Proof of the Perron-Frobenius Theorem

Based on exercise 1.20 of Introduction to Stochastic Processes
by Gregory F. Lawler

Doron Shahar

May 1, 2014

Some Definitions

Definition

Let $\vec{u} = (u^1, ..., u^n)$ and $\vec{v} = (v^1, ..., v^n)$ be vectors.

We write $\vec{u} \geq \vec{v}$ if $u^i \geq v^i$ for all i, and $\vec{u} > \vec{v}$ if $u^i > v^i$ for all i.

$|\vec{v}| = (|v^1|, ..., |v^n|)$, and $\vec{0} = (0, ..., 0)$.

Statement of Theorem

Perron-Frobenius Theorem

Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ matrix with $a_{ij} > 0$ forall i, j. $\mathbf{A}$ has a real eigenvalue $\alpha > 0$ with a unique eigenvector $\vec{v} > \vec{0}$ of norm 1 such that $|\lambda| < \alpha$ for any other eigenvalue λ of $\mathbf{A}$. Moreover, α is a simple eigenvalue (i.e., it is a root of the characteristic polynomial of $\mathbf{A}$ with multiplicity 1).

1.20(c) shows that $\alpha > 0$ is an eigenvalue of $\mathbf{A}$.
1.20(d) shows that there is α has a unique eigenvector $\vec{v} \geq \vec{0}$ with norm 1.
1.20(e) shows that, in fact, $\vec{v} > \vec{0}$.
1.20(f) shows that all other eigenvalues λ of $\mathbf{A}$, $|\lambda| < \alpha$.
1.20(i) shows that α is a simple eigenvalue of $\mathbf{A}$.

1.20(j)

Theorem

Every stochastic $n \times n$ matrix $\mathbf{P}$ with positive entries has a unique invariant probability distribution $\vec{\pi}$ with all positive components.

1.20(j)

Theorem

Every stochastic $n \times n$ matrix $\mathbf{P}$ with positive entries has a unique invariant probability distribution $\vec{\pi}$ with all positive components.

Proof:

$\mathbf{P}^T$ is a square matrix with all positive entries. By the Perron-Frobenius theorem, $\mathbf{P}^T$ has a real eigenvalue $\alpha > 0$ with a unique eigenvector $\vec{v} > \vec{0}$ of norm 1. Let $\pi^j = \frac{v^j}{\sum_{i=1}^n v_i}$. Then $\vec{\pi} = (\pi^1, ..., \pi^n) > \vec{0}$ is a probability distribution. Since $\mathbf{P}$ is a stochastic matrix, $\mathbf{P}^T\vec{\pi} = \alpha\vec{\pi}$ is a probability distribution. Therefore, $\alpha = \alpha \sum_{i=1}^n \pi^i = \sum_{i=1}^n \alpha\pi^i = 1$. So $\mathbf{P}^T\vec{\pi} = \vec{\pi}$, and $\vec{\pi}\mathbf{P} = \vec{\pi}$ if we equate $\vec{\pi}$ with $\vec{\pi}^T$. Therefore, $\vec{\pi}$ is an invariant probability distribution of $\mathbf{P}$.

1.20(j)

Theorem

Every stochastic $n \times n$ matrix $\mathbf{P}$ with positive entries has a unique invariant probability distribution $\vec{\pi}$ with all positive components.

Proof:

$\mathbf{P}^T$ is a square matrix with all positive entries. By the Perron-Frobenius theorem, $\mathbf{P}^T$ has a real eigenvalue $\alpha > 0$ with a unique eigenvector $\vec{v} > \vec{0}$ of norm 1. Let $\pi^j = \frac{v^j}{\sum_{i=1}^n v_i}$. Then $\vec{\pi} = (\pi^1, ..., \pi^n) > \vec{0}$ is a probability distribution. Since $\mathbf{P}$ is a stochastic matrix, $\mathbf{P}^T\vec{\pi} = \alpha\vec{\pi}$ is a probability distribution. Therefore, $\alpha = \alpha \sum_{i=1}^n \pi^i = \sum_{i=1}^n \alpha\pi^i = 1$. So $\mathbf{P}^T\vec{\pi} = \vec{\pi}$, and $\vec{\pi}\mathbf{P} = \vec{\pi}$ if we equate $\vec{\pi}$ with $\vec{\pi}^T$. Therefore, $\vec{\pi}$ is an invariant probability distribution of $\mathbf{P}$. If $\vec{\pi}'$ is another invariant probability distribution of $\mathbf{P}$, then $\mathbf{P}^T\vec{\pi}' = \vec{\pi}'$. So $\vec{\pi}'$ is a scalar multiple of $\vec{v}$. As $\vec{\pi}'$ is also a probability distribution, it must equal $\vec{\pi}$. Therefore, $\vec{\pi}$ is the unique invariant probability distribution of $\mathbf{P}$, and $\vec{\pi} > \vec{0}$. QED

Use of the Perron-Frobenius Theorem

Theorem

If $\mathbf{P}$ *be a stochastic matrix with positive entries and a unique invariant probability distribution* $\vec{\pi}$*, then* $\lim_{n\to\infty} \mathbf{P}^n = \begin{pmatrix} \vec{\pi} \\ \vdots \\ \vec{\pi} \end{pmatrix}$

Use of the Perron-Frobenius Theorem

Theorem

If $\mathbf{P}$ *be a stochastic matrix with positive entries and a unique invariant probability distribution* $\vec{\pi}$*, then* $\lim_{n\to\infty} \mathbf{P}^n = \begin{pmatrix} \vec{\pi} \\ \vdots \\ \vec{\pi} \end{pmatrix}$

Proof:

By the Perron-Frobenius theorem and the proof of 1.20(j), $\mathbf{P}^T$ has a simple eigenvalue of 1 and all other eigenvalues of $\mathbf{P}^T$ have absolute value less than 1. The same is true for $\mathbf{P}$.

