The Perron-Frobenius Theorem and its Application to Population Dynamics

by

Jerika Baldin

A project submitted in partial fulfillment
of the requirements for the degree of
Mathematics (April 2017)

Department of Mathematical Sciences
Lakehead University
2017

ACKNOWLEDGEMENTS

First and foremost, I wish to thank my supervisor, Dr. Christopher Chlebovec, for his guidance and support throughout the research and write up phases of my thesis. His suggestions and assistance were greatly appreciated. I wish to also extend my gratitude to the faculty, staff, and students with whom I've had the pleasure of coming into contact with at Lakehead University.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS 2

CHAPTER

I. Preliminaries 1

1.1 Positive and Non-negative Matrices 1
1.2 Jordan Normal Form 2
1.3 Norm 5
1.4 Graphs and Matrices 5
1.5 Spectrum 7

II. The Perron-Frobenius Theorem for Positive Matrices 10

2.1 The Perron-Frobenius Theorem Stated 10
2.2 Positive Eigenpair 11
2.3 Index of $\rho(\mathbf{A})$ 14
2.4 Multiplicities of $\rho(\mathbf{A})$ 17
2.5 Collaltz-Wielandt Formula 18

III. Perron Frobenius Theorem for Non-negative Matrices 20

3.1 Non-negative Eigenpair 20
3.2 Reducibility 22
3.3 Primitive Matrices 26

IV. The Leslie Model 30

4.1 The Rabbit Problem 30
4.2 Leslie Model 32

REFERENCES 35

CHAPTER I

Preliminaries

1.1 Positive and Non-negative Matrices

Let $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ be matrices of size $n \times n$. A matrix $\mathbf{A}$ is said to positive if every $a_{ij} > 0$ and is denoted by $\mathbf{A} > 0$. Also, $\mathbf{A}$ is said to be non-negative if every $a_{ij} \geq 0$ and is denoted by $\mathbf{A} \geq 0$. A matrix $\mathbf{A} > \mathbf{B}$ if $a_{ij} > b_{ij}$ for all i, j. Listed below are some useful properties about matrices.

Proposition I.1. [3, p.663] *Let* $\mathbf{P}$ *and* $\mathbf{N}$ *be matrices of size* $n \times n$ *with* $\mathbf{x}, \mathbf{u}, \mathbf{v}$, *and* $\mathbf{z}$ *vectors.*

(i) $\mathbf{P} > \mathbf{0}, \mathbf{x} \geq \mathbf{0}, \mathbf{x} \neq \mathbf{0} \Longrightarrow \mathbf{Px} > \mathbf{0}$

(ii) $\mathbf{N} \geq \mathbf{0}, \mathbf{u} \geq \mathbf{v} \geq \mathbf{0} \Longrightarrow \mathbf{Nu} \geq \mathbf{Nv}$

(iii) $\mathbf{N} \geq \mathbf{0}, \mathbf{z} > \mathbf{0}, \mathbf{Nz} = \mathbf{0} \Longrightarrow \mathbf{N} = \mathbf{0}$

(iv) $\mathbf{N} > \mathbf{0}, \mathbf{u} > \mathbf{v} > \mathbf{0} \Longrightarrow \mathbf{Nu} > \mathbf{Nv}$

Proof.

(i) Let $\mathbf{P} = (a_{ij})$ and let $\mathbf{P}_i = (a_{i1} \cdots a_{ij} \cdots a_{in})$ be the ith row of $\mathbf{P}$. Since $\mathbf{x} \geq \mathbf{0}, \mathbf{x} \neq \mathbf{0}$ there is some j with $x_j > 0$. Then the ith entry of $\mathbf{Px}$ is $\mathbf{P}_i\mathbf{x} = a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n$. Also, since $a_{ij}x_j > 0$ we have that $(\mathbf{Px})_i =$

$a_{i1}x_1 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n \geq a_{ij}x_j > 0$ for each $i = 1, 2, \ldots n$.

(ii) Let $\mathbf{N} = (b_{ij})$. Let $\mathbf{u} = (u_i)_i$ and $\mathbf{v} = (v_i)_i$. Since $\mathbf{u} \geq \mathbf{v}$ we have $u_i \geq v_i$ for each $i = 1, 2, \ldots, n$. Then, $b_{ij}u_i \geq b_{ij}v_i$ which implies that $b_{i1}u_1 + b_{i2}u_2 + \cdots + b_{in}u_n \geq b_{i1}v_1 + b_{i2}v_2 + \cdots + b_{in}v_n$. It follows that $(\mathbf{Nu})_i \geq (\mathbf{Nv})_i$ for each i and therefore, $\mathbf{Nu} \geq \mathbf{Nv}$.

(iii) Let $\mathbf{N} = (b_{ij})$ and $\mathbf{z} = (z_i)_i$. Since $\mathbf{Nz} = \mathbf{0}$ we have that $(\mathbf{Nz})_i = 0$ for each i. We know $(\mathbf{Nz})_i = b_{i1}z_1 + b_{i2}z_2 + \cdots + b_{in}z_n = 0$. Since $\mathbf{z} > \mathbf{0}$ and $\mathbf{N} \geq \mathbf{0}$ we know that $b_{ij} = 0$ for each $j = 1, 2, \ldots, n$. Therefore, since $b_{ij} = 0$ we have that $\mathbf{N} = \mathbf{0}$.

(iv) Let $\mathbf{N} = (b_{ij})$. Let $\mathbf{u} = (u_i)_i$ and $\mathbf{v} = (v_i)_i$. If $\mathbf{N} > \mathbf{0}$ then $b_{ij} > 0$. Since $u_i > v_i$ we have that $b_{i1}u_1 + b_{i2}u_2 + \cdots + b_{in}u_n > b_{i1}v_1 + b_{i2}v_2 + \cdots + b_{in}v_n$. Therefore, $\mathbf{Nu} > \mathbf{Nv}$. □

1.2 Jordan Normal Form

Theorem I.2. Jordan Normal Form [3, p.590]

A Jordan matrix or a matrix in Jordan Normal Form is a block matrix that has Jordan blocks down its diagonal and is zero everyone else. Every matrix $\mathbf{A}$ *in* $\mathbf{M}_n(\mathbb{C})$ *with distinct eigenvalues* $\sigma(\mathbf{A}) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ *is similar to a matrix in Jordan Normal Form. That is, there exists an invertible matrix* $\mathbf{P}$ *such that*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J} = \begin{pmatrix} \mathbf{J}(\lambda_1) & 0 & \cdots & 0 \\ 0 & \ddots & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{J}(\lambda_n) \end{pmatrix}.$$

For each eigenvalue $\lambda_j \in \sigma(\mathbf{A})$ there exists one Jordan segment $\mathbf{J}(\lambda_j)$ that is made up of t_j Jordan blocks where $t_j = dimN(\mathbf{A} - \lambda_j\mathbf{I})$. Indeed,

$$\mathbf{J}(\lambda_j) = \begin{pmatrix} \mathbf{J}_1(\lambda_j) & & & \\ & \mathbf{J}_2(\lambda_j) & & \\ & & \ddots & \\ & & & \mathbf{J}_{t_j}(\lambda_j) \end{pmatrix}. \tag{1.1}$$

Then, $\mathbf{J}_*(\lambda_j)$ represents an arbitrary block of $\mathbf{J}(\lambda_j)$. That is,

$$\mathbf{J}_*(\lambda_j) = \begin{pmatrix} \lambda_j & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_j \end{pmatrix}. \tag{1.2}$$

Also, $\mathbf{A}^k = \mathbf{P}\mathbf{J}^k\mathbf{P}^{-1}$ and

$$\mathbf{J}_*^k(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}^k = \begin{pmatrix} \lambda^k & \binom{k}{1}\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} & \cdots & \binom{k}{m-1}\lambda^{k-m+1} \\ & \lambda^k & \binom{k}{1}\lambda^{k-1} & & \vdots \\ & & \ddots & \ddots & \binom{k}{1}\lambda^{k-1} \\ & & & & \lambda^k \end{pmatrix}_{m\times m}. \tag{1.3}$$

Let $\mathbf{A}$ be a square matrix. The algorithm below is used to find the Jordan Normal Form of matrix $\mathbf{A}$.

(1) Find the distinct eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ of $\mathbf{A}$ such that $\lambda \in \sigma(\mathbf{A})$. For each eigenvalue λ_j we have the segment $\mathbf{J}(\lambda_j)$ which is made up of $t_j = dimNul(\mathbf{A} - \lambda_j\mathbf{I})$ Jordan blocks, $\mathbf{J}_*(\lambda_j)$. We obtain $\mathbf{J}(\lambda_j)$ as seen in (1.1) where $\mathbf{J}_*(\lambda_j)$ as seen in (1.2).

(2) Find the rank for each Jordan segment $\mathbf{J}(\lambda_1), \ldots, \mathbf{J}(\lambda_n)$ where $i = 1, \ldots, n$ and $r_i(\lambda_j) = rank((\mathbf{A} - \lambda_j\mathbf{I})^i)$.Then,

$$r_1(\lambda_i) = r_1(\mathbf{A} - \lambda_i\mathbf{I})$$

$$r_2(\lambda_i) = r_2(\mathbf{A} - \lambda_i\mathbf{I})^2$$

$$\vdots$$

$$r_n(\lambda_i) = r_n(\mathbf{A} - \lambda_i\mathbf{I})^n$$

Stop computing the rank value once the value begins to repeat. The $index(\lambda_i) =$ the smallest positive integer k such that $rank((\mathbf{A}-\lambda_i\mathbf{I})^k) = rank((\mathbf{A}-\lambda_i\mathbf{I})^{k+1})$. This k value gives the size of the largest Jordan block for $\mathbf{J}(\lambda_i)$.

(3) The number of $i \times i$ Jordan blocks in $\mathbf{J}(\lambda_j)$ is given by the formula,

$$v_i(\lambda_j) = r_{i-1}(\lambda_j) - 2r_i(\lambda_j) + r_{i+1}(\lambda_j).$$

This computed value gives the size of all of the individual Jordan blocks. Once you have established how many of each block size are needed matrix $\mathbf{J}$ can be constructed.

