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 $\vec{u} = (u^1, \ldots, u^n)$ and $\vec{v} = (v^1, \ldots, v^n)$ 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 $i$.

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

Doron Shahar
Proof of the Perron-Frobenius Theorem
May 1, 2014 2 / 18
Statement of Theorem

Perron-Frobenius Theorem

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

1.20(c) shows that $\alpha > 0$ is an eigenvalue of $A$.
1.20(d) shows that there is $\alpha$ 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 $\lambda$ of $A$, $|\lambda| < \alpha$.
1.20(i) shows that $\alpha$ is a simple eigenvalue of $A$. 

Doron Shahar
Proof of the Perron-Frobenius Theorem
May 1, 2014 3 / 18
Theorem

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

Proof:

$P^T$ is a square matrix with all positive entries. By the Perron-Frobenius theorem, $P^T$ 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 $\vec{\pi} = (\pi_1, \ldots, \pi_n) > \vec{0}$ is a probability distribution. Since $P$ is a stochastic matrix, $P^T \vec{\pi} = \alpha \vec{\pi}$ is a probability distribution. Therefore, $\alpha = \alpha \sum_{i=1}^{n} \pi_i = \sum_{i=1}^{n} \alpha \pi_i = 1$. So $P^T \vec{\pi} = \vec{\pi}$, and $\vec{\pi} P = \vec{\pi}$ if we equate $\vec{\pi}$ with $\vec{\pi}^T$. Therefore, $\vec{\pi}$ is an invariant probability distribution of $P$. If $\vec{\pi}'$ is another invariant probability distribution of $P$, then $P^T \vec{\pi}' = \vec{\pi}'$. 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 $P$, and $\vec{\pi} > \vec{0}$. QED
Theorem

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

Proof:

$P^T$ is a square matrix with all positive entries. By the Perron-Frobenius theorem, $P^T$ 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 $P$ is a stochastic matrix, $P^T \vec{\pi} = \alpha \vec{\pi}$ is a probability distribution. Therefore, $\alpha = \alpha \sum_{i=1}^{n} \pi^i = \sum_{i=1}^{n} \alpha \pi^i = 1$. So $P^T \vec{\pi} = \vec{\pi}$, and $\vec{\pi} P = \vec{\pi}$ if we equate $\vec{\pi}$ with $\vec{\pi}^T$. Therefore, $\vec{\pi}$ is an invariant probability distribution of $P$. 
Theorem

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

Proof:

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

**Theorem**

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

Proof: By the Perron-Frobenius theorem and the proof of 1.20(j), \( P^T \) has a simple eigenvalue of 1 and all other eigenvalues of \( P^T \) have absolute value less than 1. The same is true for \( P \). Therefore, the Jordan canonical form of \( P \) is

\[
J = \begin{pmatrix}
1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & M
\end{pmatrix}
\]

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

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

Proof:

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

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

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

Doron Shahar

Proof of the Perron-Frobenius Theorem

May 1, 2014 5 / 18
Proof Continued:

$P$ is similar to its Jordan canonical form. That is, $P = QJQ^{-1}$. 

$(1, ..., 1)^T = P(1, ..., 1)^T = QJQ^{-1}(1, ..., 1)^T = QJ(1, 0, ..., 0)^T = Q(1, 0, ..., 0)^T$. It follows that the first column of $Q$ is all 1's.

$\vec{\pi} = P^T\vec{\pi} = (Q^{-1})^TJ^TQ^T\vec{\pi} = (Q^{-1})^TJ^T(1, 0, ..., 0)^T = (Q^{-1})^T(1, 0, ..., 0)^T$. It follows that the first row of $Q^{-1}$ is $\vec{\pi}$.

Therefore, $\lim_{n \to \infty} P^n = Q(\lim_{n \to \infty} J^n)Q^{-1} = Q\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}Q^{-1} = \begin{bmatrix} \vec{\pi} \\ \vdots \\ \vec{\pi} \end{bmatrix}$.