Therefore, the Jordan canonical form of $\mathbf{P}$ is $\mathbf{J} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{M} & \\ 0 & & & \end{pmatrix}$

where $\mathbf{M}$ has entries with absolute value less than 1 on the diagonal, 1's or 0's on the superdiagonal, and 0's everywhere else. $\lim_{n\to\infty} \mathbf{M}^n = \mathbf{0}$.

Use of the Perron-Frobenius Theorem (continued)

Proof Continued:

$\mathbf{P}$ is similar to its Jordan canonical form. That is, $\mathbf{P} = \mathbf{QJQ}^{-1}$.
$(1, ..., 1)^T = \mathbf{P}(1, ..., 1)^T = \mathbf{QJQ}^{-1}(1, ..., 1)^T = \mathbf{QJ}(1, 0, ..., 0)^T = \mathbf{Q}(1, 0..., 0)^T$.
It follows that the first column of $\mathbf{Q}$ is all 1's.
$\vec{\pi} = \mathbf{P}^T\vec{\pi} = (\mathbf{Q}^{-1})^T\mathbf{J}^T\mathbf{Q}^T\vec{\pi} = (\mathbf{Q}^{-1})^T\mathbf{J}^T(1, 0, ..., 0)^T = (\mathbf{Q}^{-1})^T(1, 0..., 0)^T$.
It follows that the first row of $\mathbf{Q}^{-1}$ is $\vec{\pi}$.

Use of the Perron-Frobenius Theorem (continued)

Proof Continued:

$\mathbf{P}$ is similar to its Jordan canonical form. That is, $\mathbf{P} = \mathbf{QJQ}^{-1}$.
$(1, ..., 1)^T = \mathbf{P}(1, ..., 1)^T = \mathbf{QJQ}^{-1}(1, ..., 1)^T = \mathbf{QJ}(1, 0, ..., 0)^T = \mathbf{Q}(1, 0..., 0)^T$.
It follows that the first column of $\mathbf{Q}$ is all 1's.
$\vec{\pi} = \mathbf{P}^T\vec{\pi} = (\mathbf{Q}^{-1})^T\mathbf{J}^T\mathbf{Q}^T\vec{\pi} = (\mathbf{Q}^{-1})^T\mathbf{J}^T(1, 0, ..., 0)^T = (\mathbf{Q}^{-1})^T(1, 0..., 0)^T$.
It follows that the first row of $\mathbf{Q}^{-1}$ is $\vec{\pi}$. Therefore,

$$\lim_{n\to\infty} \mathbf{P}^n = \mathbf{Q}(\lim_{n\to\infty} \mathbf{J}^n)\mathbf{Q}^{-1} = \mathbf{Q}\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{0} & \\ 0 & & & \end{pmatrix}\mathbf{Q}^{-1} = \begin{pmatrix} \vec{\pi} \\ \vdots \\ \vec{\pi} \end{pmatrix}.$$

QED

Use of the Perron-Frobenius Theorem (continued)

Proof Continued:

$\mathbf{P}$ is similar to its Jordan canonical form. That is, $\mathbf{P} = \mathbf{QJQ}^{-1}$.
$(1, ..., 1)^T = \mathbf{P}(1, ..., 1)^T = \mathbf{QJQ}^{-1}(1, ..., 1)^T = \mathbf{QJ}(1, 0, ..., 0)^T = \mathbf{Q}(1, 0..., 0)^T$.
It follows that the first column of $\mathbf{Q}$ is all 1's.
$\vec{\pi} = \mathbf{P}^T\vec{\pi} = (\mathbf{Q}^{-1})^T\mathbf{J}^T\mathbf{Q}^T\vec{\pi} = (\mathbf{Q}^{-1})^T\mathbf{J}^T(1, 0, ..., 0)^T = (\mathbf{Q}^{-1})^T(1, 0..., 0)^T$.
It follows that the first row of $\mathbf{Q}^{-1}$ is $\vec{\pi}$. Therefore,

$$\lim_{n\to\infty} \mathbf{P}^n = \mathbf{Q}(\lim_{n\to\infty} \mathbf{J}^n)\mathbf{Q}^{-1} = \mathbf{Q}\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \mathbf{0} & \\ 0 & & & \end{pmatrix}\mathbf{Q}^{-1} = \begin{pmatrix} \vec{\pi} \\ \vdots \\ \vec{\pi} \end{pmatrix}.$$

QED

Remark

There is an analytic proof to the previous theorem that uses less machinery.