Remark I.3. [2, p.683] *Recall that the algebraic multiplicity is the number of times λ appears as a root of the characteristic polynomial and the geometric multiplicity is the number of linearly independent eigenvectors associated with λ or in other words, $geomult_{\mathbf{A}}(\lambda) = dimN(\mathbf{A} - \lambda\mathbf{I})$. Also, it is important to note that we can use Jordan Normal Form to find the algebraic and geometric multiplicities directly.*

For some eigenvalue λ, the algebraic multiplicity is equal to the sum of the sizes of all Jordan blocks in $\mathbf{J}(\lambda)$. The geometric multiplicity is equal to the number of Jordan blocks associated with λ. When $index(\lambda) = 1$ we have that the largest Jordan *block is of the size* 1×1*. In this case, from Theorem I.2 (1.1) it is evident that* $algmult_{\mathbf{A}}(\lambda) = geomult_{\mathbf{A}}(\lambda)$.

1.3 Norm

Definition I.4. [3, p.280] A matrix norm is a function $||*||$ from the set of all complex matrices into $\mathbb{R}$ satisfies the following properties:

(i) $||\mathbf{A}|| \geq 0$ and $||\mathbf{A}|| = 0 \Longleftrightarrow \mathbf{A} = \mathbf{0}$

(ii) $||\alpha\mathbf{A}|| = |\alpha|||\mathbf{A}||$ for all scalars α

(iii) $||\mathbf{A} + \mathbf{B}|| \leq ||\mathbf{A}|| + ||\mathbf{B}||$

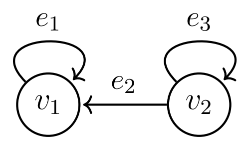
(iv) $||\mathbf{AB}|| \leq ||\mathbf{A}||||\mathbf{B}||$

An example of a general matrix norm is the infinity norm. The infinity norm of a square matrix is the maximum of the absolute row sums and is denoted by $||\mathbf{A}||_\infty$. That is, $||\mathbf{A}||_\infty = \max_{1 \leq i \leq n} \left\{ \sum_{j=1}^{n} |a_{ij}| \right\}$.

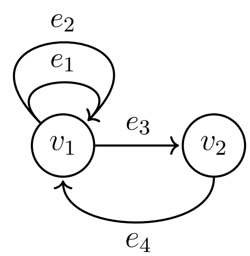
1.4 Graphs and Matrices

Definition I.5. [4, p.3] A directed graph $G = \{G^0, G^1, r, s\}$ consists of a finite set G^0 of vertices such that $G^0 = \{v_1, v_2, \ldots, v_m\}$, a finite set G^1 of edges such that $G^1 = \{e_1, e_2, \ldots, e_n\}$ and maps $r, s : G^1 \longrightarrow G^0$ where $r(e_i)$ is the range of the edge e_i and $s(e_i)$ is the source of the edge e_i. A path in G is a sequence of edges $e = e_1 e_2 \cdots e_n$ with $r(e_i) = s(e_{i+1})$ for $1 \leq i < n$ where e has length $|e| = n$. A cycle is a path $e = e_1 \cdots e_n$ with $r(e_n) = s(e_1)$. A simple cycle occurs when there is no repetition of edges or vertices along the path except maybe your starting and ending vertex.

Example I.6.



(1.4)



(1.5)

In Figure (1.4) we have that $G^0 = \{v_1, v_2\}$ and $G^1 = \{e_1, e_2, e_3\}$ where $r(e_1) = v_1, s(e_1) = v_1, r(e_2) = v_1, s(e_2) = v_2$, and $r(e_3) = v_2, s(e_3) = v_2$. In Figure (1.5) we have that $G^0 = \{v_1, v_2\}$ and $G^1 = \{e_1, e_2, e_3, e_4\}$ where $r(e_1) = v_1, s(e_1) = v_1, r(e_2) = v_1, s(e_2) = v_1, r(e_3) = v_2, s(e_3) = v_1$, and $r(e_4) = v_1, s(e_4) = v_2$.

A graph is said to be strongly connected if between any two vertices v and w there exists a path from v to w and from w to v. In Figure (1.4) the graph is not strongly connected since there is no path from v_1 to v_2. On the other hand, the graph in Figure (1.5) is strongly connected since there exists a path from v_1 to v_2 and vice versa.

A graph associated with a non-negative matrix $\mathbf{A}$ is denoted by $G(\mathbf{A})$ where the directed edge (v_i, v_j) is in E exactly when $a_{ij} \neq 0$. A non-negative matrix $\mathbf{A}$ in which each non-zero entry is replaced by 1 is called the adjacency matrix for $G(\mathbf{A})$ and is denoted by $\mathbf{A}_G$. Given a graph G, the adjacency matrix $\mathbf{A}_G$ is an $|E^0| \times |E^0|$ matrix where $\mathbf{A}_G(i, j)$ is 1 when there is a path of length 1 from v_i to v_j and 0 otherwise.

Example I.7. Consider $\mathbf{A} = \begin{pmatrix} 0.7 & 0 \\ 0.2 & 0.9 \end{pmatrix}$. Then, $G(\mathbf{A})$ is the graph from Figure

(1.4) with adjacency matrix

$$\mathbf{A_G} = \begin{array}{c c} & \begin{array}{cc} v_1 & v_2 \end{array} \\ \begin{array}{c} v_1 \\ v_2 \end{array} & \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \end{array}.$$

Given the graph in Figure (1.5) its adjacency matrix is

$$\mathbf{A_G} = \begin{array}{c c} & \begin{array}{cc} v_1 & v_2 \end{array} \\ \begin{array}{c} v_1 \\ v_2 \end{array} & \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \end{array}.$$

Lemma I.8. [3, p.672] *Let $\mathbf{A} \geq 0$. Then $(\mathbf{A}^k)_{ij} > 0$ if and only if there exists a path of length k from v_i to v_j in $G(\mathbf{A})$.*

Proof. Let $A = (a_{ij})$ and let $a_{ij}^{(k)}$ denote the (i,j)-entry in $\mathbf{A}^k$. Then, $a_{ij}^{(k)} = \sum_{h_1,\ldots,h_{k-1}} a_{ih_1} a_{h_1 h_2} \cdots a_{h_{k-1}j} > 0$ if and only if there exists a set of indices $h_1, h_2, \ldots, h_{k-1}$ such that $a_{ih_1} > 0$ and $a_{h_1h_2} > 0$ and $\cdots$ and $a_{h_{k-1}j} > 0$. Since $a_{st} > 0$ for some arbitrary s and t, this is equivalent to saying there is a path of length 1 from v_s to v_t in $G(\mathbf{A})$. The previous sentence is equivalent to finding a path of length k from v_i to v_j. Therefore, there exists a path of length k in $G(\mathbf{A})$ from v_i to v_j if and only if $a_{ij}^{(k)} > 0$. □

1.5 Spectrum

The spectrum of a square matrix $\mathbf{A}$ is denoted by $\sigma(\mathbf{A})$ and is equal to the set of all eigenvalues of $\mathbf{A}$. The spectral radius of $\mathbf{A}$ is the maximum eigenvalue of $\mathbf{A}$ in absolute value; that is, $\rho(\mathbf{A}) = \max\{|\lambda| : \lambda \in \sigma(\mathbf{A})\}$. The following are important facts to note about spectral radius.

Proposition I.9. [3]

(i) $\mathbf{A} > \mathbf{0} \Longrightarrow \rho(\mathbf{A}) > \mathbf{0}$

(ii) $\mathbf{A} > \mathbf{0} \Leftrightarrow \mathbf{A}/\rho(\mathbf{A}) > \mathbf{0}$

(iii) If c is a constant then, $\rho(c\mathbf{A}) = |c|\rho(\mathbf{A})$

(iv) Let $r > 0$ then, $\rho(\mathbf{A}) = r \Leftrightarrow \rho(\mathbf{A}/r) = 1$

(v) If $\rho(\mathbf{A}) < 1$ then, $\lim_{k\to\infty} \mathbf{A}^k = 0$*. That is,* $\lim_{k\longrightarrow\infty} (\mathbf{A}^k)_{ij} = 0$ *for each* i, j*.*

Proof.

(i) Let $\mathbf{A} = (a_{ij})$ with $a_{ij} > 0$. We know that $\rho(\mathbf{A}) \geq 0$. Suppose $\rho(\mathbf{A}) = 0$. Then it follows that $\sigma(\mathbf{A}) = \{\mathbf{0}\}$. If this were the case we would have the Jordan Normal Form of $\mathbf{A}$ would imply $\mathbf{A}^k = 0$ for some $k \geq 0$. If $\mathbf{A} > 0$ then $\mathbf{A}^k \neq 0$ for all $k \geq 0$. This poses as contradiction that $\rho(\mathbf{A}) = 0$. Therefore, $\rho(\mathbf{A}) > 0$.

(ii) ($\Longrightarrow$) If $\mathbf{A} > \mathbf{0}$ from (i) we know that $\rho(\mathbf{A}) > 0$. From the definition of spectral radius and positive matrices we know that any entry $a_{ij} > 0$ of $\mathbf{A}$ divided by $\rho(\mathbf{A}) > 0$ will also be positive.

($\Longleftarrow$) Suppose $\mathbf{A}/\rho(\mathbf{A}) > 0$. From (i) we have that $\rho(\mathbf{A}) > 0$. If $\rho(\mathbf{A}) > 0$ then for $\mathbf{A}/\rho(\mathbf{A}) > \mathbf{0}$ we must have $\mathbf{A} > \mathbf{0}$.

(iii) Let c be a constant. We know that $\rho(\mathbf{A}) = \max\{|\lambda| : \lambda \in \sigma(\mathbf{A})\}$. If λ is an eigenvalue of $\mathbf{A}$ with eigenvector $\mathbf{x}$ then multiplying $\mathbf{A}$ by some constant c we get,

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

$$c\mathbf{A}\mathbf{x} = c\lambda\mathbf{x}$$

$$(c\mathbf{A})\mathbf{x} = (c\lambda)\mathbf{x}$$

This shows the eigenvalue of $c\mathbf{A}$ is $c\lambda$. Let $\{\lambda_1, \ldots, \lambda_p\}$ be the distinct eigenvalues of $\mathbf{A}$ then it follows that, $\rho(\mathbf{A}) = \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_p|\}$. We know $\rho(\mathbf{A}) = |\lambda_M|$ for some λ_M. Since $|\lambda_i| \leq |\lambda_M|$ for each i we have that $|c||\lambda_i| \leq |c||\lambda_M|$. Indeed,

$$
\begin{aligned}
\rho(c\mathbf{A}) &= \max\{|c\lambda_1|, |c\lambda_2|, \ldots, |c\lambda_p|\} \\
&= \max\{|c||\lambda_1|, |c||\lambda_2|, \ldots, |c||\lambda_p|\} \\
&= |c||\lambda_M| \\
&= |c|\rho(\mathbf{A})
\end{aligned}
$$

Therefore, if c is constant the eigenvalues of $c\mathbf{A}$ are $\{c\lambda_1, \ldots, c\lambda_p\}$.

