Detailed Proof of The Perron–Frobenius Theorem

Arseny M. Shur
Ural Federal University

October 30, 2016

1 Introduction

This famous theorem has numerous applications, but to apply it you should understand it. The proof of a theorem provides the best way of understanding it. Many texts on the Perron–Frobenius theorem can be found on the internet, and some of them contain proofs. However, usually these are proofs of special cases: either for primitive matrices only or even for positive matrices only (the latter is Perron's Theorem). But at least for the applications in graph theory and Markov chains it is better to know the theorem in its full extent, and this is the reason for the appearance of this text. The main line of the proof follows the classical "The Theory of Matrices" book by Felix Gantmacher; some simplifications and clarifications are added.

The text is mainly addressed to graduate students. The reader is supposed to be familiar with the notions and concepts of linear algebra (a two-semester undergraduate course would suit well). For convenience, many of the used notions are recalled in the footnotes.

2 Notation

We write capitals $A, B, C, \ldots$ (sometimes with an index) for matrices, A_{ij} or, say, $(A_kB)_{ij}$ for their elements (or entries), I for the identity matrix. For vectors: $\vec{u}, \vec{v}, \vec{x}, \vec{y}, \vec{z}$; $x_j, (A\vec{x})_j$ for their components. Vectors are columns, to get a row vector we use the transpose of a column: $\vec{x}^\top$. By default, the matrices involved are $n \times n$ and the vectors have size n ($n \geq 2$).

The notation $A \geq B$ (or $A > B$) means that this inequality holds elementwise. Thus, nonnegative (positive) matrices are defined by $A \geq 0$ (resp., $A > 0$). For a complex-valued matrix C, the notation C^+ stands for the matrix obtained from C by replacing the elements with their absolute values: $(C^+)_{ij} = |C_{ij}|$ for all $i, j \in \{1, \ldots, n\}$. All the same applies for vectors.

We write $\imath$ for the imaginary unit and use exponential form $\rho e^{\imath\phi}$ for complex numbers, assuming $\rho \geq 0$ and $0 \leq \phi < 2\pi$.

3 Digraphs and Irreducibility

Every nonnegative $n \times n$ matrix A has an associated digraph G_A with the vertex set $\{1, \ldots, n\}$ in which the edge (i, j) exists iff $A_{ij} > 0$. (We can put this another way: if we replace each positive entry in A with 1, the resulting matrix will be the adjacency matrix of G_A.)

A nonnegative matrix A is called *irreducible* if its associated digraph G_A is strongly connected. This means that for any $i, j \in \{1, \ldots, n\}$ there exists an integer k such that there is an (i, j)-walk of length k in G_A. This fact leads to an alternative, purely algebraic definition of irreducibility: $A \geq 0$ is irreducible if for any $i, j \in \{1, \ldots, n\}$ there exists k such that $(A^k)_{ij} > 0$.

Note that all irreducible matrices are nonzero. The matrices that are not irreducible are called *reducible.*

Let π be an arbitrary permutation of $\{1,\dots,n\}$. Then $\pi(A)$ is the matrix whose entries are defined by the equalities $A_{ij} = (\pi(A))_{\pi(i),\pi(j)}$ for all i, j. In other terms, one can write $\pi(A) = \Pi^{-1}A\Pi$, where Π is the matrix of π; indeed, computing $(\pi(A))_{\pi(i),\pi(j)}$, we get the only nonzero term:

$$(\pi(A))_{\pi(i),\pi(j)} = (\Pi^{-1})_{\pi(i),i} \cdot A_{ij} \cdot \Pi_{j,\pi(j)}.$$

Thus, $\pi(A)$ defines the same linear operator as A in a different basis, obtained from the initial one by permuting the coordinates[1]; as for the digraph, the permutation just renames its vertices.

Remark 1. *A matrix $A \geq 0$ is reducible iff for some permutation π one has $\pi(A) = \left[\begin{smallmatrix} B & C \\ 0 & D \end{smallmatrix}\right]$, where B and D are square matrices. An easy way to see this is to look at digraphs: a digraph is not strongly connected iff its vertices can be partitioned into two classes such that no edge goes from the second class to the first class; an appropriate numeration of vertices gives the adjacency matrix of the form $\left[\begin{smallmatrix} B & C \\ 0 & D \end{smallmatrix}\right]$.*

4 Perron–Frobenius Theorem for Irreducible Matrices

Theorem 1 (Perron–Frobenius Theorem). *Let $A \geq 0$ be an irreducible matrix with the spectral radius r. Then*

