The Perron–Frobenius Theorem: Statement

Theorem 1 (The Perron–Frobenius Theorem)**.** *Let $A \in \mathbb{R}^{n\times n}_{\ge 0}$, and assume that $A^K > 0$ for some $K \ge 1$. Then there is a unique $\lambda_* > 0$ and a vector $\mathbf{y}_* \in \mathbb{R}^n_{>0}$ that is unique up to scaling, with the following properties:*

1. (Root convergence) *For every nonzero start vector $\mathbf{x}^{(0)} \in \mathbb{R}^n_{\ge 0}$,*
$$\lim_{k\to\infty} \sqrt[k]{\|A^k \mathbf{x}_0\|} = \lambda_*,$$
where $\|\cdot\|$ is any norm on $\mathbb{R}^n$.

2. (Ratio convergence) *For every nonzero start vector $\mathbf{x}^{(0)} \in \mathbb{R}^n_{\ge 0}$, and every $i = 1, \dots, n$,*
$$\lim_{k\to\infty} \frac{(A^{k+1}\mathbf{x}^{(0)})_i}{(A^k\mathbf{x}^{(0)})_i} = \lambda_*.$$

3. *λ_* is the eigenvalue of A with the largest absolute value, and A has no other eigenvalue with this absolute value.*

4. *$\mathbf{y}_*$ is an eigenvector with eigenvalue λ_*, and it is the only eigenvector with this eigenvalue.*
In other words, λ_* is an eigenvalue of geometric multiplicity 1. (*The algebraic multiplicity can be higher.*)

5. *$\mathbf{y}_*$ is the only nonnegative eigenvector.*

6. (The Collatz-Wielandt inequalities) *Let $\mathbf{u} \in \mathbb{R}^n_{>0}$ be any positive vector. Determine numbers $\underline{\lambda}$ and $\overline{\lambda}$ such that the following inequalities hold:*
$$\underline{\lambda}\mathbf{u} \le A\mathbf{u} \le \overline{\lambda}\mathbf{u}$$
Then $\underline{\lambda} \le \lambda_* \le \overline{\lambda}$.

Proof

We start with an easy observation:

Lemma 1. *If B is a positive matrix, then $\mathbf{u} \ge \mathbf{v}$ and $\mathbf{u} \ne \mathbf{v}$ implies $B\mathbf{u} > B\mathbf{v}$.*

Proof. For all $i = 1, \dots, n$:
$$\sum_j b_{ij}(u_j - v_j) > 0,$$
because $b_{ij} > 0$ and $u_j - v_j \ge 0$ for all j, and $u_j - v_j > 0$ for at least one j. □

Definition 1. *For a nonnegative nonzero vector $\mathbf{x} \in \mathbb{R}^n$, we define*
$$\lambda_{\min}(\mathbf{x}) := \max\{\, \lambda \mid A\mathbf{x} \ge \lambda\mathbf{x} \,\} = \min_{1\le i\le n} \frac{(A\mathbf{x})_i}{x_i}.$$
The vector $\mathbf{x}$ may contain zeros, and in this case, fractions with denominator 0 *in the last term are interpreted as $+\infty$.*

Boundedness. We show that $\lambda_{\min}(\mathbf{x})$ is bounded by the global upper bound
$$M := \max_{1\le i\le n} \sum_{j=1}^n a_{ij}.$$

Proof. Let x_{i_0} be the largest entry of $\mathbf{x}$. Then, for every i, $(A\mathbf{x})_i = \sum_{j=1}^n a_{ij}x_j \le \sum_{j=1}^n a_{ij}x_{i_0} \le Mx_{i_0}$, and hence $\lambda_{\min}(\mathbf{x}) = \min_{1\le i\le n} \frac{(A\mathbf{x})_i}{x_i} \le \frac{(A\mathbf{x})_{i_0}}{x_{i_0}} \le M$. □

Iteration. Take an arbitrary start vector $\mathbf{x}^{(0)} \in \mathbb{R}^n$ with positive entries and consider the iteration $\mathbf{x}^{(1)} = A\mathbf{x}^{(0)}$, $\mathbf{x}^{(2)} = A\mathbf{x}^{(1)}$, etc. Let $\lambda^{(k)} := \lambda_{\min}(\mathbf{x}^{(k)})$.

Then $\lambda^{(k)} \leq \lambda^{(k+1)}$ (easy induction exercise), and since $\lambda^{(k)} \leq M$, we can form the limit

$$\tilde{\lambda} := \lim_{k\to\infty} \lambda^{(k)}.$$

Consider the sequence of normalized vectors:

$$\mathbf{y}^{(k)} := \frac{\mathbf{x}^{(k)}}{\|\mathbf{x}^{(k)}\|_1}$$

Lemma 2. *Every accumulation point $\tilde{\mathbf{y}}$ of the sequence $\mathbf{y}^{(0)}, \mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \ldots$ is an eigenvector of A with eigenvalue $\tilde{\lambda}$.*

Proof. Consider a subsequence $\mathbf{y}^{(k_1)}, \mathbf{y}^{(k_2)}, \mathbf{y}^{(k_3)}, \ldots$ converging to some accumulation point $\tilde{\mathbf{y}}$. Then the subsequence $\mathbf{y}^{(k_1+1)}, \mathbf{y}^{(k_2+1)}, \mathbf{y}^{(k_3+1)}, \ldots$ converges also, namely to $A\tilde{\mathbf{y}}/\|A\tilde{\mathbf{y}}\|_1$. More generally, the subsequence $\mathbf{y}^{(k_1+K)}, \mathbf{y}^{(k_2+K)}, \mathbf{y}^{(k_3+K)}, \ldots$ converges to $A^K\tilde{\mathbf{y}}/\|A^K\tilde{\mathbf{y}}\|_1$.