(iv) Let $\rho(\mathbf{A}) = r$ where $r > 0$.

($\Longrightarrow$) We want to show that if $\rho(\mathbf{A}) = r$ then $\rho\left(\dfrac{\mathbf{A}}{r}\right) = 1$. From (iii) we have that, $\rho\left(\dfrac{\mathbf{A}}{r}\right) = \left|\dfrac{1}{r}\right|\rho(\mathbf{A}) = \dfrac{1}{r}(r) = 1$ as required.

($\Longleftarrow$) We want to show that if $\rho\left(\dfrac{\mathbf{A}}{r}\right) = 1$ then $\rho(\mathbf{A}) = r$. If $\rho\left(\dfrac{\mathbf{A}}{r}\right) = 1$ from (iii) it follows that $\left|\dfrac{1}{r}\right|\rho(\mathbf{A}) = 1$ and since $r > 0$ we have that $\rho(\mathbf{A}) = 1(r) = r$.

(v) Since $\rho(\mathbf{A}) < 1$, then $|\lambda| < 1$ for all $\lambda \in \sigma(\mathbf{A})$. We know from Theorem I.2 (1.3) that each entry of $\mathbf{J}_*^k$ is of the form $\binom{k}{j}\lambda^{k-j}$. Since $\binom{k}{j} = \dfrac{k(k-1)\cdots(k-j+1)}{j!} \leq \dfrac{k^j}{j!}$ this implies that $\binom{k}{j}\lambda^{k-j} \leq \dfrac{k^j}{j!}\lambda^{k-j}$. Taking the absolute value we get, $\left|\binom{k}{j}\lambda^{k-j}\right| \leq \dfrac{k^j}{j!}|\lambda|^{k-j} \longrightarrow 0$. Taking $\displaystyle\lim_{k\longrightarrow\infty} \frac{k^j/j!}{|\lambda|^{j-k}} = \frac{\infty}{\infty}$ so, L'Hopital's rule can be applied. Repeatedly applying L'Hopital's rule it is evident that k^j tends to 1 and $|\lambda|^{j-k}$ will eventually approach ∞ so it follows that $\displaystyle\lim_{k\longrightarrow\infty} \frac{k^j/j!}{|\lambda|^{j-k}} \longrightarrow 0$. Since $\mathbf{J}_*^k \longrightarrow 0$ we have that $\mathbf{A}^k \longrightarrow 0$. $\square$

CHAPTER II

The Perron-Frobenius Theorem for Positive Matrices

2.1 The Perron-Frobenius Theorem Stated

Theorem II.1. [3, p.667] *If* $\mathbf{A}$ *is our* $n \times n$ *matrix, with* $\mathbf{A} > 0$ *and* $r = \rho(\mathbf{A})$*, then the following statements are true.*

(i) $r > 0$

(ii) $r \in \sigma(\mathbf{A})$ *(r is called the* **Perron-Frobenius eigenvalue***)*

(iii) $algmult_{\mathbf{A}}(r) = 1$ *and* $geomult_{\mathbf{A}}(r) = 1$

(iv) There exists an eigenvector $\mathbf{x} > 0$ *such that* $\mathbf{Ax} = r\mathbf{x}$

(v) The **Perron- Frobenius eigenvector** *is the unique vector defined by*

$$\mathbf{Ap} = r\mathbf{p}, \ \ \mathbf{p} > 0 \text{ and } ||\mathbf{p}||_1 = 1$$

and, except for positive multiples of $\mathbf{p}$*, there are no other non-negative eigenvectors for* A*, regardless of the eigenvalue.*

(vi) r *is the only eigenvalue on the spectral circle of* $\mathbf{A}$*.*

(vii) $r = \max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x})$ *(the Collatz-Wielandt formula), where*

$$f(\mathbf{x}) = \min_{\substack{1\le i\le n \\ x_i\neq 0}} \frac{[\mathbf{Ax}]_i}{x_i} \text{ and } \mathcal{N} = \{\mathbf{x} : \mathbf{x} \geq 0 \text{ with } \mathbf{x} \neq 0\}.$$

We are going to prove all these statements in the following sections.

2.2 Positive Eigenpair

Remark II.2. $|\mathbf{Ax}| \leq |\mathbf{A}||\mathbf{x}|$. *Indeed, using the triangle inequality we get,*

$$\begin{aligned}(|\mathbf{Ax}|)_i &= |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n| \\ &\leq |a_{i1}x_1| + |a_{i2}x_2| + \cdots + |a_{in}x_n| \\ &\leq |a_{i1}||x_1| + |a_{i2}||x_2| + \cdots + |a_{in}||x_n| \\ &= |\mathbf{A}|_i|\mathbf{x}|\end{aligned}$$

where $|\mathbf{A}|_i$ *represents the* i*th row of* $|\mathbf{A}|$*.*

Let $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. The eigenvalues of $\mathbf{A}$ are $\lambda_1 = 1, \lambda_2 = -5, \lambda_3 = 2$. That is, $\sigma(\mathbf{A}) = \{1, 2, -5\}$. This implies that $\rho(\mathbf{A}) = |-5| = 5$, but $5 \notin \sigma(\mathbf{A})$. Below we will show that if $\mathbf{A} > 0$, then the spectral radius of $\mathbf{A}$ is in fact an eigenvalue of $\mathbf{A}$.

Theorem II.3. Positive Eigenpair [3, p.663]

Let $\mathbf{A} > 0$ *be an* $n \times n$ *matrix then the following hold true,*

(i) $\rho(\mathbf{A}) \in \sigma(\mathbf{A})$

(ii) If $\mathbf{Ax} = \rho(\mathbf{A})\mathbf{x}$ *then* $\mathbf{A}|\mathbf{x}| = \rho(\mathbf{A})|\mathbf{x}|$ *and* $|\mathbf{x}| > 0$

Proof. We will first prove this for the case when $\rho(\mathbf{A}) = 1$. Since $\rho(\mathbf{A}) = |\lambda| = 1$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ we have that $(\lambda, \mathbf{x})$ is an eigenpair of $\mathbf{A}$. Then,

$$|\mathbf{x}| = |\lambda||\mathbf{x}| = |\lambda\mathbf{x}| = |\mathbf{A}\mathbf{x}|$$

$$\leq |\mathbf{A}||\mathbf{x}|$$

$$= \mathbf{A}|\mathbf{x}|$$

From Remark II.2 we have that $|\mathbf{x}| \leq \mathbf{A}|\mathbf{x}|$.

Let $\mathbf{z} = \mathbf{A}|\mathbf{x}|$ and $\mathbf{y} = \mathbf{z} - |\mathbf{x}|$. We know that $\mathbf{0} < |\mathbf{x}| \leq \mathbf{A}|\mathbf{x}|$, so it follows that $\mathbf{z} > \mathbf{0}$. Since, $|\mathbf{x}| \leq \mathbf{A}|\mathbf{x}|$ we have that $\mathbf{y} \geq \mathbf{0}$. We want to show $\mathbf{y} = \mathbf{0}$. Suppose on the contrary that $\mathbf{y} \neq \mathbf{0}$. Let $c_i = \left(\frac{(\mathbf{A}\mathbf{y})_i}{z_i}\right)$. Since $(\mathbf{A}\mathbf{y})_i > 0$ and $z_i > 0$ we have $c_i > 0$ for each $i = 1, 2, \ldots, n$. Let $c = \min\{c_i\}$ and take $\epsilon < c$. Then $\epsilon \leq c_i$ for each i. Therefore,

$$\left(\frac{(\mathbf{A}\mathbf{y})_i}{z_i}\right) > \epsilon$$

$$(\mathbf{A}\mathbf{y})_i > \epsilon z_i$$

$$(\mathbf{A}\mathbf{y})_i > (\epsilon\mathbf{z})_i$$

$$\mathbf{A}\mathbf{y} > \epsilon\mathbf{z}$$

We know that $\mathbf{A}\mathbf{y} > \epsilon\mathbf{z}$ or equivalently,

$$\mathbf{A}(\mathbf{z} - |\mathbf{x}|) > \epsilon(\mathbf{A}|\mathbf{x}|)$$

$$\mathbf{A}\mathbf{z} - \mathbf{A}|\mathbf{x}| > \epsilon(\mathbf{A}|\mathbf{x}|)$$

$$\mathbf{A}\mathbf{z} > \epsilon\mathbf{A}|\mathbf{x}| + \mathbf{A}|\mathbf{x}|$$

$$\mathbf{A}\mathbf{z} > \mathbf{A}|\mathbf{x}|(1+\epsilon)$$

$$\left(\frac{\mathbf{A}}{1+\epsilon}\right)\mathbf{z} > \mathbf{A}|\mathbf{x}|$$

$$\left(\frac{\mathbf{A}}{1+\epsilon}\right)\mathbf{z} > \mathbf{z}$$

Let $\mathbf{B} = \left(\frac{\mathbf{A}}{1+\epsilon}\right)$ which implies $\mathbf{B}\mathbf{z} > \mathbf{z}$. Since $\mathbf{A} > 0$ we have that $\mathbf{B} > 0$ so we get the following result,

$$\mathbf{B}\mathbf{z} > \mathbf{z}$$

$$\mathbf{B}^2\mathbf{z} > \mathbf{B}\mathbf{z} > \mathbf{z}$$

$$\mathbf{B}^3\mathbf{z} > \mathbf{B}\mathbf{z} > \mathbf{z}$$

$$\vdots$$

$$\mathbf{B}^k\mathbf{z} > \mathbf{z}$$

We know that $\rho(\mathbf{B}) = \rho\left(\frac{\mathbf{A}}{1+\epsilon}\right) = \left(\frac{1}{1+\epsilon}\right)\rho(\mathbf{A}) = \frac{1}{1+\epsilon} < 1$ so, from Proposition I.9 (v) we have that $\lim_{k\longrightarrow\infty} \mathbf{B}^k = 0$ for each k. Then $\mathbf{0} \geq \mathbf{z}$ which is a contradiction, so $\mathbf{y} = \mathbf{0}$. Since $\mathbf{y} = \mathbf{0} = \mathbf{A}|\mathbf{x}| - |\mathbf{x}|$ we have that $\mathbf{x}$ is an eigenvector for $\mathbf{A}$ with corresponding eigenvalue $\rho(\mathbf{A}) = 1$.