1.20(a)

Lemma

If $\vec{v} \geq \vec{0}$ *and* $\vec{v} \neq \vec{0}$, *then* $\mathbf{A}\vec{v} > \vec{0}$.

1.20(a)

Lemma

If $\vec{v} \geq \vec{0}$ *and* $\vec{v} \neq \vec{0}$, *then* $\mathbf{A}\vec{v} > \vec{0}$.

Proof:

If $\vec{v} \geq \vec{0}$ and $\vec{v} \neq \vec{0}$, then the v^j are all non-negative and not all zero. That is, $v^j \geq 0$ for all j and there is a j_0 such that $v^{j_0} > 0$. Thus, the ith coordinate of $\mathbf{A}\vec{v}$ equals $\sum_{j=1}^{n} a_{ij} v^j \geq a_{ij_0} v^{j_0} > 0$. Therefore, $\mathbf{A}\vec{v} > \vec{0}$.
QED

Another Definition

Definition

For any nonzero $\vec{v} \geq \vec{0}$, let $g(\vec{v})$ be the largest λ such that $\mathbf{A}\vec{v} \geq \lambda\vec{v}$

Let's verify that there indeed is a largest such λ.

Another Definition

Definition

For any nonzero $\vec{v} \geq \vec{0}$, let $g(\vec{v})$ be the largest λ such that $\mathbf{A}\vec{v} \geq \lambda\vec{v}$

Let's verify that there indeed is a largest such λ.

Consider a nonzero vector $\vec{v} \geq 0$. Let $[\mathbf{A}\vec{v}]^i$ denote the ith coordinate of $\mathbf{A}\vec{v}$. If $v^i \neq 0$, let $\lambda_i = \frac{[\mathbf{A}\vec{v}]^i}{v^i}$. For such i, $[\mathbf{A}\vec{v}]^i = \lambda_i v^i$. Let $\lambda = \min\{\lambda_i\}$. If $v_i \neq 0$, $[\mathbf{A}\vec{v}]^i = \lambda_i v^i \geq \lambda v^i$, and if $v^i = 0$, $[\mathbf{A}\vec{v}]^i > 0 = \lambda v^i$ by 1.20(a). Therefore, $\mathbf{A}\vec{v} \geq \lambda\vec{v}$.

Another Definition

Definition

For any nonzero $\vec{v} \geq \vec{0}$, let $g(\vec{v})$ be the largest λ such that $\mathbf{A}\vec{v} \geq \lambda\vec{v}$

Let's verify that there indeed is a largest such λ.

Consider a nonzero vector $\vec{v} \geq 0$. Let $[\mathbf{A}\vec{v}]^i$ denote the ith coordinate of $\mathbf{A}\vec{v}$. If $v^i \neq 0$, let $\lambda_i = \frac{[\mathbf{A}\vec{v}]^i}{v^i}$. For such i, $[\mathbf{A}\vec{v}]^i = \lambda_i v^i$. Let $\lambda = \min\{\lambda_i\}$. If $v_i \neq 0$, $[\mathbf{A}\vec{v}]^i = \lambda_i v^i \geq \lambda v^i$, and if $v^i = 0$, $[\mathbf{A}\vec{v}]^i > 0 = \lambda v^i$ by 1.20(a). Therefore, $\mathbf{A}\vec{v} \geq \lambda\vec{v}$.

If $\lambda' > \lambda$, then $\lambda' > \lambda_i$ for some i such that $v_i \neq 0$. In that case, $[\mathbf{A}\vec{v}]^i = \lambda_i v^i < \lambda' v^i$. So it is not true that $\mathbf{A}\vec{v} \geq \lambda'\vec{v}$. Thus, λ is the largest number such that $\mathbf{A}\vec{v} \geq \lambda\vec{v}$.

1.20(b)

Lemma

For any nonzero $\vec{v} \geq \vec{0}$, $g(\vec{v}) > 0$, *and if* $c > 0$, *then* $g(c\vec{v}) = g(\vec{v})$

1.20(b)

Lemma

For any nonzero $\vec{v} \geq \vec{0}$, $g(\vec{v}) > 0$, *and if* $c > 0$, *then* $g(c\vec{v}) = g(\vec{v})$

Proof:

From the previous slide, we know that $g(\vec{v}) = \min\left\{\frac{[\mathbf{A}\vec{v}]^i}{v^i} : v_i \neq 0\right\}$. By 1.20(a), $[\mathbf{A}\vec{v}]^i > 0$ for all i. Therefore, $g(\vec{v}) > 0$.

Lemma

For any nonzero $\vec{v} \geq \vec{0}$, $g(\vec{v}) > 0$, *and if* $c > 0$, *then* $g(c\vec{v}) = g(\vec{v})$

Proof:

From the previous slide, we know that $g(\vec{v}) = \min\left\{\frac{[\mathbf{A}\vec{v}]^i}{v^i} : v_i \neq 0\right\}$. By 1.20(a), $[\mathbf{A}\vec{v}]^i > 0$ for all i. Therefore, $g(\vec{v}) > 0$.