1. *r is an eigenvalue of A;*
2. *r has (algebraic and geometric) multiplicity 1;*
3. *r possesses a positive eigenvector;*
4. *all eigenvalues of A with the absolute value r have multiplicity 1; if there are h of them, then they are exactly the (complex) solutions to the equation $\lambda^h = r^h$;*
5. *the spectrum of A, viewed as a multiset, maps to itself under the rotation of the complex plane by the angle $\frac{2\pi}{h}$;*
6. *if $h > 1$, then for some permutation π one has $\pi(A) = \begin{bmatrix} 0 & B_1 & 0 & \cdots & 0 \\ 0 & 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & B_{h-1} \\ B_h & 0 & 0 & \cdots & 0 \end{bmatrix}$, where all blocks on the main diagonal are square*[2].

[1] Recall that a linear operator $\mathcal{A}$ has matrix A in the basis $Y = \{\vec{y}^{(1)},\dots,\vec{y}^{(n)}\}$ if $\mathcal{A}(\vec{x}) = A\vec{x}$ for any $\vec{x}$, where the vectors are given by their coordinates in Y. If we change the basis to $Z = \{\vec{z}^{(1)},\dots,\vec{z}^{(n)}\}$, then the new coordinates of $\vec{x}$ are given by $T^{-1}\vec{x}$, and the matrix of $\mathcal{A}$ changes to an *equivalent* matrix $T^{-1}AT$, where $T = \left[\vec{z}^{(1)},\dots,\vec{z}^{(n)}\right]$ is the *transition* matrix, in which the vectors are given by their coordinates in Y. Equivalent matrices share characteristic polynomials, Jordan normal forms, and other spectral characteristics.

[2] Recall that the *spectrum* of a matrix is the (multi)set of its eigenvalues, and the spectral radius is the maximum absolute value of an eigenvalue (or the minimum radius of a zero-centered circle in the complex plane, containing the spectrum). *Algebraic* multiplicity of an eigenvalue is its multiplicity as a root of the characteristic polynomial of the matrix, and its *geometric* multiplicity is the dimension of its eigenspace (the number of linearly independent eigenvectors it possesses). For the *Jordan normal form* of a matrix A (see footnote 7 for details), the meaning of the multiplicities of an eigenvalue λ is as follows: its geometric multiplicity is the number of Jordan cells with λ, and its algebraic multiplicity is the total size of these cells.

5 Proof of Statements 1–3

Lemma 1. *Let $A \geq 0$ be an irreducible matrix. Then $(A+I)^{n-1} > 0$.*

Proof. Note that a matrix B is positive iff the vector $B\vec{x}$ for any $\vec{x} \geq 0$, $\vec{x} \neq 0$, is positive[3]. Compare the number of nonzero coordinates of an arbitrary nonzero vector $\vec{x} \geq 0$ and of the vector $\vec{y} = (A+I)\vec{x}$. If $x_j > 0$, then $\vec{y}_j = (A\vec{x})_j + x_j > 0$. Thus the set of zero coordinates of $\vec{y}$ is a subset of such a set for $\vec{x}$. Assume that these two sets coincide. Change the basis with the permutation π such that the new coordinates of $\vec{x}$ and $\vec{y}$ are $\left[\begin{smallmatrix}\vec{u}\\0\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}\vec{v}\\0\end{smallmatrix}\right]$ respectively, where $\vec{u}, \vec{v} > 0$ have the same size. Represent $\pi(A)$ as the block matrix $\left[\begin{smallmatrix}B & C\\D & F\end{smallmatrix}\right]$, where B, F are square matrices, and the size of B equals the size of $\vec{u}$. Now the equality

$$\left[\begin{smallmatrix}B & C\\D & F\end{smallmatrix}\right]\left[\begin{smallmatrix}\vec{u}\\0\end{smallmatrix}\right] = \left[\begin{smallmatrix}\vec{v}\\0\end{smallmatrix}\right]$$

implies $D = 0$ because $\vec{u} > 0$. Then A is reducible by Remark 1, contradicting the conditions of the lemma. Hence $\vec{y}$ has strictly less zero coordinates than $\vec{x}$.