In the general case, suppose $\rho(\mathbf{A}) = k$. By Proposition I.9 (i), we know $k > 0$. Since $\rho(\mathbf{A}) = k$ from Proposition I.9 (iii) we have $\rho\left(\frac{\mathbf{A}}{k}\right) = 1$ so 1 is an eigenvalue of $\frac{\mathbf{A}}{k}$. We want to show $\rho(\mathbf{A}) \in \sigma(\mathbf{A})$. Since 1 is an eigenvalue of $\frac{\mathbf{A}}{k}$ we have that $\frac{\mathbf{A}}{k}\mathbf{x} = \mathbf{x}$ which implies that $\mathbf{A}\mathbf{x} = k\mathbf{x} = \rho(\mathbf{A})\mathbf{x}$. Taking the absolute value we get, $|\mathbf{A}\mathbf{x}| = |\rho(\mathbf{A})\mathbf{x}|$ and it follows that $\mathbf{A}|\mathbf{x}| = \rho(\mathbf{A})|\mathbf{x}|$. □

The above proof of positive eigenpair yielded the following important result: $\rho(\mathbf{A})$ is an eigenvalue for $\mathbf{A} > 0$.

Remark II.4. *In the previous proofs we have used $\rho(\mathbf{A}) = 1$. In light of Proposition I.9 (iii), in further proofs we can assume without loss of generality that this is true.*

2.3 Index of $\rho(\mathbf{A})$

Definition II.5. [3, p.510] If $\mathbf{A}$ is a square matrix then λ is a *semi-simple* eigenvalue if and only if $algmult_{\mathbf{A}}(\lambda) = geomult_{\mathbf{A}}(\lambda)$. An eigenvalue is called *simple* if $algmult_{\mathbf{A}}(\lambda) = 1$.

Lemma II.6. [3, p.664] *For nonzero vectors $\{\mathbf{z}_1, \ldots, \mathbf{z}_n\}$ it follows that $||\sum_j \mathbf{z}_j||_2 = \sum_j ||\mathbf{z}_j||_2 \iff \mathbf{z}_j = \alpha_j \mathbf{z}_1$ where $\alpha_j > 0$.*

Theorem II.7. [3, p.664] *Let $\mathbf{A} > 0$ be an $n \times n$ matrix. Then, the following statements are true.*

(i) $\rho(\mathbf{A})$ is the only eigenvalue of $\mathbf{A}$ on the spectral circle. That is, $|\lambda| = \rho(\mathbf{A})$ then $\lambda = \rho(\mathbf{A})$.

(ii) $index(\rho(\mathbf{A})) = 1$. In other words, $\rho(\mathbf{A})$ is a semi-simple eigenvalue.

Proof.

(i) Assume without loss of generality that $\rho(\mathbf{A}) = 1$. Then $(\lambda, \mathbf{x})$ is an eigenpair for $\mathbf{A}$ such that $|\lambda| = 1$. From Theorem II.3 (ii), we know that $\mathbf{0} < |\mathbf{x}| \leq \mathbf{A}|\mathbf{x}|$. Consider the kth entry such that $0 < |x_k| = (\mathbf{A}|\mathbf{x}|)_k$. Let $j = 1, 2, \ldots, n$. We know that $(\mathbf{A}|\mathbf{x}|)_k = \sum_{j=1}^{n} a_{kj}|x_j|$. Also,

$$|x_k| = |\lambda||x_k|$$

$$= |(\lambda \mathbf{x})_k|$$

$$= |(\mathbf{Ax})_k|$$

$$= \left| \sum_{j=1}^{n} a_{kj} x_j \right|$$

Therefore,

$$\left| \sum_{j=1}^{n} a_{kj} x_j \right| = \sum_{j=1}^{n} a_{kj} |x_j| = \sum_{j=1}^{n} |a_{kj} x_j|.$$

From Lemma II.6 we have that, $a_{kj}x_j = \alpha_j(a_{k1}x_1)$ for each j. We know $a_{kj} > 0$ so it follows that,

$$a_{kj}x_j = \alpha_j(a_{k1}x_1)$$

$$x_j = \frac{\alpha_j(a_{k1}x_1)}{a_{kj}}$$

$$x_j = \frac{\alpha_j a_{k1}}{a_{kj}} x_1$$

Let $\pi_j = \dfrac{\alpha_j(a_{k1})}{a_{kj}} > 0$ then we have that $x_j = \pi_j x_1$. Let $\mathbf{x} = x_1\mathbf{p}$ where $\mathbf{p} = (1, \pi_2, \ldots, \pi_n)^T > 0$. Then, $\mathbf{Ax} = \lambda\mathbf{x} \Longrightarrow \mathbf{A}(x_1\mathbf{p}) = \lambda(x_1\mathbf{p}) \Longrightarrow \mathbf{Ap} = \lambda\mathbf{p}$. Then, $\lambda\mathbf{p} = \mathbf{Ap} = |\mathbf{Ap}| = |\lambda\mathbf{p}| = |\lambda|\mathbf{p} = \mathbf{p}$. So we have that $\lambda\mathbf{p} = \mathbf{p}$. Setting $\lambda\mathbf{p} - \mathbf{p} = 0$ we get that $\mathbf{p}(\lambda - 1) = 0$. We know that $\mathbf{p} > 0$ so $\lambda - 1 = 0$ which implies $\lambda = 1$. Therefore, $\lambda = 1$ is the only eigenvalue of $\mathbf{A}$ on the spectral circle.

(ii) We want to show that $index(1) = 1$. Suppose on the contrary, $index(1) = m > 1$. That is, the largest Jordan block for $\mathbf{J}_*(1)$ is of m x m size. We know that $\lambda = 1$ so from Theorem I.2 (*1.3*) we have that $||\mathbf{J}^k||_\infty \geq ||\mathbf{J}_*^k||_\infty$ so each individual Jordan block, $||\mathbf{J}_*^k||_\infty \longrightarrow \infty$. Since $||\mathbf{J}^k||_\infty \geq k$ for all k we have that $||\mathbf{J}^k||_\infty \longrightarrow \infty$ as $k \longrightarrow \infty$. From Theorem I.2 we know that $\mathbf{J}^k = \mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}$ and since $||\mathbf{J}^k||_\infty \longrightarrow \infty$

it directly follows that $||\mathbf{A}^k||_\infty \longrightarrow \infty$. Indeed,

$$\mathbf{J}^k = \mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}$$

$$||\mathbf{J}^k||_\infty = ||\mathbf{P}^{-1}\mathbf{A}^k\mathbf{P}||_\infty$$

$$||\mathbf{J}^k||_\infty \leq ||\mathbf{P}^{-1}||_\infty||\mathbf{A}^k||_\infty||\mathbf{P}||_\infty$$

Since $\mathbf{P}$ is invertible $\mathbf{P} \neq 0$ it follows that, $||\mathbf{P}||_\infty \neq 0$ and $||\mathbf{P}^{-1}||_\infty \neq 0$. Therefore,

$$||\mathbf{A}^k||_\infty \leq \frac{||\mathbf{J}^k||_\infty}{||\mathbf{P}^{-1}||_\infty||\mathbf{P}||_\infty} \longrightarrow \infty.$$

Let $\mathbf{A}^k = [a_{ij}^{(k)}]$ where $[a_{ij}^{(k)}]$ represents (i, j) entries of $\mathbf{A}^k$. Then, $||\mathbf{A}^k||_\infty = \sum_j |a_{i_k j}^{(k)}|$ for some $1 \leq i_k \leq n$. From above, we know that there exists a vector $\mathbf{p} > 0$ such that $\mathbf{p} = \mathbf{A}\mathbf{p}$. Continually multiplying both sides by $\mathbf{A}$ we eventually get $\mathbf{p} = \mathbf{A}^k\mathbf{p}$. Then,

$$\begin{aligned} ||\mathbf{p}||_\infty &\geq p_{i_k} \\ &= \sum_j a_{i_k j}^{(k)} p_j \\ &\geq \Big(\sum_j a_{i_k j}^{(k)}\Big)(\min_i p_i) \\ &= ||\mathbf{A}^k||_\infty(\min_i p_i) \longrightarrow \infty \end{aligned}$$

Since $\mathbf{p}$ was a constant vector it cannot approach ∞ so, we have a contradiction. Therefore, $index(1) = 1$. If the $index(1) = 1$ we want to show that $\lambda = 1$ is a semi-simple eigenvalue. Since $index(1) = 1$ from Remark I.3 it directly follows that $algmult_{\mathbf{A}}(1) = geomult_{\mathbf{A}}(1) = 1$ as desired. □

2.4 Multiplicities of $\rho(\mathbf{A})$

Theorem II.8. [3, p.664] *If* $\mathbf{A} > 0$*, then* $algmult_{\mathbf{A}}(\rho(\mathbf{A})) = 1$*. In other words, the spectral radius of* $\mathbf{A}$ *is a simple eigenvalue of* $\mathbf{A}$*. Therefore,* $dimN(\mathbf{A} - \rho(\mathbf{A})\mathbf{I}) = geomult_{\mathbf{A}}\rho(\mathbf{A}) = algmult_{\mathbf{A}}\rho(\mathbf{A}) = 1$*.*

Proof. Assume without loss of generality that $\rho(\mathbf{A}) = 1$. We want to show that $algmult_{\mathbf{A}}(1) = 1$. Suppose on the contrary, $algmult_{\mathbf{A}}(1) = m > 1$. From Theorem II.7 (ii) we know $\lambda = 1$ is a semi-simple eigenvalue so it follows that $algmult_{\mathbf{A}}(1) = geomult_{\mathbf{A}}(1)$. Since $geomult_{\mathbf{A}}(1) = m > 1$, we must have $m \geq 2$ which means $dimN(\mathbf{A} - 1\mathbf{I}) \geq 2$. Thus, there exists at least two linearly independent eigenvectors say, $(\mathbf{x}, \mathbf{y})$ such that $\mathbf{x} \neq \alpha\mathbf{y}$ for all $\alpha \in \mathbb{C}$. Select a non-zero component from $\mathbf{y}$, say $y_i \neq 0$ and let $\mathbf{z} = \mathbf{x} - \left(\frac{x_i}{y_i}\right)\mathbf{y}$. Since $\mathbf{x}$ and $\mathbf{y}$ are linearly independent eigenvectors we have that $\mathbf{Ax} = \mathbf{x}$ and $\mathbf{Ay} = \mathbf{y}$ so, $\mathbf{Az} = \mathbf{Ax} - \left(\frac{x_i}{y_i}\right)\mathbf{Ay} = \mathbf{x} - \left(\frac{x_i}{y_i}\right)\mathbf{y}$ which implies that $\mathbf{Az} = \mathbf{z}$. From Theorem II.3 we know that $\mathbf{A}|\mathbf{z}| = |\mathbf{z}| > \mathbf{0}$. Then $z_i = x_i - \left(\frac{x_i}{y_i}\right)y_i = x_i - x_i = 0$. This shows $z_i = 0$, but $|\mathbf{z}| > \mathbf{0}$ so we have a contradiction. Therefore, $m = 1$ and $algmult_{\mathbf{A}}(1) = algmult_{\mathbf{A}}(\rho(\mathbf{A})) = 1$.