QED

Remark: There is an analytic proof to the previous theorem that uses less machinery.
Proof Continued:

\( P \) is similar to its Jordan canonical form. That is, \( P = QJQ^{-1} \).

\[
(1, \ldots, 1)^T = P(1, \ldots, 1)^T = QJQ^{-1}(1, \ldots, 1)^T = QJ(1, 0, \ldots, 0)^T = Q(1, 0, \ldots, 0)^T.
\]

It follows that the first column of \( Q \) is all 1’s.

\[
\overrightarrow{\pi} = P^T \overrightarrow{\pi} = (Q^{-1})^T J^T Q^T \overrightarrow{\pi} = (Q^{-1})^T J^T (1, 0, \ldots, 0)^T = (Q^{-1})^T (1, 0, \ldots, 0)^T.
\]

It follows that the first row of \( Q^{-1} \) is \( \overrightarrow{\pi} \). Therefore,

\[
\lim_{n \to \infty} P^n = Q(\lim_{n \to \infty} J^n)Q^{-1} = Q \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} Q^{-1} = \begin{pmatrix} \overrightarrow{\pi} \\ \vdots \\ \overrightarrow{\pi} \end{pmatrix}.
\]

QED
Proof Continued:

\( P \) is similar to its Jordan canonical form. That is, \( P = Q J Q^{-1} \).

\((1, ..., 1)^T = P(1, ..., 1)^T = Q J Q^{-1} (1, ..., 1)^T = Q J (1, 0, ..., 0)^T = Q (1, 0, ..., 0)^T. \)

It follows that the first column of \( Q \) is all 1’s.

\( \vec{\pi} = P^T \vec{\pi} = (Q^{-1})^T J^T Q^T \vec{\pi} = (Q^{-1})^T J^T (1, 0, ..., 0)^T = (Q^{-1})^T (1, 0, ..., 0)^T. \)

It follows that the first row of \( Q^{-1} \) is \( \vec{\pi} \). Therefore,

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

QED

Remark

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

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

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

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 $i$th coordinate of $A\vec{v}$ equals $\sum_{j=1}^{n} a_{ij} v^j \geq a_{ij_0} v^{j_0} > 0$. Therefore, $A\vec{v} > \vec{0}$. QED
Another Definition

Definition

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

Let’s verify that there indeed is a largest such $\lambda$. 

Doron Shahar
Proof of the Perron-Frobenius Theorem
May 1, 2014 8 / 18
Another Definition

Definition

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

Let’s verify that there indeed is a largest such $\lambda$.

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

Doron Shahar

Proof of the Perron-Frobenius Theorem

May 1, 2014 8 / 18
Definition

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

Let’s verify that there indeed is a largest such \( \lambda \).

Consider a nonzero vector \( \vec{v} \geq 0 \). Let \( [A\vec{v}]^i \) denote the \( i \)th coordinate of \( A\vec{v} \). If \( v^i \neq 0 \), let \( \lambda_i = \frac{[A\vec{v}]^i}{v^i} \). For such \( i \), \( [A\vec{v}]^i = \lambda_i v^i \). Let \( \lambda = \min\{\lambda_i\} \).

If \( v^i \neq 0 \), \( [A\vec{v}]^i = \lambda_i v^i \geq \lambda v^i \), and if \( v^i = 0 \), \( [A\vec{v}]^i > 0 = \lambda v^i \) by 1.20(a).

Therefore, \( A\vec{v} \geq \lambda \vec{v} \).

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

For any nonzero $\vec{v} \geq 0$, $g(\vec{v}) > 0$, and if $c > 0$, then $g(c\vec{v}) = g(\vec{v})$.
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 $g(\vec{v}) = \min \left\{ \frac{[A\vec{v}]^i}{v_i} : v_i \neq 0 \right\}$. By 1.20(a), $[A\vec{v}]^i > 0$ for all $i$. Therefore, $g(\vec{v}) > 0$. 