Applying this fact to the vectors $\vec{x}$, $(A+I)\vec{x}$, ..., $(A+I)^{n-2}\vec{x}$ and given that the number of zero coordinates of $\vec{x}$ is at most $n-1$, we obtain $(A+I)^{n-1}\vec{x} > 0$, whence the result. $\square$

Step 1. Consider the function

$$r(\vec{x}) = \min_{\substack{j=1,\dots,n\\ x_j \neq 0}} \frac{(A\vec{x})_j}{x_j} \tag{1}$$

in the orthant $X = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} \geq 0,\ \vec{x} \neq 0\}$. Since A represents a linear operator, $r(\vec{x})$ is continuous at any point $\vec{x} > 0$; but it can have the points of discontinuity at the borders of the orthant[4]. A straightforward property of $r(\vec{x})$ is $r(\vec{x})\vec{x} \leq A\vec{x}$ and, moreover,

$$r(\vec{x}) = \max\{\rho \mid \rho\vec{x} \leq A\vec{x}\}. \tag{2}$$

We are going to prove the existence of an "extremal" vector $\vec{z}$ such that

$$r(\vec{z}) = \sup\{r(\vec{x}) \mid \vec{x} \in X\}. \tag{3}$$

The *Weierstrass theorem* (extreme value theorem) says that a continuous function from a non-empty compact set to a subset of $\mathbb{R}$ attains its maximum (and minimum). Further, the *Bolzano–Weierstrass theorem* says that a subset of $\mathbb{R}^n$ is compact iff it is bounded and closed. The set X is neither bounded nor closed. However, $r(\vec{x}) = r(\alpha\vec{x})$ for any $\alpha > 0$, and thus $\sup\{r(\vec{x}) \mid \vec{x} \in X\} = \sup\{r(\vec{x}) \mid \vec{x} \in X^{(1)}\}$, where $X^{(1)}$ consists of vectors of unit length: $X^{(1)} = \{\vec{x} \in X \mid |\vec{x}| = 1\}$. The set $X^{(1)}$ is bounded and closed, and thus compact, but $r(\vec{x})$ can have discontinuities at some elements of $X^{(1)}$ which are not positive vectors. So instead of $X^{(1)}$ we consider the set

$$Y = \{\vec{y} \mid \vec{y} = (A+I)^{n-1}\vec{x} \text{ for some } \vec{x} \in X^{(1)}\}.$$

This set is compact as the image of a compact set under a continuous function (a linear operator). Further, all vectors $\vec{y} \in Y$ are positive by Lemma 1. Hence the function r is continuous on Y and, by the Weierstrass theorem, there exists $\vec{z}$ such that $r(\vec{z}) = \max\{r(\vec{y}) \mid \vec{y} \in Y\}$. Finally, let $\vec{y} = (A+I)^{n-1}\vec{x}$ for some $x \geq 0$, $x \neq 0$, and $\rho\vec{x} \leq A\vec{x}$. Applying the positive matrix $(A+I)^{n-1}$ to both sides of this inequality[5], we get $\rho\vec{y} \leq A\vec{y}$. By (2), this means $r(x) \leq r(y)$. So we have found the vector required by (3):

$$\sup\{r(\vec{x}) \mid \vec{x} \in X\} = \sup\{r(\vec{x}) \mid \vec{x} \in X^{(1)}\} \leq \sup\{r(\vec{y}) \mid \vec{y} \in Y\} = \max\{r(\vec{y}) \mid \vec{y} \in Y\} = r(\vec{z}).$$

[3] Indeed, if $B > 0$ and $x_j > 0$ then for any i one has $(B\vec{x})_i \geq B_{ij}x_j > 0$; and if $B_{ij} = 0$, we can take $x_j > 0$, $x_k = 0$ for $k \neq j$, obtaining $(B\vec{x})_i = 0$.

[4] For example, let $A = \left[\begin{smallmatrix}0 & 1\\1 & 0\end{smallmatrix}\right]$, $\vec{x}^{(k)} = \left[\begin{smallmatrix}1/k\\1\end{smallmatrix}\right]$, $\vec{x} = \lim_{k\to\infty}\vec{x}^{(k)}$. Then $r(\vec{x}^{(k)}) = k$ while $r(\vec{x}) = 0$.

[5] Note that the matrices A and $(A+I)^{n-1}$ commute.