Now we want to show that $dimN(\mathbf{A} - 1\mathbf{I}) = geomult_{\mathbf{A}}(1) = algmult_{\mathbf{A}}(1) = 1$. In Theorem II.7 (ii) it was shown that $geomult_{\mathbf{A}}(\lambda) = algmult_{\mathbf{A}}(\lambda)$. From above, we found that that the $algmult_{\mathbf{A}}(1) = 1$ so it directly follows that $geomult_{\mathbf{A}}(1) = 1$. ☐

Since $dimN(\mathbf{A} - \rho(\mathbf{A})\mathbf{I})$ is a one-dimensional space that can be spanned by some $\mathbf{v} > 0$ there exists a unique eigenvector $\mathbf{p} \in N(\mathbf{A} - \rho(\mathbf{A})\mathbf{I})$ such that $\sum_j p_j = 1$. This unique eigenvector is called the *Perron-Frobenius eigenvector* for $\mathbf{A} > 0$ and the eigenvalue $r = \rho(\mathbf{A})$ is called the *Perron-Frobenius eigenvalue* for $\mathbf{A}$. Since $\mathbf{A} > 0 \iff \mathbf{A}^T > 0$ and $\rho(\mathbf{A}) = \rho(\mathbf{A}^T)$ then in addition to the eigenpair $(r, \mathbf{p})$ for $\mathbf{A}$ there exists an eigenpair $(r, \mathbf{q})$ for $\mathbf{A}^T$. Since $\mathbf{q}^T\mathbf{A} = r\mathbf{q}^T$ the vector, $\mathbf{q}^T > 0$, is called the *left-hand Perron-Frobenius eigenvector* for $\mathbf{A}$.

Proposition II.9. [3, p.666] *Although eigenvalues of* $\mathbf{A} > 0$ *other than* $\rho(\mathbf{A})$ *may or may not be positive there are no non-negative eigenvectors for* $\mathbf{A} > 0$ *other than the Perron vector* $\mathbf{p}$ *and its positive multiples.*

Proof. Let $r = \rho(\mathbf{A})$ and $(\lambda, \mathbf{y})$ be an eigenpair for $\mathbf{A}$ such that $\mathbf{y} \geq 0$. Given $(r, \mathbf{q})$ where $\mathbf{q}$ is the left-hand Perron-Frobenius eigenvector for $\mathbf{A}$ we know that $r\mathbf{q} = \mathbf{A}^T\mathbf{q}$. Since $\mathbf{q}^T\mathbf{y} > 0$ we get,

$$(r\mathbf{q})^T = (\mathbf{A}^T\mathbf{q})^T$$

$$r\mathbf{q}^T = \mathbf{q}^T\mathbf{A}$$

$$r\mathbf{q}^T\mathbf{y} = (\mathbf{q}^T\mathbf{A})\mathbf{y}$$

$$r\mathbf{q}^T\mathbf{y} = \mathbf{q}^T(\mathbf{A}\mathbf{y})$$

$$r\mathbf{q}^T\mathbf{y} = \lambda\mathbf{q}^T\mathbf{y}$$

$$r = \lambda$$

Therefore, $r = \rho(\mathbf{A}) = \lambda$. From Theorem II.8 we know the eigenspace for the Perron-Frobenius eigenvalue of r is one-dimensional so $\mathbf{p}$ can be written as $\mathbf{p} = \alpha\mathbf{y}$ for some α. Since $\mathbf{p} > 0$ and $\mathbf{y} \geq 0$ we must have $\alpha > 0$. Then, $\mathbf{y} = \frac{1}{\alpha}\mathbf{p}$ and therefore, $\mathbf{y}$ is a positive multiple of the Perron-Frobenius one. □

2.5 Collaltz-Wielandt Formula

Theorem II.10. [3, p.666] *The Perron eigenvalue of* $\mathbf{A} > 0$ *is given by* $r = \max_{\mathbf{x} \in \mathcal{N}} f(\mathbf{x})$, *where*

$$f(\mathbf{x}) = \min_{\substack{1 \leq i \leq n \\ x_i \neq 0}} \frac{[\mathbf{A}\mathbf{x}]_i}{x_i} \text{ and } \mathcal{N} = \{\mathbf{x} : \mathbf{x} \geq 0 \text{ with } \mathbf{x} \neq 0\}.$$

Proof. Let $\beta = f(\mathbf{x})$ for some arbitrary $\mathbf{x} \in \mathcal{N}$. Since $f(\mathbf{x}) \leq \frac{[\mathbf{Ax}]_i}{x_i}$ for each i we have that $\beta \leq \frac{[\mathbf{Ax}]_i}{x_i}$ for each i so, $\beta x_i \leq [\mathbf{Ax}]_i$ which implies that $\mathbf{0} \leq \beta\mathbf{x} \leq \mathbf{Ax}$. Let $\mathbf{p}$ and $\mathbf{q}^T$ be the respective right-hand and left-hand Perron-Frobenius eigenvectors for $\mathbf{A}$ associated with the Perron-Frobenius eigenvalue r. Then,

$$\beta\mathbf{x} \leq \mathbf{Ax}$$

$$\beta\mathbf{q}^T\mathbf{x} \leq (\mathbf{q}^T\mathbf{A})\mathbf{x}$$

$$\beta\mathbf{q}^T\mathbf{x} \leq r\mathbf{q}^T\mathbf{x}$$

$$\beta \leq r$$

Therefore, $f(\mathbf{x}) \leq r$ for all $\mathbf{x} \in \mathcal{N}$. We know $\max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x}) \leq r$. We want to show that in fact $r = \max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x})$. Indeed,

$$\begin{aligned} f(\mathbf{p}) &= \min_{1\leq i\leq n} \frac{[\mathbf{Ap}]_i}{p_i} \\ &= \min_{1\leq i\leq n} \frac{[\mathbf{rp}]_i}{p_i} \\ &= r \min_{1\leq i\leq n} \frac{p_i}{p_i} \\ &= r \end{aligned}$$

Since $f(\mathbf{p}) = r$ for $\mathbf{p} \in \mathcal{N}$ it follows that, $r = \max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x})$. $\square$

CHAPTER III

Perron Frobenius Theorem for Non-negative Matrices

3.1 Non-negative Eigenpair

A standard fact to note about spectral radius is that $\rho(\mathbf{A}) = \lim_{k\longrightarrow\infty} ||\mathbf{A}^k||^{1/k}$ [3, p.619]. We will use this result to prove the following Lemma.

Lemma III.1. [3, p.619] *If* $|\mathbf{A}| \leq \mathbf{B}$ *then* $\rho(\mathbf{A}) \leq \rho(|\mathbf{A}|) \leq \rho(\mathbf{B})$, *where* $|\mathbf{A}|$ *denotes the matrix having entries* $|a_{ij}|$ *and define* $\mathbf{B} \leq \mathbf{C}$ *to mean* $b_{ij} \leq c_{ij}$ *for each* i *and* j.

Proof. Using the triangle inequality we get that, $|\mathbf{A}^k| \leq |\mathbf{A}|^k$ for every positive integer k. So, $|\mathbf{A}| \leq \mathbf{B}$ implies that $|\mathbf{A}|^k \leq \mathbf{B}^k$. Using our result from above we get, $||\mathbf{A}^k||_\infty = |||\mathbf{A}^k|||_\infty \leq |||\mathbf{A}|^k||_\infty \leq ||\mathbf{B}^k||_\infty$. Then,

$$||\mathbf{A}^k||_\infty^{1/k} \leq |||\mathbf{A}|^k||_\infty^{1/k} \leq ||\mathbf{B}^k||_\infty^{1/k}$$

$$\lim_{k\longrightarrow\infty} ||\mathbf{A}^k||_\infty^{1/k} \leq \lim_{k\longrightarrow\infty} |||\mathbf{A}|^k||_\infty^{1/k} \leq \lim_{k\longrightarrow\infty} ||\mathbf{B}^k||_\infty^{1/k}$$

$$\rho(\mathbf{A}) \leq \rho(|\mathbf{A}|) \leq \rho(\mathbf{B})$$

□

Theorem III.2. [3, p.670] *Let $\mathbf{A} \geq 0$ be an $n \times n$ matrix with $r = \rho(\mathbf{A})$, the following statements are true.*

(i) $r \in \sigma(\mathbf{A})$, *but* $r = 0$ *is possible.*

(ii) $\mathbf{Az} = r\mathbf{z}$ *for some* $\mathbf{z} \in \mathcal{N} = \{\mathbf{x} | \mathbf{x} \geq 0 \text{ with } \mathbf{x} \neq 0\}$.