For the second part of the statement, let $c > 0$.
$\mathbf{A}(c\vec{v}) = c(\mathbf{A}\vec{v}) \geq cg(\vec{v})\vec{v} = g(\vec{v})(c\vec{v})$. Thus, $g(c\vec{v}) \geq g(\vec{v})$.
$\mathbf{A}\vec{v} = \frac{1}{c}(\mathbf{A}(c\vec{v})) \geq \frac{1}{c}g(c\vec{v})(c\vec{v}) = g(c\vec{v})\vec{v}$. Thus, $g(\vec{v}) \geq g(c\vec{v})$.
Therefore, $g(c\vec{v}) = g(\vec{v})$. QED

Defining α

Let $\alpha = \sup\{g(\vec{v}) : \vec{v} \geq \vec{0}, \vec{v} \neq \vec{0}\}$. I will show that α is an eigenvalue of **A**.

Defining α

Let $\alpha = \sup\{g(\vec{v}) : \vec{v} \geq \vec{0}, \vec{v} \neq \vec{0}\}$. I will show that α is an eigenvalue of $\mathbf{A}$.

For any nonzero vector $\vec{v} \geq \vec{0}$, there is a vector $\vec{w} = \frac{\vec{v}}{\|\vec{v}\|} \geq \vec{0}$ with norm 1 $(\|\vec{w}\| = \|\frac{\vec{v}}{\|\vec{v}\|}\| = \frac{\|\vec{v}\|}{\|\vec{v}\|} = 1)$ such that $g(\vec{w}) = g(\frac{\vec{v}}{\|\vec{v}\|}) = g(\vec{v})$ by 1.20(b).

Therefore, $\alpha = \sup\{g(\vec{v}) : \vec{v} \geq \vec{0}, \|\vec{v}\| = 1\}$.

Defining α

Let $\alpha = \sup\{g(\vec{v}) : \vec{v} \geq \vec{0}, \vec{v} \neq \vec{0}\}$. I will show that α is an eigenvalue of **A**.

For any nonzero vector $\vec{v} \geq \vec{0}$, there is a vector $\vec{w} = \frac{\vec{v}}{\|\vec{v}\|} \geq \vec{0}$ with norm 1 $(\|\vec{w}\| = \|\frac{\vec{v}}{\|\vec{v}\|}\| = \frac{\|\vec{v}\|}{\|\vec{v}\|} = 1)$ such that $g(\vec{w}) = g(\frac{\vec{v}}{\|\vec{v}\|}) = g(\vec{v})$ by 1.20(b).

Therefore, $\alpha = \sup\{g(\vec{v}) : \vec{v} \geq \vec{0}, \|\vec{v}\| = 1\}$.

The function $g(\vec{v})$ can be shown to be a continuous function, and the set $\{\vec{v} \in \mathbb{R}^n : \vec{v} \geq \vec{0}, \|\vec{v}\| = 1\}\}$ is a compact subset of $\mathbb{R}^n$. Therefore, g attains a maximum value on $\{\vec{v} \in \mathbb{R}^n : \vec{v} \geq \vec{0}, \|\vec{v}\| = 1\}\}$.

That is to say, there is a vector $\vec{v} \geq \vec{0}$ with norm 1 such that $g(\vec{v}) = \alpha$.

1.20(c)

Theorem (part 1)

For any vector $\vec{v} \geq \vec{0}$ $(\vec{v} \neq \vec{0})$ *with* $g(\vec{v}) = \alpha$, $\mathbf{A}\vec{v} = \alpha\vec{v}$

1.20(c)

Theorem (part 1)

For any vector $\vec{v} \geq \vec{0}$ $(\vec{v} \neq \vec{0})$ *with* $g(\vec{v}) = \alpha$, $\mathbf{A}\vec{v} = \alpha\vec{v}$

Proof:

Suppose $\mathbf{A}\vec{v} \neq \alpha\vec{v}$. Since $g(\vec{v}) = \alpha$, $\mathbf{A}\vec{v} \geq \alpha\vec{v}$. Thus, $\mathbf{A}\vec{v} - \alpha\vec{v} \geq \vec{0}$ and $\mathbf{A}\vec{v} - \alpha\vec{v} \neq \vec{0}$. So by part (a), $\mathbf{A}(\mathbf{A}\vec{v} - \alpha\vec{v}) > \vec{0}$. Therefore, there is a $\lambda > 0$ small enough such that $\mathbf{A}(\mathbf{A}\vec{v} - \alpha\vec{v}) \geq \lambda(\mathbf{A}\vec{v})$. Then $\mathbf{A}(\mathbf{A}\vec{v}) \geq (\alpha + \lambda)(\mathbf{A}\vec{v})$, and $\mathbf{A}\vec{v} > \vec{0}$ by part (a). Thus, $g(\mathbf{A}\vec{v}) \geq \alpha + \lambda > \alpha$. This contradicts the fact that α is the supremum of $g(\vec{w})$ over all nonzero $\vec{w} \geq 0$. Therefore, $\mathbf{A}\vec{v} = \alpha\vec{v}$. QED

1.20(d)

Theorem (part 2)

There is a unique $\vec{v} \geq \vec{0}$ *with* $\|\vec{v}\| = 1$ *such that* $g(\vec{v}) = \alpha$.