Step 2. Let $\hat{r} = r(\vec{z})$ and $\vec{u} \in X$ be any vector satisfying $r(\vec{u}) = \hat{r}$. We establish some properties of $\hat{r}$ and $\vec{u}$. The main conclusions are written in boldface.
(i) If $\vec{x} > 0$, then $A\vec{x} > 0$ and thus $r(\vec{x}) > 0$. Hence $\hat{r} > 0$.
(ii) Let $\vec{y} = (A+I)^{n-1}\vec{u}$. By (2), $\hat{r}\vec{u} \leq A\vec{u}$. If this is not an equality, then $(A+I)^{n-1}(A\vec{u}-\hat{r}\vec{u}) > 0$ by Lemma 1, implying $\hat{r}\vec{y} < A\vec{y}$. Then there exists $\varepsilon > 0$ such that $(\hat{r}+\varepsilon)\vec{y} < A\vec{y}$. Hence $r(\vec{y}) \geq \hat{r}+\varepsilon$, contradicting (3). Therefore, $\hat{r}\vec{u} = A\vec{u}$. Thus, $\hat{r}$ **is an eigenvalue of** A **with the eigenvector** $\vec{u}$.
(iii) The vector $\vec{u}$ **is positive**, because $\vec{y} = (A+I)^{n-1}\vec{u} = (r+1)^{n-1}\vec{u}$ and $\vec{y} > 0$ by Lemma 1.
(iv) Let α be a (complex) eigenvalue of A with an eigenvector $\vec{v}$. The equality $A\vec{v} = \alpha\vec{v}$ implies $(A\vec{v})^+ = (\alpha\vec{v})^+$. We have, for any $i = j, \ldots, n$,

$$((\alpha\vec{v})^+)_j = |\alpha v_j| = |\alpha||v_j| = (|\alpha|\vec{v}^+)_j, \tag{4}$$
$$((A\vec{v})^+)_j = |A_{j1}v_1 + \cdots + A_{jn}v_n| \leq A_{j1}|v_1| + \cdots + A_{jn}|v_n| = (A\vec{v}^+)_j, \tag{5}$$

and then $|\alpha|\vec{v}^+ \leq A\vec{v}^+$. Since $\vec{v}^+ \geq 0$, $\vec{v}^+ \neq 0$, we have $|\alpha| \leq r(\vec{v}^+) \leq \hat{r}$. Therefore, $\hat{r}$ **is an eigenvalue which is greater than or equal to the absolute value of any other eigenvalue of** A. Hence $\hat{r} = r$, and we are done with statement 1 of the theorem; as $r\vec{u} = A\vec{u}$ (ii) and $\vec{u}$ is positive (iii), we have also proved statement 3.
(v) Reproduce the argument of (iv) for $\alpha = \hat{r} = r$. Since $r\vec{v}^+ \leq A\vec{v}^+$, we have $r(\vec{v}^+) = r$; then $\vec{v}^+$ is a positive eigenvector of A corresponding to r by (ii),(iii). But $A\vec{v}^+ = r\vec{v}^+$ implies the equality in (5). Therefore, the complex numbers $v_1, \ldots, v_n$ must have the same argument ϕ, yielding $\vec{v} = e^{\imath\phi}\vec{v}^+$. Thus any eigenvector corresponding to r is collinear to a positive eigenvector. It remains to note that r cannot have two linearly independent positive eigenvectors, since two such vectors would have a nonzero linear combination $\vec{u}$ which is a non-negative but not positive eigenvector; this is impossible by (iii). Thus, r **has a unique eigenvector**[6], **which can be taken positive**. This positive vector, denoted below by $\vec{z}$, is the *principal eigenvector of* A.

Step 3. We have proved in step 2, (v) that r has geometric multiplicity 1, and thus the Jordan form[7] J of A has a unique cell with the number r. For statement 2, it remains to prove that

($*$) the Jordan form J of A has a cell of size 1 with the number r

Consider the transpose $A^\top$ of A. It is nonnegative, irreducible, and shares the characteristic polynomial with A. Hence all the proof above works for $A^\top$, and there is a positive vector $\vec{y}$ such that $A^\top\vec{y} = r\vec{y}$.

Consider A as the matrix of a linear operator $\mathcal{A}$ in some basis. The orthocomplement $\vec{y}^\perp = \{\vec{x} \mid \vec{y}^\top\vec{x} = 0\}$ of $\vec{y}$ is invariant under $\mathcal{A}$, since $\vec{y}^\top\vec{x} = 0$ implies $\vec{y}^\top(A\vec{x}) = r\vec{y}^\top\vec{x} = 0$. Note that $\vec{z} \notin \vec{y}^\perp$ because both $\vec{y}$ and $\vec{z}$ are positive. Thus, $\mathbb{R}^n$ is the direct sum of two invariant subspaces: $\vec{y}^\perp$ and the eigenspace $\langle\vec{z}\rangle$. Then the matrix of $\mathcal{A}$ in the basis $(\vec{z}, \vec{y}^{(2)}, \ldots, \vec{y}^{(n)})$, where $\vec{y}^\perp = \langle\vec{y}^{(2)}, \ldots, \vec{y}^{(n)}\rangle$, has the block diagonal form $A' = \left[\begin{smallmatrix} r & 0 \\ 0 & Y \end{smallmatrix}\right]$ with some matrix Y of size $n-1$. Bringing Y to the Jordan normal form with some change of basis, we also bring A' to this form (note that A' and A, being equivalent, have the same Jordan normal form). In the upper corner this form has a 1×1 cell with the number r. So we have shown ($*$) and then statement 2.