(iii) $r = \max_{\mathbf{x} \in \mathcal{N}} f(\mathbf{x}), where f(\mathbf{x}) = \min_{\substack{1 \leq i \leq n \\ x_i \neq 0}} \frac{[\mathbf{Ax}]_i}{x_i}$ *(ie.the Collatz-Wieldant Formula remains valid).*

Proof. Let $\mathbf{A}_k = \mathbf{A} + \left(\frac{1}{k}\right)\mathbf{E} > 0$ where $\mathbf{E}$ is a matrix containing all 1's and $k = 1, 2, \ldots, n$. Then it follows that $\mathbf{A}_k > 0$ since $\mathbf{A} \geq 0, k > 0$ and $\mathbf{E} > 0$. Let $r_k > 0$ and $\mathbf{p}_k > 0$ be the Perron-Frobenius eigenvalue and the Perron-Frobenius eigenvector corresponding to $\mathbf{A}_k$, respectively. We know $\{\mathbf{p}_k\}_{k=1}^{\infty}$ is bounded since $\{\mathbf{p}_k\}_{k=1}^{\infty} \subseteq \{\mathbf{x} \in \mathbb{R} : ||\mathbf{x}||_1 = 1\}$. The Bolanzo-Wierstrass Theorem states that every bounded sequence has a convergent subsequence. Thus, since $\{\mathbf{p}_k\}_{k=1}^{\infty}$ is bounded, a convergent subsequence exists. Let $\mathbf{z}$ be an arbitrary point in $\mathbb{R}^n$. We have that $\{p_{k_i}\}_{i=1}^{\infty} \longrightarrow \mathbf{z}$ for some increasing sequence k_i where $\mathbf{z} \geq 0$ and $\mathbf{z} \neq 0$ since $p_{k_i} > 0$ and $||p_{k_i}||_1 = \sum_{i=1}^{n} |p_{k_i}| = 1$. Hence, since $\mathbf{A_1} > \mathbf{A_2} > \ldots > \mathbf{A}$ from Lemma III.1 it is clear that $r_1 \geq r_2 \geq \ldots \geq r$ so $\{\mathbf{r}_k\}_{k=1}^{\infty}$ is a decreasing sequence of positive numbers bounded below by r. In other words, $\lim_{k \longrightarrow \infty} r_k = r^*$ exists and $r^* \geq r$. In particular, $\lim_{i \longrightarrow \infty} r_{k_i} = r^* \geq r$. But, from our definition of $\mathbf{A}_k$ we have $\lim_{k \longrightarrow \infty} \mathbf{A}_k = \mathbf{A}$ which implies that $\lim_{i \longrightarrow \infty} \mathbf{A}_{k_i} \longrightarrow \mathbf{A}$ so it is also true that,

$$\begin{aligned}
\mathbf{Az} &= \lim_{i \longrightarrow \infty} \mathbf{A}_{k_i} \lim_{i \longrightarrow \infty} p_{k_i} \\
&= \lim_{i \longrightarrow \infty} \mathbf{A}_{k_i} p_{k_i} \\
&= \lim_{i \longrightarrow \infty} r_{k_i} p_{k_i} \\
&= \lim_{i \longrightarrow \infty} r_{k_i} \lim_{i \longrightarrow \infty} p_{k_i}
\end{aligned}$$

$$= r^* \mathbf{z}$$

Since $\mathbf{Az} = r^*\mathbf{z}$ we have that $r^* \in \sigma(\mathbf{A})$ and therefore, $r^* \leq r$. Since $r = \rho(\mathbf{A})$, we know r^* will always be less than r as it is not the maximum eigenvalue so therefore, $r^* = r$.

Let $\mathbf{q}_k{}^T$ be the left-hand Perron-Frobenius eigenvector for $\mathbf{A}_k$. From Proposition II.9 we know that for every $x \in \mathcal{N}$ and $k > 0$ we have that $\mathbf{q}_k{}^T > 0$. Then it follows that, $0 \leq f(\mathbf{x})\mathbf{x} \leq \mathbf{Ax} \leq \mathbf{A}_k\mathbf{x}$. Indeed,

$$\mathbf{q}_k{}^T[f(\mathbf{x})\mathbf{x}] \leq \mathbf{q}_k{}^T(\mathbf{A}_k\mathbf{x})$$

$$f(\mathbf{x})\mathbf{q}_k{}^T\mathbf{x} \leq (\mathbf{q}_k{}^T\mathbf{A}_k)\mathbf{x}$$

$$f(\mathbf{x})\mathbf{q}_k{}^T\mathbf{x} = r_k\mathbf{q}_k{}^T\mathbf{x}$$

$$f(\mathbf{x}) \leq r_k$$

Since $r_k \longrightarrow r^* = r$ we have that $f(\mathbf{x}) \leq r$. Since $f(\mathbf{z}) = r$ and $\mathbf{z} \in \mathcal{N}$ it directly follows that $\max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x}) = r$. Therefore, the Collatz-Wieldant formula is satisfied. □

3.2 Reducibility

Definition III.3. [3, p.671] A matrix $\mathbf{A} \geq 0$ is said to be a reducible matrix when there exists a permutation matrix $\mathbf{P}$ such that,

$$\mathbf{P}^T\mathbf{AP} = \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{0} & \mathbf{Z} \end{pmatrix}$$

where $\mathbf{X}$ and $\mathbf{Z}$ are both square. Otherwise, $\mathbf{A}$ is said to be an irreducible matrix. To determine whether a matrix is irreducible it is useful to use the following fact: a matrix $\mathbf{A}$ is irreducible if and only if $G(\mathbf{A})$ is strongly connected [3, p. 671].

Example III.4. The matrix $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ is reducible because we are able to find a permutation matrix $\mathbf{P}$ such that $\mathbf{P}^T\mathbf{AP} = \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{0} & \mathbf{Z} \end{pmatrix}$. Namely, $\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and $\mathbf{P}^T\mathbf{AP} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. We notice that the matrix $\mathbf{A}$ gives the graph $G(\mathbf{A})$ as seen in Figure (1.4) on page 6. We also notice that the graph is not strongly connected since there does not exist a path from v_1 to v_2. Hence, this matrix is not irreducible and therefore, it must be reducible. On the other hand, the graph in Figure (1.5) on page 6 is strongly connected which implies its corresponding adjacency matrix is in fact irreducible.

Remark III.5. *A matrix* $\mathbf{A} = \begin{pmatrix} * & * \\ * & 0 \end{pmatrix}$ *where the $*$'s are non-negative and non-zero numbers, is reducible since $G(\mathbf{A})$ is the same graph as in Figure* (1.4) *and the discussion in Example III.4 above will follow.*

The following Lemma shows a simple process that allows us to convert non-negative irreducible matrices into a positive matrix.

Lemma III.6. [1, p.192] Converting Non-negativity and Irreducibility to Positivity
Let $\mathbf{A} \geq 0$ be an $n \times n$ irreducible matrix then, $(\mathbf{I}+\mathbf{A})^{n-1} > 0$ where $\mathbf{I}$ is the identity matrix.

Proof. From Lemma I.8, it follows that if $(\mathbf{A}^k)_{ij} > 0$ if and only if there exists a path of length k from v_i to v_j in $G(\mathbf{A})$. If there exists a path of length at most k between every pair of vertices, then $\sum_{L=0}^{k} \mathbf{A}^L > 0$.

Since $G(\mathbf{A})$ has n vertices, it follows from a simple induction proof with respect to n that between any two pairs of vertices there exists a path of length at most

$n-1$. This implies that $\sum_{L=0}^{n-1} \mathbf{A}^L > 0$, by Lemma I.8. Then, for any positive integers $c_0, c_1, \ldots, c_{n-1}$ it follows that $\sum_{L=0}^{n-1} \mathbf{A}^L > 0$ if and only if $\sum_{L=0}^{n-1} c_L \mathbf{A}^L > 0$. Let $\mathbf{I}$ be the identity matrix. Since $\mathbf{A}$ and $\mathbf{I}$ commute the Binomial Theorem holds. Indeed,

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

Choose $c_L = \binom{n-1}{L}$. Since $\sum_{L=0}^{n-1} \mathbf{A}^L > 0$, we have that $(\mathbf{A}+\mathbf{I})^{n-1} > 0$. □

Definition III.7. [3, p.600] Function of Jordan Blocks

For a $k \times k$ Jordan block, $\mathbf{J}_*$, with eigenvalue λ, and for a function $f : \mathbb{C} \longrightarrow \mathbb{C}$ is said to be defined (or to exist) at $\mathbf{A}$ when $f(\lambda), f'(\lambda), \ldots, f^{(k-1)}(\lambda)$ exist such that, $f(\mathbf{J}_*)$ is defined to be

$$f(\mathbf{J}_*) = f\begin{pmatrix} \lambda & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \ddots & 1 \\ 0 & \cdots & \cdots & \lambda \end{pmatrix} = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(k-1)}(\lambda)}{(k-1)!} \\ & f(\lambda) & f'(\lambda) & \ddots & \vdots \\ & & \ddots & \ddots & \frac{f''(\lambda)}{2!} \\ & & & f(\lambda) & f'(\lambda) \\ & & & & f(\lambda) \end{pmatrix}.$$

Definition III.8. [3, p.601] Function of a Matrix

Suppose that $\mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1}$, where $\mathbf{J} = \begin{pmatrix} \ddots & & \\ & J_*(\lambda_j) & \\ & & \ddots \end{pmatrix}$ is in Jordan form. If f exists at $\mathbf{A}$ then the value of f at $\mathbf{A}$ is defined to be

$$f(\mathbf{A}) = \mathbf{P} f(\mathbf{J}) \mathbf{P}^{-1} = \mathbf{P} \begin{pmatrix} \ddots & & \\ & f(\mathbf{J}_*(\lambda_j)) & \\ & & \ddots \end{pmatrix} \mathbf{P}^{-1}.$$

Remark III.9. [3, p.601] *The matrix function $f(\mathbf{A})$ produces a uniquely defined matrix.*

Based on the definitions above, if f exists at $\mathbf{A}$, then $f(\mathbf{A}) = \mathbf{P}f(\mathbf{J})\mathbf{P}^{-1}$. Thus, if $(\lambda, \mathbf{x})$ is an eigenpair for $\mathbf{A}$, it follows that $(f(\lambda), \mathbf{x})$ is an eigenpair for $f(\mathbf{A})$. This result will be used in the following proof.

Theorem III.10. [3, p.673] *Let $\mathbf{A} \geq 0$ be an $n \times n$ irreducible matrix, then the following statements are true.*

(i) $r = \rho(\mathbf{A}) \in \sigma(\mathbf{A})$ *and* $r > 0$.

(ii) $algmult_{\mathbf{A}}(r) = 1$.

(iii) *There exists an eigenvector* $\mathbf{x} > 0$ *such that* $\mathbf{A}\mathbf{x} = r\mathbf{x}$.