1.20(d)

Theorem (part 2)

There is a unique $\vec{v} \geq \vec{0}$ *with* $\|\vec{v}\| = 1$ *such that* $g(\vec{v}) = \alpha$.

Proof:

Let $\vec{v}_1, \vec{v}_2 \geq \vec{0}$ have norm 1 and be such that $g(\vec{v}_1) = g(\vec{v}_2) = \alpha$.
Suppose $\vec{v}_1 \neq \vec{v}_2$. Then $|\vec{v}_1 - \vec{v}_2| \geq \vec{0}$ and $|v_1 - v_2| \neq \vec{0}$.
$\mathbf{A}|\vec{v}_1 - \vec{v}_2| \geq |\mathbf{A}(\vec{v}_1 - \vec{v}_2)| = |\mathbf{A}\vec{v}_1 - \mathbf{A}\vec{v}_2| = |\alpha\vec{v}_1 - \alpha\vec{v}_2| = \alpha|\vec{v}_1 - \vec{v}_2|$.
Thus, $g(|\vec{v}_1 - \vec{v}_2|) = \alpha$. So $\mathbf{A}|\vec{v}_1 - \vec{v}_2| = \alpha|\vec{v}_1 - \vec{v}_2|$ by 1.20(c). Therefore,
$\mathbf{A}|\vec{v}_1 - \vec{v}_2| \geq |\mathbf{A}(\vec{v}_1 - \vec{v}_2)| = \alpha|\vec{v}_1 - \vec{v}_2| = \mathbf{A}|\vec{v}_1 - \vec{v}_2|$. So
$\mathbf{A}|\vec{v}_1 - \vec{v}_2| = |\mathbf{A}(\vec{v}_1 - \vec{v}_2)|$.

1.20(d)

Theorem (part 2)

There is a unique $\vec{v} \geq \vec{0}$ *with* $\|\vec{v}\| = 1$ *such that* $g(\vec{v}) = \alpha$.

Proof:

Let $\vec{v}_1, \vec{v}_2 \geq \vec{0}$ have norm 1 and be such that $g(\vec{v}_1) = g(\vec{v}_2) = \alpha$. Suppose $\vec{v}_1 \neq \vec{v}_2$. Then $|\vec{v}_1 - \vec{v}_2| \geq \vec{0}$ and $|v_1 - v_2| \neq \vec{0}$. $\mathbf{A}|\vec{v}_1 - \vec{v}_2| \geq |\mathbf{A}(\vec{v}_1 - \vec{v}_2)| = |\mathbf{A}\vec{v}_1 - \mathbf{A}\vec{v}_2| = |\alpha\vec{v}_1 - \alpha\vec{v}_2| = \alpha|\vec{v}_1 - \vec{v}_2|$. Thus, $g(|\vec{v}_1 - \vec{v}_2|) = \alpha$. So $\mathbf{A}|\vec{v}_1 - \vec{v}_2| = \alpha|\vec{v}_1 - \vec{v}_2|$ by 1.20(c). Therefore, $\mathbf{A}|\vec{v}_1 - \vec{v}_2| \geq |\mathbf{A}(\vec{v}_1 - \vec{v}_2)| = \alpha|\vec{v}_1 - \vec{v}_2| = \mathbf{A}|\vec{v}_1 - \vec{v}_2|$. So $\mathbf{A}|\vec{v}_1 - \vec{v}_2| = |\mathbf{A}(\vec{v}_1 - \vec{v}_2)|$. Thus, the ith coordinate of $\mathbf{A}|\vec{v}_1 - \vec{v}_2|$ and $|\mathbf{A}(\vec{v}_1 - \vec{v}_2)|$ are equal: $\sum_{j=1}^{n} a_{ij}|v_1^j - v_2^j| = |\sum_{j=1}^{n} a_{ij}(v_1^j - v_2^j)|$. It then follows from properties of the absolute values that $\vec{v}_1 \geq \vec{v}_2$ or $\vec{v}_2 \geq \vec{v}_1$. As $\|\vec{v}_1\| = \|\vec{v}_2\|$, it must be that $\vec{v}_1 = \vec{v}_2$ contradicting the supposition. Therefore, $\vec{v}_1 = \vec{v}_2$. QED

1.20(e)

Theorem (part 3)

If $\vec{v} \geq \vec{0}$ is the unique vector with norm 1 such that $g(\vec{v}) = \alpha$, then $\vec{v} > \vec{0}$.

1.20(e)

Theorem (part 3)

If $\vec{v} \geq \vec{0}$ is the unique vector with norm 1 such that $g(\vec{v}) = \alpha$, then $\vec{v} > \vec{0}$.