[6]Up to multiplication by a scalar, of course.

[7]Recall some facts. A *Jordan matrix* has the form $J = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_s \end{bmatrix}$, where $J_k = \begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_k & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_k \end{bmatrix}$ are *Jordan cells* (a 1×1 cell contains the number λ_k). *Jordan's theorem* says that any matrix $A \in \mathbb{C}^{n\times n}$ has a unique, up to the order of blocks, equivalent Jordan matrix, called *Jordan normal form* of A. Since equivalent matrices share the same characteristic polynomial, the numbers λ_k in the Jordan normal form of A are eigenvalues of A, each appearing the number of times equal to its algebraic multiplicity. Note that if a jth column of J is the first column of some cell J_k, it is an eigenvector corresponding to λ_k.

6 Proof of Statements 4–6

Lemma 2. *Let α be an eigenvalue of a matrix $C \in \mathbb{C}^{n\times n}$ such that $C^+ \le A$. Then $|\alpha| \le r$, and the equality implies $C = e^{\imath\phi} DAD^{-1}$, where $\alpha = re^{\imath\phi}$ and D is a diagonal matrix with $D^+ = I$.*

Proof. Let $C\vec{y} = \alpha\vec{y}$. Repeating the argument of step 2 (iv) and using the relation $C^+ \le A$, we get

$$|\alpha|\vec{y}^+ \le C^+\vec{y}^+ \le A\vec{y}^+. \tag{6}$$

Then $|\alpha| \le r(\vec{y}^+) \le r$ by (2),(3) (recall that $r = r(\vec{z})$). Now move to the second statement. If $|\alpha| = r$, then $r(\vec{y}^+) = r$, and thus $\vec{y}^+$ is a positive eigenvector of A corresponding to r by Step 2 (ii),(iii). Hence we have equalities in (6):

$$r\vec{y}^+ = C^+\vec{y}^+ = A\vec{y}^+; \tag{7}$$

from $(A - C^+)\vec{y}^+ = 0$, $A - C^+ \ge 0$, and $\vec{y}^+ > 0$ we get $C^+ = A$. Let

$$\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} |y_1|e^{\imath\phi_1} \\ \vdots \\ |y_n|e^{\imath\phi_n} \end{bmatrix}, \quad D = \begin{bmatrix} e^{\imath\phi_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\imath\phi_n} \end{bmatrix}.$$

Then $\vec{y} = D\vec{y}^+$. Substituting $\alpha = re^{\imath\phi}$, we obtain

$$C\vec{y} = CD\vec{y}^+ = re^{\imath\phi}D\vec{y}^+ \quad \text{and} \quad \underbrace{e^{-\imath\phi}D^{-1}CD}_{F}\,\vec{y}^+ = r\vec{y}^+.$$

So $F\vec{y}^+ = A\vec{y}^+$ by (7) and clearly $F^+ = C^+ (= A)$, implying $F\vec{y}^+ = F^+\vec{y}^+$ by (7). Since $\vec{y}^+ > 0$ and the real parts of all elements of $(F^+ - F)$ are nonnegative, these parts must all be zero, which is possible only if $F = F^+$. Hence $F = A = e^{-\imath\phi}D^{-1}CD$, implying $C = e^{\imath\phi}DAD^{-1}$. □

Step 4. Let $\lambda_0 = r, \lambda_1 = re^{\imath\phi_1}, \ldots, \lambda_{h-1} = re^{\imath\phi_{h-1}}$ be all eigenvalues of A of absolute value r, and $0 = \phi_0 < \phi_1 < \cdots < \phi_{h-1} < 2\pi$. Applying Lemma 2 for $C = A$, $\alpha = \lambda_k$, we get the set of equations

$$A = e^{\imath\phi_k} D_k A D_k^{-1} \quad (k = 0, \ldots, h{-}1;\ D_k^+ = I). \tag{8}$$