(iv) *The unique vector defined by*

$$\mathbf{A}\mathbf{p} = r\mathbf{p}, \mathbf{p} > 0, \text{ and } ||\mathbf{p}||_1 = 1$$

is called the **Perron vector**. *There are no other non-negative eigenvectors for* $\mathbf{A}$ *except for positive multiples of* $\mathbf{p}$, *regardless of the eigenvalue.*

(v) *The Collatz-Wieldant formula* $r = \max_{\mathbf{x} \in \mathcal{N}} f(\mathbf{x})$, *where*

$$f(\mathbf{x}) = \min_{\substack{1 \leq i \leq n \\ x_i \neq 0}} \frac{[\mathbf{A}\mathbf{x}]_i}{x_i} \text{ and } \mathcal{N} = \{\mathbf{x} : \mathbf{x} \geq 0 \text{ with } \mathbf{x} \neq 0\}.$$

Proof. From Theorem III.2 we know that $r = \rho(\mathbf{A}) \in \sigma(\mathbf{A})$ which shows statement (i) is true. Now we want to show that the $algmult_{\mathbf{A}}(r) = 1$ is also true. Since $\mathbf{A}$ is irreducible we will let $\mathbf{B} = (\mathbf{I} + \mathbf{A})^{n-1} > 0$ be our matrix of size $n \times n$. Let $f(\mathbf{A}) = \mathbf{B}$. From Theorem III.8 since f is differentiable it follows that

$f(x) = (1+x)^{n-1}$. Then, $\lambda \in \sigma(\mathbf{A})$ if and only if $(1+\lambda)^{n-1} \in \sigma(\mathbf{B})$. From Remark I.3 we know that in terms of Jordan blocks the $algmult_{\mathbf{A}}(\lambda)$ is equal to the sum of the sizes of all Jordan blocks associated with λ. Then the size of the Jordan blocks corresponding to $f(\lambda)$ will be the same size as our matrix function $f(\mathbf{A})$ so we have that $algmult_{\mathbf{A}}(\lambda) = algmult_{\mathbf{B}}((1+\lambda)^{n-1})$. If $u = \rho(\mathbf{B})$ then by definition, $u = \max_{\lambda \in \sigma(\mathbf{B})} |u| = \max_{\lambda \in \sigma(\mathbf{A})} |(1+\lambda)|^{n-1} = \left(\max_{\lambda \in \sigma(\mathbf{A})} |(1+\lambda)| \right)^{n-1} = (1+r)^{n-1}$ since $|1+\lambda| \leq 1 + |\lambda| \leq 1 + r$ where r is an eigenvalue of $\mathbf{A}$. Since r is an eigenvalue of $\mathbf{A}$ from Theorem II.8 we have that$algmult_{\mathbf{A}}(r) = 1$ as desired.

From Theorem III.2 we know there exists a non-negative eigenvector $\mathbf{x} \geq 0$ associated with our eigenvalue r. Then we can find a corresponding eigenvector for our eigenvalue r. So if $(\lambda, \mathbf{x})$ is an eigenpair for $\mathbf{A}$ then it follows that $(f(\lambda), \mathbf{x})$ is an eigenpair for $f(\mathbf{A})$. In our case, since $(r, \mathbf{x})$ is an eigenpair for $\mathbf{A}$ implies that $(u, \mathbf{x})$ is an eigenpair for $\mathbf{B}$. The Perron-Frobenius Theorem for Positive Matrices ensures that $\mathbf{x}$ must be a positive multiple of the Perron vector $\mathbf{B}$ which implies that $\mathbf{x} > 0$. We want to show that $\mathbf{Ax} = r\mathbf{x}$. If $r = 0$ then $\mathbf{Ax} = 0$ which is not possible since $\mathbf{A} \neq 0$ and is non-negative and $\mathbf{x} > 0$. Therefore, $r > 0$ and we obtain our result for (iii).

To show $\mathbf{Ap} = r\mathbf{p}$ we simply scale our eigenvector $\mathbf{x}$ by $\dfrac{\mathbf{x}}{||\mathbf{x}||_1}$ to get our result for (iv). Part (v) was proven in Theorem III.2 (iii). □

3.3 Primitive Matrices

Definition III.11. [3, p.674] A non-negative irreducible matrix $\mathbf{A}$ having only one eigenvalue, $r = \rho(\mathbf{A})$, on its spectral circle is said to be a primitive matrix. Equivalently, a matrix is called primitive if there exists a positive integer k such that $\mathbf{A}^k$ is a positive matrix [1, p.198].

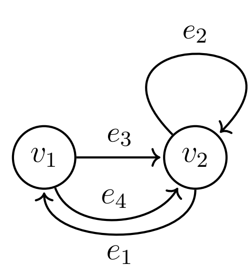
Example III.12. The matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ is primitive since for any $k > 0$ it follows that $\mathbf{A}^k > 0$. For instance, $\mathbf{A}^2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} > 0$. On the contrary, the matrix $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is irreducible, but not primitive since its eigenvalues are ± 1 which both lie on the unit circle.

Remark III.13. [1, p.202] *Every primitive matrix is irreducible, but not every irreducible matrix is primitive. We can tell whether a matrix is primitive by looking at the greatest common divisor of the cycles within the graph.*

Theorem III.14. [1, p.202] *A non-negative matrix* $\mathbf{A}$ *is primitive if and only if* $G(\mathbf{A})$ *is strongly connected and has two relatively prime cycle lengths.*

Corollary III.15. [1, p.203] *A non-negative matrix is primitive if and only if the corresponding graph is strongly connected and the gcd of its simple cycles is 1. We call such a graph a primitive graph.*

Example III.16. The matrix $\mathbf{A} = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$ is primitive with eigenvalues 2 and -1, but the matrix $\mathbf{B} = \begin{pmatrix} 0 & 4 \\ 1 & 0 \end{pmatrix}$ with eigenvalues 2, and -2 is not primitive. If you were to graph matrices $\mathbf{A}$ and $\mathbf{B}$ it would be clear that matrix $\mathbf{A}$ is the only non-negative primitive matrix. The graph of matrix $\mathbf{A}$ is shown below.



(3.1)

From Theorem III.14, for matrix $\mathbf{A}$ to be primitive its corresponding graph must be strongly connected and have two relatively prime cycle lengths. It is evident the graph below is strongly connected as there exists a path from v_1 to v_2 and vise versa. In addition, the graph has two relatively prime cycles; the first cycle starting at v_2 and travelling along e_1 to e_3 back to v_2 and the second cycle starting at v_1 travelling along e_3 to e_2 to e_1 and then back to v_1. The first cycle is of length 2, the second is of length 3 which are two relatively prime cycles.

Definition III.17. Projection [3, p.385]

Let X and Y be vector spaces over $\mathbb{C}$. Suppose $V = X \oplus Y$. Then, for each $\mathbf{v} \in V$ there are unique vectors $\mathbf{x} \in X$ and $\mathbf{y} \in Y$ such that $V = X + Y$.

(i) The vector $\mathbf{x}$ is called the projection of $\mathbf{v}$ onto X along Y.

(ii) The vector $\mathbf{y}$ is called the projection of $\mathbf{v}$ onto Y along X.

Theorem III.18. Projector [3, p.386]

Let X and Y be vector spaces over $\mathbb{C}$ and let $\mathbf{v}$ be in vector space V. The unique linear operator $\mathbf{P}$ defined by $\mathbf{Pv} = \mathbf{x}$ is called the projector onto X along Y and $\mathbf{P}$ has the following properties:

(i) $\mathbf{P}^2 = \mathbf{P}$

(ii) $R(\mathbf{P}) = \{\mathbf{x} | \mathbf{Px} = \mathbf{x}\}$

(iii) $R(\mathbf{P}) = N(\mathbf{I} - \mathbf{P}) = X$ *and* $R(\mathbf{I} - \mathbf{P}) = N(\mathbf{P}) = Y$

Theorem III.19. [3, p.339] *A linear projector $\mathbf{P}$ on V is a projector if and only if $\mathbf{P}^2 = \mathbf{P}$.*

We showed in Proposition I.9 that if $\rho(\mathbf{A}) < 1$ then $\lim_{k \longrightarrow \infty} \mathbf{A}^k$ exists. The next result gives necessary and sufficient conditions for the existence of $\lim_{k \longrightarrow \infty} \mathbf{A}^k$.

Lemma III.20. [3, p.630] *For* $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\lim_{k\longrightarrow\infty} \mathbf{A}^k$ *exists if and only if* $\rho(\mathbf{A}) < 1$ *or else* $\rho(\mathbf{A}) = 1$, *where* $\lambda = 1$ *is the only eigenvalue on the unit circle, and* $\lambda = 1$ *is semi-simple. When it exists,* $\lim_{k\longrightarrow\infty} \mathbf{A}^k$ *is equal to the projector onto* $N(\mathbf{I} - \mathbf{A})$ *along* $R(\mathbf{I} - \mathbf{A})$.

Theorem III.21. Primitive Matrices [3, p.674]

Let $\mathbf{A}$ *be a non-negative irreducible matrix and* $r = \rho(\mathbf{A})$. *Then* $\mathbf{A}$ *is primitive if and only if*

$$\lim_{k\longrightarrow\infty} \left(\frac{\mathbf{A}}{r}\right)^k = \mathbf{G} = \frac{\mathbf{p}\mathbf{q}^T}{\mathbf{q}^T\mathbf{p}} > 0$$

where $\mathbf{p}$ *and* $\mathbf{q}$ *are the respective Perron vectors for* $\mathbf{A}$ *and* $\mathbf{A}^T$. *Also,* $\mathbf{G}$ *is the projector onto* $N(\mathbf{A} - r\mathbf{I})$ *along* $R(\mathbf{A} - r\mathbf{I})$.

Proof. By the Perron-Frobenius Theorem for Irreducible Matrices III.10, $1 = \rho\left(\frac{\mathbf{A}}{r}\right)$ is a simple eigenvalue. We know $\mathbf{A}$ is primitive if and only if $\frac{\mathbf{A}}{r}$ is primitive. Matrix $\mathbf{A}$ is primitive if and only if there is a $k > 0$ such that $\mathbf{A}^k > 0$. Then, $\left(\frac{\mathbf{A}}{r}\right)^k = \frac{1}{r}\mathbf{A}^k$. Since $\mathbf{A}^k > 0$ and $r > 0$ it follows that $\frac{\mathbf{A}}{r}$ is also primitive. By definition III.11, $\frac{\mathbf{A}}{r}$ is primitive if and only if $1 = \rho\left(\frac{\mathbf{A}}{r}\right)$ is the only eigenvalue on its spectral circle. By Lemma III.20, we have that $\lim_{k\longrightarrow\infty} \left(\frac{\mathbf{A}}{r}\right)^k$ exists since $1 = \rho\left(\frac{\mathbf{A}}{r}\right)$ is a simple eigenvalue and the only eigenvalue on its spectral circle. In addition, $\lim_{k\longrightarrow\infty} \left(\frac{\mathbf{A}}{r}\right)^k$ is the projector onto $N\left(\mathbf{I} - \frac{\mathbf{A}}{r}\right)$ along $R\left(\mathbf{I} - \frac{\mathbf{A}}{r}\right)$. Then it follows from Lemma III.20 that the $\lim_{k\longrightarrow\infty} \mathbf{A}^k$ exists and is equal to the projector onto $N(\mathbf{I} - \mathbf{A})$ along $R(\mathbf{I} - \mathbf{A})$. Also, $\mathbf{G} = \lim_{k\longrightarrow\infty} \left(\frac{\mathbf{A}}{r}\right)^k = \frac{\mathbf{p}\mathbf{q}^T}{\mathbf{q}^T\mathbf{p}} > \mathbf{0}$ is a projector by Theorem III.19. □

CHAPTER IV

The Leslie Model

In this chapter, we will take all of our previous knowledge and apply it to real world applications. We will begin with a simple population model developed through the Rabbit Problem.