Proof:

By 1.20(a), $\mathbf{A}\vec{v} > 0$. $\mathbf{A}(\mathbf{A}\vec{v}) = \mathbf{A}(\alpha\vec{v}) = \alpha(\mathbf{A}\vec{v})$. So $g(\mathbf{A}\vec{v}) = \alpha$. Let $\vec{w} = \mathbf{A}\vec{v}/\|\mathbf{A}\vec{v}\|$. Then $\vec{w} > \vec{0}$, $\|\vec{w}\| = 1$, and by 1.20(b), $g(\vec{w}) = g(\mathbf{A}\vec{v}/\|\mathbf{A}\vec{v}\|) = g(\mathbf{A}\vec{v}) = \alpha$. But $\vec{v}$ is the unique vector satisfying these properties, so $\vec{v} = \vec{w} > \vec{0}$. QED

1.20(f)

Theorem (part 4)

If $\lambda \neq \alpha$ is an eigenvalue of $\mathbf{A}$*, then* $|\lambda| < \alpha$.

1.20(f)

Theorem (part 4)

If $\lambda \neq \alpha$ is an eigenvalue of $\mathbf{A}$, then $|\lambda| < \alpha$.

Proof:

Let $\vec{u}$ be a eigenvector that corresponds to the eigenvalue λ. $\mathbf{A}\vec{u} = \lambda\vec{u}$. Then $|\vec{u}| \geq \vec{0}$ is nonzero. Furthermore, $\mathbf{A}|\vec{u}| \geq |\mathbf{A}\vec{u}| = |\lambda\vec{u}| = |\lambda||\vec{u}|$. Therefore, $|\lambda| \leq g(|\vec{u}|) \leq \alpha$. Suppose $|\lambda| = \alpha$. Then $\mathbf{A}|\vec{u}| \geq |\mathbf{A}\vec{u}| = \alpha|\vec{u}| = \mathbf{A}|\vec{u}|$. So $\mathbf{A}|\vec{u}| = |\mathbf{A}\vec{u}|$.

1.20(f)

Theorem (part 4)

If $\lambda \neq \alpha$ is an eigenvalue of $\mathbf{A}$, then $|\lambda| < \alpha$.

Proof:

Let $\vec{u}$ be a eigenvector that corresponds to the eigenvalue λ. $\mathbf{A}\vec{u} = \lambda\vec{u}$. Then $|\vec{u}| \geq \vec{0}$ is nonzero. Furthermore, $\mathbf{A}|\vec{u}| \geq |\mathbf{A}\vec{u}| = |\lambda\vec{u}| = |\lambda||\vec{u}|$. Therefore, $|\lambda| \leq g(|\vec{u}|) \leq \alpha$. Suppose $|\lambda| = \alpha$. Then $\mathbf{A}|\vec{u}| \geq |\mathbf{A}\vec{u}| = \alpha|\vec{u}| = \mathbf{A}|\vec{u}|$. So $\mathbf{A}|\vec{u}| = |\mathbf{A}\vec{u}|$. Thus, the ith coordinate of $\mathbf{A}|\vec{u}|$ and $|\mathbf{A}\vec{u}|$ are equal: $\sum_{j=1}^{n} a_{ij}|u^j| = |\sum_{j=1}^{n} a_{ij}u^j|$. It then follows from properties of the absolute values that $\vec{u} = e^{i\theta}\vec{w}$ for some angle θ and vector $\vec{w} \geq \vec{0}$. Then $\mathbf{A}\vec{w} = \lambda\vec{w}$, so it must be that $\lambda \geq 0$. Hence, $\lambda = \alpha$ contradicting the fact that $\lambda \neq \alpha$. Therefore, $|\lambda| < \alpha$. QED

1.20(g)

Lemma

Let $\mathbf{B}^k$ *be the submatrix of* $\mathbf{A}$ *obtained by deleting the* k*th row and the* k*th column. All the eigenvalues of* $\mathbf{B}^k$ *have absolute value strictly less than* α*.*

1.20(g)

Lemma

Let $\mathbf{B}^k$ *be the submatrix of* $\mathbf{A}$ *obtained by deleting the* k*th row and the* k*th column. All the eigenvalues of* $\mathbf{B}^k$ *have absolute value strictly less than* α.

Proof:

$\mathbf{B}^k = (b_{ij})$ is a matrix with positive entries. Therefore, we may apply 1.20(a)-(f) to $\mathbf{B}^k$. For any nonzero $\vec{w} \geq \vec{0}$, let $h(\vec{w})$ be the largest λ such that $\mathbf{B}^k \vec{w} \geq \lambda \vec{w}$, and let $\beta = \sup\{h(\vec{w}) : \vec{w} \geq \vec{0}, \vec{w} \neq \vec{0}\}$. Then there is a unique vector $\vec{w} > \vec{0}$ with norm 1 such that $\mathbf{B}^k \vec{w} = \beta \vec{w}$. Thus β is an eigenvalue of $\mathbf{B}^k$, and $|\lambda| < \beta$ for all other eigenvalues λ of $\mathbf{B}^k$.

1.20(g)

Lemma

Let $\mathbf{B}^k$ be the submatrix of $\mathbf{A}$ obtained by deleting the kth row and the kth column. All the eigenvalues of $\mathbf{B}^k$ have absolute value strictly less than α.

