|OP| + |OQ| = |OR|$, where $|OP|$, for example, denotes the length of the line segment OP with O as the point 0 of the complex plane. Thus O, P, Q, and R lie on the same line and, therefore,

$$\theta = \frac{x_1}{|x_1|} = \frac{x_2}{|x_2|} = \frac{x_1 + x_2}{|x_1 + x_2|}.$$

This is the same as saying $x_i = \theta |x_i|$, $i = 1, 2$, where $\theta = \frac{x_1 + x_2}{|x_1 + x_2|}$. ■

The last assertion in Theorem 1.4.4 does not appear to be based on game theoretic ideas. We first prove the last assertion for the special case $\lambda_0 = 1$, from which the general case easily follows. The first step in this direction is the following:

Lemma 1.4.7. *Let $A \geq 0$ be an $n \times n$ irreducible matrix with the maximal eigenvalue (in modulus) $\lambda_0 = 1$. Then the set of eigenvalues of absolute value unity form a group under multiplication.*

Proof. Let λ and μ be two eigenvalues (not necessarily distinct) of A with $|\lambda| = |\mu| = 1$. Let $Au = \lambda u$ and $Av = \mu v$. As before, let u^+ denote the vector $u^+ = (|u_1|, |u_2|, \ldots, |u_n|)^T$, where $u = (u_1, u_2, \ldots, u_n)$. Without loss of generality, we can assume that u^+ and v^+ are probability vectors. We have

$$\sum_j a_{ij} u_j = \lambda u_i, \quad i = 1, 2, \ldots, n.$$

Taking absolute values on both sides of the above equation, we get

$$|u_i| = |\lambda||u_i| = |\lambda u_i| = \left|\sum_j a_{ij}u_j\right| \leq \sum_j a_{ij}|u_j|, \quad i = 1, 2, \ldots, n. \tag{1.4.3}$$

We thus have $(A-I)u^+ \geq 0$. Since $\lambda_0 = 1$, the game $A-I$ is completely mixed with value zero by Lemma 1.4.3, and so $(A-I)u^+ = 0$. That is, $Au^+ = u^+$ and, similarly, $Av^+ = v^+$. From (1.4.3) we see that

$$\sum_j a_{ij}|u_j| = \left|\sum_j a_{ij}u_j\right|, \quad i = 1, 2, \ldots, n.$$

Applying Lemma 1.4.6 n times we see that, for some θ_i, π_i of absolute value unity,

$$a_{ij}u_j = \theta_i a_{ij}|u_j|, \qquad a_{ij}v_j = \pi_i a_{ij}|v_j| \quad \text{for all } i, j. \tag{1.4.4}$$

However, the probability vector u^+ is the unique eigenvector for the maximal eigenvalue 1 and, therefore, $u^+ = v^+$. Now consider the vector $(\pi_1 u_1, \ldots, \pi_n u_n)$. We show that it is an eigenvector. We have

$$\begin{aligned}
\sum_j a_{ij}\pi_j u_j &= \sum_j a_{ij}u_j\pi_j \\
&= \sum_j \pi_j\theta_i a_{ij}|u_j| \quad \text{[from (1.4.4)]} \\
&= \sum_j \pi_j\theta_i a_{ij}|v_j| \quad (\text{since } |u_j| = |v_j|) \\
&= \theta_i \sum_j \pi_j a_{ij}\left(\sum_k a_{jk}|v_k|\right) \quad (\text{since } Av^+ = v^+) \\
&= \theta_i \sum_j a_{ij}\left(\sum_k \pi_j a_{jk}|v_k|\right) \\
&= \theta_i \sum_j a_{ij}\left(\sum_k a_{jk}v_k\right) \quad \text{[from (1.4.4))]} \\
&= \theta_i \sum_j a_{ij}\mu v_j \quad (\text{since } Av = \mu v) \\
&= \mu\theta_i \sum_j a_{ij}v_j \\
&= \mu\theta_i \sum_j \pi_i a_{ij}|v_j| \quad \text{[from 1.4.4]}
\end{aligned}$$

$$= \mu\theta_i \sum_j \pi_i a_{ij} |u_j| \quad (\text{since } u^+ = v^+)$$

$$= \mu\pi_i \sum_j \theta_i a_{ij} |u_j|$$

$$= \mu\pi_i \sum_j a_{ij} u_j \quad [\text{from (1.4.4)}]$$

$$= \mu\lambda\pi_i u_i.$$

Thus $(\pi_1 u_1, \ldots, \pi_n u_n)^T$ is an eigenvector for the eigenvalue $\mu\lambda$. This shows that the set of eigenvalues of absolute value unity is closed under multiplication. Since there are only a finite number of distinct eigenvalues, we have, for any arbitrary eigenvalue λ of absolute value unity, that $\lambda, \lambda^2, \ldots, \lambda^l$ are eigenvalues and $\lambda^s = \lambda^l$ for some $s > l$. In particular, $\lambda^{s-l} = 1$. If $s - l = 1$, then $\lambda = 1$ is the only eigenvalue of absolute value unity. If $s - l > 1$, then λ^{s-l-1} is also an eigenvalue of A and, further, since $\lambda^{s-l-1} = \lambda^{-1}$, the multiplicative inverse of λ, the set of eigenvalues of absolute value unity form a group under multiplication. ■

The proof of Lemma 1.4.7 is based on Karlin (1959), who attributes it to Bohnenblust. Rota (1961) extended the technique to positive operators on L_1 and L_∞ spaces.

We recall the following property of finite groups:

Lemma 1.4.8. *Let id denote the identity element of a finite group G with N elements. Let $x \neq id$ be any arbitrary element of G with $x^k \neq id$ for $1 \leq k < h$ and $x^h = id$. Then h divides N.*

Proof. This is essentially Lagrange's Theorem. [See, Herstein (1964), p. 35.] ■

Now, continuing with our proof of Theorem 1.4.4, by a direct application of Lemmas 1.4.7 and 1.4.8, $\lambda^k - 1 = 0$ has $\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_{(k-1)}$ as the solution set. The proof for the general case in assertion (vi) of Theorem 1.4.4 follows by considering the matrix $\frac{1}{\lambda_0} A$. This completes the proof of Theorem 1.4.4.

We refer to λ_0 as the *Perron eigenvalue* (or the *Perron root*) and to y as a *(right) Perron eigenvector* of A. A *left Perron eigenvector* of A is simply a right Perron eigenvector of A^T.

We will often refer to basic concepts from the theory of linear programming [see, for example, Gale (1960), Chvátal (1983)] and, in particular, the Duality