Thus the matrix $e^{-\imath\phi_k}A = D_k A D_k^{-1}$ is equivalent to A and has the same spectrum. On the other hand, multiplying a matrix by a scalar, we multiply all its eigenvalues by this scalar. Thus, **the spectrum of A is preserved by the rotation of the complex plane by any angle ϕ_k**. Since the rotation by ϕ_k maps r to λ_k, their multiplicities coincide; so, **each λ_k has multiplicity 1**. Further, the spectrum is preserved, for any $k, l \in \{0, \ldots, h-1\}$, by the rotation by $\phi_k + \phi_l$, $-\phi_k$, and also by 0. Each of these rotations also maps r to some λ_j; so the sums and differences, taken modulo 2π, of angles from $\Phi = \{\phi_0, \ldots, \phi_{h-1}\}$ also belong to this set. Hence Φ is a group under addition modulo 2π. Note that $\phi_1 + \phi_1 = \phi_2$ (otherwise $0 < \phi_1 + \phi_1 - \phi_2 < \phi_1 \notin \Phi$), $\phi_1 + \phi_2 = \phi_3$, …, $\phi_1 + \phi_{h-1} = 0 \pmod{2\pi}$. So we have $\phi_k = k\phi_1$ for all k and $h\phi_1 = 2\pi$. Therefore, $\boldsymbol{\phi_k = \frac{2\pi k}{h}}$, implying statements 4 and 5 of the theorem. We can also write $\lambda_k = r\varepsilon^k$, where $\varepsilon = e^{\imath\frac{2\pi}{h}}$ is the first root of unity of degree h.

Step 5. (i) To prove statement 6, we first return to the equations (8). Such an equation remains true if D_k is replaced by αD_k for any $\alpha \in \mathbb{C}\backslash\{0\}$. Let $D = (D_1)_{11}^{-1} \cdot D_1$. Then $D_{11} = 1$; taking into account that $\lambda_k = r\varepsilon^k$, we have

$$A = \varepsilon DAD^{-1} = \varepsilon^2 D^2 AD^{-2} = \cdots = \varepsilon^{h-1}D^{h-1}AD^{-(h-1)} = D^h AD^{-h}. \tag{9}$$

(ii) The last equation in (9) implies $A_{ij} = (D^h)_{ii} A_{ij} (D^{-h})_{jj}$ for all $i, j \in \{1, \dots, n\}$. If $A_{ij} > 0$, this means $(D^h)_{ii} = (D^h)_{jj}$. Since the digraph G_A of A is strongly connected, it has a $(1, j)$-path for any $j \in \{2, \dots, n\}$. Then the elements $A_{1j_1}, A_{j_1 j_2}, \dots, A_{j_k j}$, representing the edges of this path, are all positive. Hence $(D^h)_{jj} = (D^h)_{11}$. On the other hand, we know that $(D^h)_{11} = (D_{11})^h = 1$, so $\boldsymbol{D^h = I}$. Thus, **all diagonal elements of D are the roots of unity of degree h.**
(iii) The first equality (9) means that

$$A_{ij} = \varepsilon D_{ii} A_{ij} D_{jj}^{-1} \text{ for any } i, j \in \{1, \dots, n\}\,. \tag{10}$$

Let $D_{ii} = \varepsilon^k$. Since A is irreducible, $A_{ij} > 0$ for some $j \neq i$. So we have $D_{jj} = \varepsilon^{k+1}$. Therefore, **all roots of unity of degree h are elements of D.**
(iv) Consider the permutation π which sorts the diagonal entries of D such that if $D_{ii} = \varepsilon^k$ and $D_{i+1,i+1} = \varepsilon^l$, then $k \leq l$ (with the ties broken arbitrarily; 1 is considered as ε^0). Since (10) can be rewritten as

$$A_{\pi(i),\pi(j)} = \varepsilon D_{\pi(i),\pi(i)} A_{\pi(i),\pi(j)} D^{-1}_{\pi(j),\pi(j)}\,,$$

we have $\pi(A) = \varepsilon\pi(D)\pi(A)\pi(D)^{-1}$. The diagonal entries of $\pi(D)$ are arranged into h blocks, with the elements of kth block containing ε^{k-1}. Take some i, j such that $(\pi(A))_{ij} > 0$. If $(\pi(D))_{ii} = \varepsilon^k$, then we have, as in (iii), $(\pi(D))_{jj} = \varepsilon^{k+1}$ (or $(\pi(D))_{jj} = 1$ if $k = h - 1$). Due to the way π sorts the entries of D, all nonempty entries of $\pi(A)$ occur inside the blocks B_j above the main diagonal: $\pi(A) = \begin{bmatrix} 0 & B_1 & 0 & \cdots & 0 \\ 0 & 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & B_{h-1} \\ B_h & 0 & 0 & \cdots & 0 \end{bmatrix}$. Statement 6 and Theorem 1 are thus proved.

7 Useful Related Results on Nonnegative Matrices

Here we collect a few useful facts that are densely related to Theorem 1 and its proof.