Proof:

$\mathbf{B}^k = (b_{ij})$ is a matrix with positive entries. Therefore, we may apply 1.20(a)-(f) to $\mathbf{B}^k$. For any nonzero $\vec{w} \geq \vec{0}$, let $h(\vec{w})$ be the largest λ such that $\mathbf{B}^k \vec{w} \geq \lambda \vec{w}$, and let $\beta = \sup\{h(\vec{w}) : \vec{w} \geq \vec{0}, \vec{w} \neq \vec{0}\}$. Then there is a unique vector $\vec{w} > \vec{0}$ with norm 1 such that $\mathbf{B}^k \vec{w} = \beta \vec{w}$. Thus β is an eigenvalue of $\mathbf{B}^k$, and $|\lambda| < \beta$ for all other eigenvalues λ of $\mathbf{B}^k$.

Let $\vec{w}_0 = (w^1, ..., w^{k-1}, 0, w^k, ..., w^{n-1})$. $\vec{w}_0 \geq \vec{0}$ and $\|\vec{w}_0\| = \|\vec{w}\| = 1$.
If $i = k$, $[\mathbf{A}\vec{w}_0]^i = \sum_{j=1}^n a_{ij} w_0^j = \sum_{j \neq k} a_{ij} w_0^j > 0 = \beta w_0^i$. And if $i \neq k$,
$[\mathbf{A}\vec{w}_0]^i = \sum_{j=1}^n a_{ij} w_0^j = \sum_{j \neq k} a_{ij} w_0^j = \sum_{j=1}^{n-1} b_{ij} w^j = [\mathbf{B}^k \vec{w}]^i = \beta w^i = \beta w_0^i$.
Therefore, $\mathbf{A}\vec{w}_0 \geq \beta \vec{w}_0$, but the two are not equal. So $\beta \leq g(\vec{w}_0) \leq \alpha$. If $\beta = \alpha$, then $\mathbf{A}\vec{w}_0$ must equal $\beta \vec{w}_0$ by 1.20(c), which is not the case. Thus, $\beta < \alpha$, and for all eigenvalues λ of $\mathbf{B}^k$, $|\lambda| < \beta < \alpha$. QED

1.20(h)

Lemma

Let $f(\lambda) = \det(\lambda \mathbf{I} - A)$ *be the characteristic polynomial of* $\mathbf{A}$.
$f'(\lambda) = \sum_{k=1}^{n} \det(\lambda \mathbf{I} - \mathbf{B}^k)$

1.20(h)

Lemma

Let $f(\lambda) = \det(\lambda \mathbf{I} - A)$ *be the characteristic polynomial of* $\mathbf{A}$.
$f'(\lambda) = \sum_{k=1}^{n} \det(\lambda \mathbf{I} - \mathbf{B}^k)$

Proof:

Let $\lambda \mathbf{I} - \mathbf{A} = (c_{ij})$, and $\lambda \mathbf{I} - \mathbf{B}^k = (d_{ij}^k)$.

$$\frac{d}{d\lambda} f(\lambda) = \frac{d}{d\lambda} \sum_{\sigma \in S^n} \operatorname{sgn}(\sigma) c_{1\sigma(1)} \cdots c_{n\sigma(n)}$$

1.20(h)

Lemma

Let $f(\lambda) = \det(\lambda\mathbf{I} - A)$ *be the characteristic polynomial of* $\mathbf{A}$.
$f'(\lambda) = \sum_{k=1}^{n} \det(\lambda\mathbf{I} - \mathbf{B}^k)$

Proof:

Let $\lambda\mathbf{I} - \mathbf{A} = (c_{ij})$, and $\lambda\mathbf{I} - \mathbf{B}^k = (d_{ij}^k)$.

$$\begin{aligned}\frac{d}{d\lambda} f(\lambda) =& \frac{d}{d\lambda} \sum_{\sigma \in S^n} \operatorname{sgn}(\sigma) c_{1\sigma(1)} \cdots c_{n\sigma(n)} \\ &= \sum_{\sigma \in S^n} \operatorname{sgn}(\sigma) \sum_{k} (c_{1\sigma(1)} \cdots \frac{d}{d\lambda} c_{k\sigma(k)} \cdots c_{n\sigma(n)})\end{aligned}$$

1.20(h)

Lemma

Let $f(\lambda) = \det(\lambda\mathbf{I} - A)$ *be the characteristic polynomial of* $\mathbf{A}$*.*
$f'(\lambda) = \sum_{k=1}^{n} \det(\lambda\mathbf{I} - \mathbf{B}^k)$

Proof:

Let $\lambda\mathbf{I} - \mathbf{A} = (c_{ij})$, and $\lambda\mathbf{I} - \mathbf{B}^k = (d_{ij}^k)$.

$$
\begin{aligned}
\frac{d}{d\lambda} f(\lambda) =& \frac{d}{d\lambda} \sum_{\sigma \in S^n} \operatorname{sgn}(\sigma) c_{1\sigma(1)} \cdots c_{n\sigma(n)} \\
=& \sum_{\sigma \in S^n} \operatorname{sgn}(\sigma) \sum_{k} (c_{1\sigma(1)} \cdots \frac{d}{d\lambda} c_{k\sigma(k)} \cdots c_{n\sigma(n)}) \\
=& \sum_{k} \sum_{\substack{\sigma \in S^n \\ \sigma(k)=k}} \operatorname{sgn}(\sigma) c_{1\sigma(1)} \cdots c_{(k-1)\sigma(k-1)} c_{(k+1)\sigma(k+1)} \cdots c_{n\sigma(n)}
\end{aligned}
$$