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

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

Proof:

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

For the second part of the statement, let $c > 0$.

$A(c\vec{v}) = c(A\vec{v}) \geq cg(\vec{v})\vec{v} = g(\vec{v})(c\vec{v})$. Thus, $g(c\vec{v}) \geq g(\vec{v})$.

$A\vec{v} = \frac{1}{c}(A(c\vec{v})) \geq \frac{1}{c}g(c\vec{v})(c\vec{v}) = g(c\vec{v})\vec{v}$. Thus, $g(\vec{v}) \geq g(c\vec{v})$.

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

Let $\alpha = \sup\{ g(\vec{v}) : \vec{v} \geq \vec{0}, \vec{v} \neq \vec{0} \}$. I will show that $\alpha$ 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, $\alpha = \sup\{ g(\vec{v}) : \vec{v} \geq \vec{0}, \|\vec{v}\| = 1 \}$. 

Doron Shahar  
Proof of the Perron-Frobenius Theorem  
May 1, 2014  
10 / 18
Defining \( \alpha \)

Let \( \alpha = \sup \{ g(\vec{v}) : \vec{v} \geq \vec{0}, \vec{v} \neq \vec{0} \} \). I will show that \( \alpha \) 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 (\( \|\vec{w}\| = \|\frac{\vec{v}}{\|\vec{v}\|}\| = \|\vec{v}\| = 1 \)) 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 \( g(\vec{v}) = \alpha \).
Theorem (part 1)

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

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

**Theorem (part 1)**

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

**Proof:**

Suppose $A\vec{v} \neq \alpha\vec{v}$. Since $g(\vec{v}) = \alpha$, $A\vec{v} \geq \alpha\vec{v}$. Thus, $A\vec{v} - \alpha\vec{v} \geq 0$ and $A\vec{v} - \alpha\vec{v} \neq \vec{0}$. So by part (a), $A(A\vec{v} - \alpha\vec{v}) > 0$. Therefore, there is a $\lambda > 0$ small enough such that $A(A\vec{v} - \alpha\vec{v}) \geq \lambda(A\vec{v})$. Then $A(A\vec{v}) \geq (\alpha + \lambda)(A\vec{v})$, and $A\vec{v} > 0$ by part (a). Thus, $g(A\vec{v}) \geq \alpha + \lambda > \alpha$. This contradicts the fact that $\alpha$ is the supremum of $g(\vec{w})$ over all nonzero $\vec{w} \geq 0$. Therefore, $A\vec{v} = \alpha\vec{v}$. QED
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 $|v_1 - v_2| \neq \vec{0}$.

A $|\vec{v}_1 - \vec{v}_2| \geq |A(\vec{v}_1 - \vec{v}_2)| = |A\vec{v}_1 - A\vec{v}_2| = |\alpha v_1 - \alpha v_2| = \alpha |\vec{v}_1 - \vec{v}_2|$.

Thus, $g(|\vec{v}_1 - \vec{v}_2|) = \alpha$. So $A|\vec{v}_1 - \vec{v}_2| = \alpha |\vec{v}_1 - \vec{v}_2|$ by 1.20(c).

Therefore, $A|\vec{v}_1 - \vec{v}_2| \geq |A(\vec{v}_1 - \vec{v}_2)|$.

Thus, the $i$th coordinate of $A|\vec{v}_1 - \vec{v}_2|$ and $|A(\vec{v}_1 - \vec{v}_2)|$ are equal:

$\sum_{j=1}^{n} a_{ij} |v_j_1 - v_j_2| = |\sum_{j=1}^{n} a_{ij} (v_j_1 - v_j_2)|$.

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(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 $|v_1 - v_2| \neq \vec{0}$.