Proposition 1. *An irreducible matrix has exactly one nonnegative eigenvector.*

Proof. In fact, all necessary argument is contained in step 3 (see Sect. 5), but we repeat it in a direct way. Let A be an irreducible matrix, r be its spectral radius, $\vec{z}, \vec{y} > 0$ be the principal eigenvectors of A and $A^\top$ respectively. Assume to the contrary that $\vec{u} \geq 0$ is an eigenvector of A not collinear to $\vec{z}$. Since r possesses a unique eigenvector, $A\vec{u} = \alpha\vec{u}$ for some $\alpha \neq r$. One has

$$r\vec{u}^\top\vec{y} = \vec{u}^\top A^\top \vec{y} = (A\vec{u})^\top \vec{y} = \alpha\vec{u}^\top\vec{y},$$

implying $\vec{u}^\top\vec{y} = 0$, which is impossible since $\vec{y} > 0$, $\vec{u} \geq 0$, $\vec{u} \neq 0$. $\square$

In steps 1,2 (Sect. 5) we have shown the following characteristic of the spectral radius r of an irreducible matrix $A \geq 0$:

$$r = \max_{\vec{x} \geq 0} r(\vec{x}) = \max_{\vec{x} \geq 0} \min_{\substack{j=1,\dots,n \\ x_j \neq 0}} \frac{(A\vec{x})_j}{x_j}\,.$$

The choice of maximin looks somewhat arbitrary; what would happen if we would choose the minimax? Let

$$\tilde{r}(\vec{x}) = \begin{cases} +\infty & \text{if } (A\vec{x})_j > 0,\ x_j = 0 \text{ for some } j, \\ \max\limits_{\substack{j=1,\dots,n \\ x_j \neq 0}} \frac{(A\vec{x})_j}{x_j} & \text{otherwise} \end{cases} \tag{11}$$

and

$$\tilde{r} = \tilde{r}(\vec{v}) = \min_{\vec{x} \geq 0} \tilde{r}(\vec{x}). \tag{12}$$

(The existence of the minimum can be proved in exactly the same way as the existence of the maximum of $r(\vec{x})$ in step 1 (Sect. 5.)

Proposition 2. *For the number $\tilde{r}$ and the vector $\vec{v}$ defined by* (11),(12), *$\tilde{r} = r$ and $\vec{v}$ is the principal eigenvector of A.*

Proof. From (12) we get $\tilde{r}\vec{v} - A\vec{v} \geq 0$. Assume that this vector is nonzero. Then by Lemma 1

$$(A+I)^{n-1}(\tilde{r}\vec{v} - A\vec{v}) > 0 \quad \text{and} \quad \vec{u} = (A+I)^{n-1}\vec{v} > 0$$

Since A and $(A+I)^{n-1}$ commute, we have $\tilde{r}\vec{u} - A\vec{u} > 0$; so there exists $\varepsilon > 0$ such that $(\tilde{r} - \varepsilon)\vec{u} - A\vec{u} \geq 0$. This implies $\tilde{r}(\vec{u}) < \tilde{r}$, which is impossible. Therefore, $A\vec{v} = \tilde{r}\vec{v}$. Thus, v is a nonnegative eigenvector of A. By the Theorem 1 (3) and Proposition 1, v is collinear to the principal eigenvector $\vec{z}$ and thus $\tilde{r} = r$. □

Proposition 3. *Let s and S be the minimum and the maximum row sum of an irreducible matrix A with the spectral radius r. Then $s \leq r \leq S$ and, moreover, $s < S$ implies $s < r < S$.*

Proof. Let $\vec{v} = (1, 1, \ldots, 1)^{\perp}$. Then $s = r(\vec{v})$, $S = \tilde{r}(\vec{v})$; see (1),(11). From Proposition 2 we have $r(\vec{v}) \leq r = \tilde{r} \leq \tilde{r}(\vec{v})$. Moreover, if the right inequality turns into equality, then $\vec{v}$ is the principal eigenvector of A; the same is true for the left inequality (see step 2, Sect. 5). The result now follows. □

Proposition 4. *Let $A, B \geq 0$ be two unequal irreducible matrices with the spectral radii r and ρ respectively. Then $A \leq B$ implies $r < \rho$.*

Proof. Let $\vec{z}$ be the principal eigenvector of A, i.e., $A\vec{z} = r\vec{z}$. Further, let

$$r(\vec{x}) = \min_{\substack{j=1,\ldots,n \\ x_j \neq 0}} \frac{(A\vec{x})_j}{x_j}, \quad \rho(\vec{x}) = \min_{\substack{j=1,\ldots,n \\ x_j \neq 0}} \frac{(B\vec{x})_j}{x_j}$$

for any $\vec{x} \geq 0$, $\vec{x} \neq 0$. Recall from steps 1,2 of the proof of Theorem 1 that

$$r = \max_{\substack{\vec{x} \geq 0 \\ \vec{x} \neq 0}} r(\vec{x}), \quad \rho = \max_{\substack{\vec{x} \geq 0 \\ \vec{x} \neq 0}} \rho(\vec{x}).$$