1.20(h)

Lemma

Let $f(\lambda) = \det(\lambda\mathbf{I} - A)$ *be the characteristic polynomial of* $\mathbf{A}$*.*
$f'(\lambda) = \sum_{k=1}^{n} \det(\lambda\mathbf{I} - \mathbf{B}^k)$

Proof:

Let $\lambda\mathbf{I} - \mathbf{A} = (c_{ij})$, and $\lambda\mathbf{I} - \mathbf{B}^k = (d_{ij}^k)$.

$$\begin{aligned}
\frac{d}{d\lambda} f(\lambda) =& \frac{d}{d\lambda} \sum_{\sigma \in S^n} \operatorname{sgn}(\sigma) c_{1\sigma(1)} \cdots c_{n\sigma(n)} \\
=& \sum_{\sigma \in S^n} \operatorname{sgn}(\sigma) \sum_{k} (c_{1\sigma(1)} \cdots \frac{d}{d\lambda} c_{k\sigma(k)} \cdots c_{n\sigma(n)}) \\
=& \sum_{k} \sum_{\substack{\sigma \in S^n \\ \sigma(k)=k}} \operatorname{sgn}(\sigma) c_{1\sigma(1)} \cdots c_{(k-1)\sigma(k-1)} c_{(k+1)\sigma(k+1)} \cdots c_{n\sigma(n)} \\
=& \sum_{k} \sum_{\sigma \in S^{n-1}} \operatorname{sgn}(\sigma) d_{1\sigma(1)}^k \cdots d_{n-1\sigma(n-1)}^k = \sum_{k=1}^{n} \det(\lambda\mathbf{I} - \mathbf{B}^k) \quad \text{QED}
\end{aligned}$$

Theorem (part 5)

$f'(\alpha) > 0$. *Hence, α is a simple eigenvalue for* $\mathbf{A}$.

1.20(i)

Theorem (part 5)

$f'(\alpha) > 0$. *Hence,* α *is a simple eigenvalue for* $\mathbf{A}$*.*

Proof:

$\det(\lambda\mathbf{I} - \mathbf{B}^k)$ is polynomial in λ of degree $n-1$ with leading term λ^{n-1}
Therefore, $\det(\lambda\mathbf{I} - \mathbf{B}^k) > 0$ for all λ greater than the largest real eigenvalue of $\mathbf{B}^k$. By 1.20(g), α is greater than the largest real eigenvalue of $\mathbf{B}^k$. Therefore, $f'(\alpha) = \sum_{k=1}^{n} \det(\alpha\mathbf{I} - \mathbf{B}^k) > 0$.

1.20(i)

Theorem (part 5)

$f'(\alpha) > 0$. Hence, α is a simple eigenvalue for $\mathbf{A}$.

Proof:

$\det(\lambda\mathbf{I} - \mathbf{B}^k)$ is polynomial in λ of degree $n-1$ with leading term λ^{n-1}
Therefore, $\det(\lambda\mathbf{I} - \mathbf{B}^k) > 0$ for all λ greater than the largest real eigenvalue of $\mathbf{B}^k$. By 1.20(g), α is greater than the largest real eigenvalue of $\mathbf{B}^k$. Therefore, $f'(\alpha) = \sum_{k=1}^{n} \det(\alpha\mathbf{I} - \mathbf{B}^k) > 0$.

$f(\lambda) = (\lambda - \alpha)^m q(\lambda)$ where q is a nonzero polynomial whose roots have absolute value less than α. $m \in \mathbb{N}$ is the multiplicity of α. If $m \neq 1$, then $f'(\alpha) = m(\alpha - \alpha)^{m-1} q(\alpha) + (\alpha - \alpha)^m q'(\alpha) = 0$, which contradicts the fact that $f'(\alpha) > 0$. Therefore, α is a root of the characteristic polynomial of $\mathbf{A}$ with multiplicity 1. That is, α is a simple eigenvalue for $\mathbf{A}$. QED

Summary

Perron-Frobenius Theorem

Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ matrix with $a_{ij} > 0$ forall i, j. $\mathbf{A}$ has a real eigenvalue $\alpha > 0$ with a unique eigenvector $\vec{v} > \vec{0}$ of norm 1 such that $|\lambda| < \alpha$ for any other eigenvalue λ of $\mathbf{A}$. Moreover, α is a simple eigenvalue (i.e., it is a root of the characteristic polynomial of $\mathbf{A}$ with multiplicity 1).

1.20(c) shows that $\alpha > 0$ is an eigenvalue of $\mathbf{A}$.
1.20(d) shows that there is α has a unique eigenvector $\vec{v} \geq \vec{0}$ with norm 1.
1.20(e) shows that, in fact, $\vec{v} > \vec{0}$.
1.20(f) shows that all other eigenvalues λ of $\mathbf{A}$, $|\lambda| < \alpha$.
1.20(i) shows that α is a simple eigenvalue of $\mathbf{A}$.