$A|\vec{v}_1 - \vec{v}_2| \geq |A(\vec{v}_1 - \vec{v}_2)| = |A\vec{v}_1 - A\vec{v}_2| = |\alpha\vec{v}_1 - \alpha\vec{v}_2| = \alpha|\vec{v}_1 - \vec{v}_2|.$

Thus, $g(|\vec{v}_1 - \vec{v}_2|) = \alpha$. So $A|\vec{v}_1 - \vec{v}_2| = \alpha|\vec{v}_1 - \vec{v}_2|$ by 1.20(c). Therefore,

$A|\vec{v}_1 - \vec{v}_2| \geq |A(\vec{v}_1 - \vec{v}_2)| = \alpha|\vec{v}_1 - \vec{v}_2| = A|\vec{v}_1 - \vec{v}_2|.$

So $A|\vec{v}_1 - \vec{v}_2| = |A(\vec{v}_1 - \vec{v}_2)|$. 

Therefore, $\vec{v}_1 = \vec{v}_2$. QED
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 $|\nu_1 - \nu_2| \neq \vec{0}$.

Thus, $g(|\vec{v}_1 - \vec{v}_2|) = \alpha$. So $A|\vec{v}_1 - \vec{v}_2| = \alpha|\vec{v}_1 - \vec{v}_2|$ by 1.20(c). Therefore, $A|\vec{v}_1 - \vec{v}_2| \geq |A(\vec{v}_1 - \vec{v}_2)| = |\alpha \vec{v}_1 - \alpha \vec{v}_2| = \alpha|\vec{v}_1 - \vec{v}_2|$. Thus, $g(|\vec{v}_1 - \vec{v}_2|) = \alpha$. So $A|\vec{v}_1 - \vec{v}_2| = A(\vec{v}_1 - \vec{v}_2)$. Thus, the $i$th coordinate of $A|\vec{v}_1 - \vec{v}_2|$ and $|A(\vec{v}_1 - \vec{v}_2)|$ 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
Theorem (part 3)

If \( \vec{v} \geq \vec{0} \) is the unique vector with norm 1 such that \( g(\vec{v}) = \alpha \), then \( \vec{v} > \vec{0} \).
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), \( A\vec{v} > 0 \). \( A(A\vec{v}) = A(\alpha \vec{v}) = \alpha (A\vec{v}) \). So \( g(A\vec{v}) = \alpha \). Let \( \vec{w} = \frac{A\vec{v}}{\|A\vec{v}\|} \). Then \( \vec{w} > \vec{0} \), \( \|\vec{w}\| = 1 \), and by 1.20(b),
\[
g(\vec{w}) = g\left(\frac{A\vec{v}}{\|A\vec{v}\|}\right) = g(A\vec{v}) = \alpha.
\]
But \( \vec{v} \) is the unique vector satisfying these properties, so \( \vec{v} = \vec{w} > \vec{0} \). QED
Theorem (part 4)

If \( \lambda \neq \alpha \) is an eigenvalue of \( A \), then \( |\lambda| < \alpha \).

Proof: Let \( \vec{u} \) be an eigenvector that corresponds to the eigenvalue \( \lambda \).

\[ A \vec{u} = \lambda \vec{u} \]

Then \( |\vec{u}| \geq \vec{0} \) is nonzero. Furthermore,

\[ |A| |\vec{u}| \geq |A \vec{u}| = |\lambda \vec{u}| = |\lambda||\vec{u}|. \]

Therefore,

\[ |\lambda| \leq \frac{|\vec{u}|}{g(|\vec{u}|)} \leq \alpha. \]

Suppose \( |\lambda| = \alpha \). Then

\[ A |\vec{u}| \geq |A \vec{u}| = \alpha |\vec{u}| = A |\vec{u}|. \]

So

\[ A |\vec{u}| = |A \vec{u}|. \]

Thus, the \( i \)th coordinate of

\[ A |\vec{u}| \]

and

\[ |A \vec{u}| \]

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 \( \theta \) and vector \( \vec{w} \geq \vec{0} \).