4.1 The Rabbit Problem

In this problem, let A_t represent the number of adult pairs of rabbits at the end of month t and let Y_t be the number of youth pairs of rabbits at the end of month t. To begin, we will start with one pair of youth rabbits. Each youth pair takes two months to mature into adulthood. In this particular model, both adults and youth give birth to a pair at the end of every month, but once a youth pair matures to adulthood and reproduces, it then becomes extinct. The procedure goes as follows:

(1) At the end of Month 0, $A_t = 0$ and $Y_t = 1$.

(2) At the end of Month 1, $A_t = 1$ and $Y_t = 1$.

(3) At the end of Month 2, $A_t = 1$ and $Y_t = 2$. This gives us a total of 3 pairs of rabbits at the end of Month 2.

(4) At the end of Month 3, $A_t = 2$ and $Y_t = 3$. Hence, we have a total of 5 pairs.

(5) At the end of Month 4, $A_t = 3$ and $Y_t = 5$, giving us a total of 8 pairs at the end of Month 4.

The table below is a simple illustration of the above procedure where $P_1, P_2, \ldots, P_n$ represent the first pair of rabbits, the second pair of rabbits, and so on. It is also important to note the underline represents the pair(s) that reproduce a new pair during that specific month.

Month(t)	Pairs	A_t	Y_t	Total Number of Pairs
0	P1	0	1	1
1	<u>P1</u> P2	1	1	2
2	<u>P1</u> P2 P3	1	2	3
3	<u>P1</u> <u>P2</u> P3 P4 P5	2	3	5
4	<u>P1</u> <u>P2</u> <u>P3</u> P4 P5 P6 P7 P8	3	5	8
⋮	⋮	⋮	⋮	⋮

Remark IV.1. *Thus, the total number of rabbit pairs at the end of the n^{th} month is equal to the sum of the number of pairs at the end of the previous two months.*

We can make this model more reasonable by utilizing the age classes through a mathematical model. Indeed,

$$\begin{pmatrix} Y_{t+1} \\ A_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} Y_t \\ A_t \end{pmatrix} \text{ where } \begin{pmatrix} Y_0 \\ A_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

We can write this as,

$$\mathbf{f}(t+1) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \mathbf{f}(t), \text{ where } \mathbf{f}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Hence, the number of youth pairs at the end of month $t+1$ should equal the number of youth pairs at the end of month t plus the number of adult pairs at the end of month t. Also, the number of adults pairs at the end of month $t+1$ is equal to the number of youth pairs at the end of month t.

This model has some unrealistic features such as, every young surviving to adulthood or the same number of offspring being produced by both the adult pair and young pair. To avoid these setbacks we can generalize this model using survival and fertility rates. Since the 1's in the first row represent the number of offspring produced so we can replace these 1's with birth rates b_1 and b_2 where $b_1, b_2 \geq 0$. Since the lower 1 in our matrix represents a youth surviving into adulthood we will replace it by s, which is called the survival rate, with $0 < s \leq 1$. Indeed,

$$\mathbf{f}(t+1) = \begin{pmatrix} b_1 & b_2 \\ s & 0 \end{pmatrix} \mathbf{f}(t).$$

It is evident that this simple model has some features of population such as, the population size may increase, decrease or die out, but it can be improved even further. The Fibonacci Model with two age classes can be generalized to k age classes and we call this the Leslie Model.

4.2 Leslie Model

The Leslie Model is stated as $\mathbf{f}(t+1) = \mathbf{L}\mathbf{f}(t)$ where $\mathbf{f}(t)$ and $\mathbf{f}(t+1)$ are population vectors and $\mathbf{L}$ is our Leslie Matrix. We define $\mathbf{L}$ by,

$$\mathbf{L} = \begin{pmatrix} b_1 & b_2 & \cdots & & b_n \\ s_1 & 0 & \cdots & & 0 \\ & s_2 & & & \\ & & \ddots & & \vdots \\ & & & s_{n-1} & 0 \end{pmatrix}.$$

Each component is the number of individuals of a particular age class such that $\mathbf{f}(t) = (f_1(t), f_2(t), \ldots, f_n(t))^T$. Then at time t, $f_i(t)$ is the number of individuals of age class i. That is, the number of individuals of age class a where $i - 1 \leq a < i$ for $i = 1, 2, \ldots, n$. Let b_k and s_k denote the birth rate and survival rate of $\mathbf{L}$, respectively. That is, the first row of matrix $\mathbf{L}$ consists of birth rates where b_k is the number of off-spring produced by an individual of age class i at time t. The diagonal entries of $\mathbf{L}$ consists of survival rates, where s_k is the probability that an individual in age class i will survive to age class $i + 1$. All other entries of our Leslie matrix are zeroes. For each k, we assume $0 < s_i \leq 1$ and $b_i \geq 0$ and we will assume the width of a particular age class is exactly one time unit.

Suppose you were to divide a population into specific age classes say, $G_1, G_2, \ldots, G_n$ such that each age class covers the same number of years. If $f_k(t)$ is the number of individuals in G_k at time t, then it follows that

$$f_1(t+1) = f_1(t)b_1 + f_2(t)b_2 + \cdots f_n(t)b_n \tag{4.1}$$

and

$$f_k(t+1) = f_{k-1}(t)s_{k-1} \text{ for } k = 2, 3, \ldots, n. \tag{4.2}$$

Matrix equations (4.1) and (4.2) become $\mathbf{f}(t+1) = \mathbf{Lf}(t)$, where $\mathbf{L}$ is our Leslie matrix. We also notice that,

$$F_k(t) = \frac{f_k(t)}{f_1(t) + f_2(t) + \cdots + f_n(t)}$$

is the percentage of population in G_k at time t. The vector $\mathbf{F}(t) = (F_1(t), F_2(t), \ldots, F_n(t))^T$ represents the age distribution at time t.

We will now show that $\mathbf{F}^* = \lim_{t \longrightarrow \infty} \mathbf{F}(t)$ exists and will determine its value. Let

$\mathbf{f}(t+1) = \mathbf{L}\mathbf{f}(t)$ where $\mathbf{L}$ is our Leslie Matrix stated above. The graph of $G(\mathbf{L})$ is strongly connected so from Theorem III.14 we know that $\mathbf{L}$ is primitive. Since $G(\mathbf{L})$ is strongly connected it also follows that $\mathbf{L}$ is irreducible. Then we have that, $r = \rho(\mathbf{L}) > \mathbf{0}$ is an eigenvalue for our matrix, by Theorem 2.1. Also, since $\mathbf{L}$ is primitive we have that $\lim_{k\longrightarrow\infty}\left(\frac{\mathbf{L}}{r}\right)^t = \mathbf{G} = \frac{\mathbf{p}\mathbf{q}^T}{\mathbf{q}^T\mathbf{p}} > \mathbf{0}$, by Theorem III.21. By a simple induction proof it is true that $\mathbf{f}(t) = \mathbf{L}^t\mathbf{f}(0)$. From here if $\mathbf{f}(0) = \mathbf{0}$ we get that,

$$\lim_{t\longrightarrow\infty}\frac{\mathbf{f}(t)}{r^t} = \lim_{t\longrightarrow\infty}\frac{\mathbf{L}^t\mathbf{f}(0)}{r^t} = \lim_{t\longrightarrow\infty}\left(\frac{\mathbf{L}}{r}\right)^t\mathbf{f}(0) = \mathbf{G}\mathbf{f}(0) = \frac{\mathbf{p}\mathbf{q}^T}{\mathbf{q}^T\mathbf{p}}\mathbf{f}(0) = \mathbf{p}\left(\frac{\mathbf{q}^T\mathbf{f}(0)}{\mathbf{q}^T\mathbf{p}}\right).$$

We know that, $F_k(t) = \frac{f_k(t)}{||\mathbf{f}(t)||_1}$ is the percentage of population that is in G_k at time t. Indeed,

$$\mathbf{F}^* = \lim_{t\longrightarrow\infty}\mathbf{F}(t) = \lim_{t\longrightarrow\infty}\frac{\mathbf{f}(t)}{||\mathbf{f}(t)||_1} = \lim_{t\longrightarrow\infty}\frac{\mathbf{f}(t)/r^t}{||\mathbf{f}(t)||_1/r^t} = \frac{\lim_{t\longrightarrow\infty}\mathbf{f}(t)/r^t}{\lim_{t\longrightarrow\infty}||\mathbf{f}(t)||_1/r^t} = \mathbf{p}.$$

It is evident that $\mathbf{f}(t)$ approaches $\mathbf{0}$ if $r < 1$. If $r = 1$ then $\mathbf{f}(t)$ approaches $\mathbf{G}\mathbf{f}(0) = \mathbf{p}\left(\frac{\mathbf{q}^T\mathbf{f}(0)}{\mathbf{q}^T\mathbf{p}}\right) > \mathbf{0}$. Finally, if $r < 1$ the results from our Leslie analysis imply that $f_k(t)$ approaches ∞ for each k.

REFERENCES

[1] Flahive M. Robson R. Cull, P., *Difference equations: from rabbits to chaos*, Springer, 2005.

[2] Leslie Hogben, *Handbook of linear algebra*, CRC Press, 2006.

[3] Carl D. Meyer, *Matrix analysis and applied linear algebra*, vol. 2, Siam, 2000.

[4] M. Tomforde, *The structure of graph C^*-algebras and their generalizations*, vol. 46, Mathematics Subject Classification, 2006.