Since $(B\vec{x})_j \geq (A\vec{x})_j$ for any j, one has $r(\vec{x}) \leq \rho(\vec{x})$ for any $\vec{x}$, and thus $r \leq \rho$. Note that $\rho(\vec{z}) \geq r$. If $\rho = r$, then $\vec{z}$ is the principal eigenvector of B (step 2 of the proof of Theorem 1); hence $B\vec{z} = r\vec{z} = A\vec{z}$, which is impossible since $(B - A) \geq 0$ is a non-zero matrix and $\vec{z} > 0$. Therefore, $\rho > r$. □

For arbitrary nonnegative matrices the following weak version of Theorem 1 holds.

Theorem 2. *Let A be a nonnegative matrix with the spectral radius $r > 0$. Then*

1. *r is an eigenvalue of A;*

2. *r possesses a nonnegative eigenvector.*

Proof. Consider a sequence $\{A_m\}_{m=1}^{\infty}$ of *positive* matrices such that $A = \lim_{m\to\infty} A_m$. Say, we set $(A_m)_{ij} = 1/m$ whenever $A_{ij} = 0$ and $(A_m)_{ij} = A_{ij}$ otherwise. Let r_m be the spectral radius of A_m. By Proposition 4, the sequence $\{r_m\}_1^{\infty}$ is decreasing and thus has a limit, say $\bar{r}$. One has

$$\bar{r} \geq r \tag{13}$$

because $r_m > r$ by Proposition 4. Let $\vec{z}^{(m)}$ be the positive eigenvector of A_m of length 1. Then we have

$$A_m \vec{z}^{(m)} = r_m \vec{z}^{(m)} \tag{14}$$

(cf. Proposition 1). The sequence $\{\vec{z}^{(m)}\}_{m=1}^{\infty}$ is bounded and belongs to the compact set of all size-n nonnegative vectors of length 1. Hence it has a subsequence converging to some length 1 nonnegative vector $\vec{z}$. Taking limits for both sides of (14) along this subsequence, we get $A\vec{z} = \bar{r}\vec{z}$. So $\bar{r}$ is an eigenvalue of A; then (13) and the definition of spectral radius imply $\bar{r} = r$. Both statements now follow. □

Theorem 2 says nothing about the multiplicity of r. So we end this note with the statement on the algebraic multiplicity of r. Recall that the *condensation* of a digraph G is the acyclic digraph $\mathsf{Con}(G)$ such that
- the vertices of $\mathsf{Con}(G)$ are the strongly connected components of G;
- two components are connected by a directed edge in $\mathsf{Con}(G)$ if some of their vertices are connected by an edge of the same direction in G.

Let us take $\mathsf{Con}(G_A)$ (suppose it has s vertices) and assign numbers $1, \ldots, s$ to these vertices in topological order[8]. Next we apply to the matrix A (and the graph G_A) any permutation π such that for any $j = 1, \ldots, s-1$ any vertex of the jth component precedes any vertex of the $(j+1)$th component. So we obtain the block matrix

$$\pi(A) = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1,s-1} & B_{1s} \\ 0 & B_{22} & \cdots & B_{2,s-1} & B_{2s} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & B_{ss} \end{bmatrix}, \tag{15}$$

where the graphs of square matrices B_{jj} are exactly the strongly connected components of G_A.

Proposition 5. *Let A be a nonnegative matrix with the spectral radius $r > 0$. Then the algebraic multiplicity of r as the eigenvalue of A equals the number of matrices B_{jj} with the spectral radius r in the representation* (15) *of A.*

Proof. According to (15), the characteristic polynomial of $\pi(A)$ (and thus of A) equals the product of characteristic polynomials of the matrices $B_{11}, \ldots, B_{ss}$. Then the algebraic multiplicity of r as the eigenvalue of A coincides with the sum of its algebraic multiplicities for $B_{11}, \ldots, B_{ss}$. These matrices are irreducible since their graphs are strongly connected. By Theorem 1, if r is an eigenvalue of B_{jj}, it has algebraic multiplicity 1 (since r is the spectral radius of A, it is the spectral radius of such matrix B_{jj}). The result now follows. □

Acknowledgement

The author is grateful to Pavel Gein for careful reading of this note and useful remarks.

[8]That is, if (i, j) is an edge, then $i < j$.