Then

\[ A \vec{w} = \lambda \vec{w}, \]

so it must be that \( \lambda \geq 0 \). Hence,

\[ \lambda = \alpha \]

contradicting the fact that \( \lambda \neq \alpha \). Therefore,

\[ |\lambda| < \alpha. \]

QED
Theorem (part 4)

If \( \lambda \neq \alpha \) is an eigenvalue of \( A \), then \( |\lambda| < \alpha \).

Proof:

Let \( \vec{u} \) be a eigenvector that corresponds to the eigenvalue \( \lambda \). \( A \vec{u} = \lambda \vec{u} \).

Then \( |\vec{u}| \geq 0 \) is nonzero. Furthermore, \( A|\vec{u}| \geq |A\vec{u}| = |\lambda \vec{u}| = |\lambda||\vec{u}| \).

Therefore, \( |\lambda| \leq g(|\vec{u}|) \leq \alpha \). Suppose \( |\lambda| = \alpha \). Then

\[ A|\vec{u}| \geq |A\vec{u}| = \alpha|\vec{u}| = A|\vec{u}|. \]

So \( A|\vec{u}| = |A\vec{u}| \).
Theorem (part 4)

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

Proof:

Let $\vec{u}$ be a eigenvector that corresponds to the eigenvalue $\lambda$. $A\vec{u} = \lambda \vec{u}$.
Then $|\vec{u}| \geq 0$ is nonzero. Furthermore, $A|\vec{u}| \geq |A\vec{u}| = |\lambda \vec{u}| = |\lambda||\vec{u}|$.
Therefore, $|\lambda| \leq g(|\vec{u}|) \leq \alpha$. Suppose $|\lambda| = \alpha$. Then $A|\vec{u}| \geq |A\vec{u}| = \alpha|\vec{u}| = A|\vec{u}|$. So $A|\vec{u}| = |A\vec{u}|$. Thus, the $i$th coordinate of $A|\vec{u}|$ and $|A\vec{u}|$ 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 $\theta$ and vector $\vec{w} \geq 0$. Then $A\vec{w} = \lambda \vec{w}$, so it must be that $\lambda \geq 0$. Hence, $\lambda = \alpha$ contradicting the fact that $\lambda \neq \alpha$. Therefore, $|\lambda| < \alpha$. QED
Lemma

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

Proof:

$B^k = (b_{ij})$ is a matrix with positive entries. Therefore, we may apply 1.20(a)-(f) to $B^k$. For any nonzero $\vec{w} \geq \vec{0}$, let $h(\vec{w})$ be the largest $\lambda$ such that $B^k\vec{w} \geq \lambda\vec{w}$, and let $\beta = \sup \{ h(\vec{w}) : \vec{w} \geq \vec{0}, \vec{w} \neq \vec{0} \}$. Then there is a unique vector $\vec{w}_0 > \vec{0}$ with norm 1 such that $B^k\vec{w}_0 = \beta\vec{w}_0$. Thus $\beta$ is an eigenvalue of $B^k$, and $|\lambda| < \beta$ for all other eigenvalues $\lambda$ of $B^k$.

Let $\vec{w}_0 = (w_1, \ldots, w_{k-1}, 0, w_k, \ldots, w_{n-1})$. $\vec{w}_0 \geq \vec{0}$ and $\|\vec{w}_0\| = 1$.

If $i = k$, $[A\vec{w}_0]_i = \sum_{j=1}^n a_{ij}w_j_0 = \sum_{j \neq k} a_{ij}w_j_0 > 0 = \beta w_i_0$. And if $i \neq k$, $[A\vec{w}_0]_i = \sum_{j=1}^n a_{ij}w_j_0 = \sum_{j \neq k} a_{ij}w_j_0 = \sum_{j=1}^{n-1} b_{ij}w_j = [B^k\vec{w}_0]_i = \beta w_i_0$.

Therefore, $A\vec{w}_0 \geq \beta\vec{w}_0$, but the two are not equal. So $\beta \leq g(\vec{w}_0) \leq \alpha$. If $\beta = \alpha$, then $A\vec{w}_0$ must equal $\beta\vec{w}_0$ by 1.20(c), which is not the case. Thus, $\beta < \alpha$, and for all eigenvalues $\lambda$ of $B^k$, $|\lambda| < \beta < \alpha$. QED
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 $\alpha$.

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 $\lambda$ 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 $\beta$ is an eigenvalue of $\mathbf{B}^k$, and $|\lambda| < \beta$ for all other eigenvalues $\lambda$ of $\mathbf{B}^k$. QED
Lemma

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

Proof:

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

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 $i = k$, $[A \vec{w}_0]^i = \sum_{j=1}^n a_{ij} w^j_0 = \sum_{j \neq k} a_{ij} w^j_0 > 0 = \beta w^i_0$. And if $i \neq k$,

$[A \vec{w}_0]^i = \sum_{j=1}^n a_{ij} w^j_0 = \sum_{j \neq k} a_{ij} w^j_0 = \sum_{j=1}^{n-1} b_{ij} w^j = [B^k \vec{w}]^i = \beta w^i = \beta w^i_0$.

Therefore, $A \vec{w}_0 \geq \beta \vec{w}_0$, but the two are not equal. So $\beta \leq g(\vec{w}_0) \leq \alpha$. If $\beta = \alpha$, then $A \vec{w}_0$ must equal $\beta \vec{w}_0$ by 1.20(c), which is not the case. Thus, $\beta < \alpha$, and for all eigenvalues $\lambda$ of $B^k$, $|\lambda| < \beta < \alpha$. QED
Lemma

Let $f(\lambda) = \det(\lambda I - A)$ be the characteristic polynomial of $A$. 

\[ f'(\lambda) = \sum_{k=1}^{n} \det(\lambda I - B^k) \]
Lemma

Let \( f(\lambda) = \det(\lambda I - A) \) be the characteristic polynomial of \( A \).

\[ f'(\lambda) = \sum_{k=1}^{n} \det(\lambda I - B^k) \]

Proof:

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

\[
\frac{d}{d\lambda} f(\lambda) = \frac{d}{d\lambda} \sum_{\sigma \in S^n} \text{sgn}(\sigma) c_{\sigma(1)} \cdots c_{\sigma(n)}
\]
Lemma

Let $f(\lambda) = \det(\lambda I - A)$ be the characteristic polynomial of $A$. 

$f'(\lambda) = \sum_{k=1}^{n} \det(\lambda I - B^k)$

Proof:

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

\[
\frac{d}{d\lambda} f(\lambda) = \frac{d}{d\lambda} \sum_{\sigma \in S^n} \text{sgn}(\sigma) c_{1\sigma(1)} \cdots c_{n\sigma(n)} \\
= \sum_{\sigma \in S^n} \text{sgn}(\sigma) \sum_k (c_{1\sigma(1)} \cdots \frac{d}{d\lambda} c_{k\sigma(k)} \cdots c_{n\sigma(n)})
\]
1.20(h)

Lemma

Let \( f(\lambda) = \det(\lambda I - A) \) be the characteristic polynomial of \( A \).

\[
f'(\lambda) = \sum_{k=1}^{n} \det(\lambda I - B^k)
\]

Proof:

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

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

\[
= \sum_{\sigma \in S^n} \text{sgn}(\sigma) \sum_{k} (c_{1\sigma(1)} \cdots \frac{d}{d\lambda} c_{k\sigma(k)} \cdots c_{n\sigma(n)})
\]

\[
= \sum_{k} \sum_{\substack{\sigma \in S^n \sigma(k)=k}} \text{sgn}(\sigma) c_{1\sigma(1)} \cdots c_{(k-1)\sigma(k-1)} c_{(k+1)\sigma(k+1)} \cdots c_{n\sigma(n)}
\]

QED

Doron Shahar
Proof of the Perron-Frobenius Theorem
May 1, 2014 16 / 18
Lemma

Let \( f(\lambda) = \det(\lambda \mathbf{I} - \mathbf{A}) \) 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 \( \lambda \mathbf{I} - \mathbf{B}^k = (d_{ij}^k) \).

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

\[
= \sum_{\sigma \in S^n} \text{sgn}(\sigma) \sum_k (c_{1\sigma(1)} \cdots d_k \frac{d}{d\lambda} c_{k\sigma(k)} \cdots c_{n\sigma(n)})
\]

\[
= \sum_k \sum_{\sigma \in S^n \atop \sigma(k) = k} \text{sgn}(\sigma) c_{1\sigma(1)} \cdots c_{(k-1)\sigma(k-1)} c(k+1)_{\sigma(k+1)} \cdots c_{n\sigma(n)}
\]

\[
= \sum_k \sum_{\sigma \in S^{n-1}} \text{sgn}(\sigma) d_{1\sigma(1)}^k \cdots d_{n-1\sigma(n-1)}^k = \sum_{k=1}^{n} \det(\lambda \mathbf{I} - \mathbf{B}^k) \quad \text{QED}
\]
Theorem (part 5)

\[ f'(\alpha) > 0. \text{ Hence, } \alpha \text{ is a simple eigenvalue for } A. \]
Theorem (part 5)

\[ f'(\alpha) > 0. \quad \text{Hence, } \alpha \text{ is a simple eigenvalue for } A. \]

Proof:

\[ \det(\lambda I - B^k) \text{ is polynomial in } \lambda \text{ of degree } n - 1 \text{ with leading term } \lambda^{n-1} \]
Therefore, \( \det(\lambda I - B^k) > 0 \) for all \( \lambda \) greater than the largest real eigenvalue of \( B^k \). By 1.20(g), \( \alpha \) is greater than the largest real eigenvalue of \( B^k \). Therefore, \( f'(\alpha) = \sum_{k=1}^{n} \det(\alpha I - B^k) > 0. \)
Theorem (part 5)

\( f'(\alpha) > 0. \) Hence, \( \alpha \) is a simple eigenvalue for \( A \).

Proof:

\[
\det(\lambda I - B^k) \text{ is polynomial in } \lambda \text{ of degree } n - 1 \text{ with leading term } \lambda^{n-1}.
\]

Therefore, \( \det(\lambda I - B^k) > 0 \) for all \( \lambda \) greater than the largest real eigenvalue of \( B^k \). By 1.20(g), \( \alpha \) is greater than the largest real eigenvalue of \( B^k \). Therefore, \( f'(\alpha) = \sum_{k=1}^{n} \det(\alpha I - B^k) > 0. \)

\( f(\lambda) = (\lambda - \alpha)^m q(\lambda) \) where \( q \) is a nonzero polynomial whose roots have absolute value less than \( \alpha \). \( m \in \mathbb{N} \) is the multiplicity of \( \alpha \). 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, \( \alpha \) is a root of the characteristic polynomial of \( A \) with multiplicity 1. That is, \( \alpha \) is a simple eigenvalue for \( A \). QED
Let $A = (a_{ij})$ be an $n \times n$ matrix with $a_{ij} > 0$ for all $i, j$. $A$ 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 $\lambda$ of $A$. Moreover, $\alpha$ is a simple eigenvalue (i.e., it is a root of the characteristic polynomial of $A$ with multiplicity 1).

1.20(c) shows that $\alpha > 0$ is an eigenvalue of $A$.
1.20(d) shows that there is $\alpha$ 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 $\lambda$ of $A$, $|\lambda| < \alpha$.
1.20(i) shows that $\alpha$ is a simple eigenvalue of $A$. 