By definition,

$$A\mathbf{y}^{(k_i)} \geq \lambda^{(k_i)}\mathbf{y}^{(k_i)}$$

for all i. Taking the limit, we get

$$A\tilde{\mathbf{y}} \geq \tilde{\lambda}\tilde{\mathbf{y}}$$

The claim of the lemma is that this inequality holds as an equation. Suppose not. Lemma 1 implies

$$A^K(A\tilde{\mathbf{y}}) > \tilde{\lambda}A^K\tilde{\mathbf{y}},$$

or more specifically,

$$A^K(A\tilde{\mathbf{y}}) = A(A^K\tilde{\mathbf{y}}) \geq (\tilde{\lambda}+\varepsilon)(A^K\tilde{\mathbf{y}})$$

for some $\varepsilon > 0$. Equivalently,

$$A\frac{A^K\tilde{\mathbf{y}}}{\|A^K\tilde{\mathbf{y}}\|_1} \geq (\tilde{\lambda}+\varepsilon)\frac{A^K\tilde{\mathbf{y}}}{\|A^K\tilde{\mathbf{y}}\|_1} \tag{1}$$

Since $A^K\tilde{\mathbf{y}}/\|A^K\tilde{\mathbf{y}}\|_1$ is the limit of the sequence $\mathbf{y}^{(k_1+K)}, \mathbf{y}^{(k_2+K)}, \mathbf{y}^{(k_3+K)}, \ldots$ this implies that there is some element $\mathbf{y}^{(k_i+K)}$ in this sequence which is close enough to the limit $A^K\tilde{\mathbf{y}}/\|A^K\tilde{\mathbf{y}}\|_1$ such that (1) holds with a small error $\varepsilon/2$:

$$A\mathbf{y}^{(k_i+K)} \geq (\tilde{\lambda}+\varepsilon-\varepsilon/2)\mathbf{y}^{(k_i+K)},$$

contradicting the fact that $\lambda_{\max}(\mathbf{y}^{(k_i+K)}) = \lambda^{(k_i+K)} \leq \tilde{\lambda}$. □

Uniqueness of the eigenvalue. Every nonnegative eigenvector is positive.

Proof. $A\mathbf{y} = \lambda\mathbf{y}$ implies $A^K\mathbf{y} = \lambda^K\mathbf{y}$, and by Lemma 1, $A^K\mathbf{y}$ is positive. □

All nonnegative eigenvectors have the same eigenvalue, which we denote by λ_*.

Proof. Assume $A\mathbf{y}_1 = \lambda_1\mathbf{y}_1$ and $A\mathbf{y}_2 = \lambda_2\mathbf{y}_2$. Since $\mathbf{y}_2$ is positive, we can, by rescaling $\mathbf{y}_1$ if necessary, assume that $\mathbf{y}_1 \leq \mathbf{y}_2$. It follows that

$$\lambda_1^k\mathbf{y}_1 = A^k\mathbf{y}_1 \leq A^k\mathbf{y}_2 = \lambda_2^k\mathbf{y}_2$$

for all k, and this implies $\lambda_1 \leq \lambda_2$. The reverse inequality follows in the same way. □

The Collatz-Wielandt inequalities.

Lemma 3. *1. If $A\mathbf{y} \geq \underline{\lambda}\mathbf{y}$ for some nonnegative vector $\mathbf{y}$, then $\underline{\lambda} \leq \lambda_*$.*
In other words, $\lambda_{\min}(\mathbf{y}) \leq \lambda_*$ for all nonnegative vectors $\mathbf{y}$.
2. If $A\mathbf{y} \leq \overline{\lambda}\mathbf{y}$ for some nonnegative vector $\mathbf{y}$, then $\overline{\lambda} \geq \lambda_*$.

Proof. If we start the iteration with $\mathbf{x}^{(0)} = \mathbf{y}$, then the assumption means that $\lambda^{(0)} = \lambda_{\min}(\mathbf{x}^{(0)}) \geq \overline{\lambda}$, and since the sequence $\lambda^{(k)}$ converges monotonically to λ_*, $\overline{\lambda} \leq \lambda^{(0)} \leq \lambda_*$.

The second inequality follows in an analogous way, using instead of $\lambda_{\min}$ the quantity

$$\lambda_{\max}(\mathbf{x}) := \min\{\, \lambda \mid A\mathbf{x} \leq \lambda\mathbf{x} \,\} = \max_{1\leq i\leq n} \frac{(A\mathbf{x})_i}{x_i},$$

and arguing that the sequence $\lambda_{\max}(\mathbf{x}^{(k)})$ decreases monotonically. □

Corollary 4. *There can be no nonnegative vector $\mathbf{x}$ with $A\mathbf{x} > \lambda^*\mathbf{x}$.* □

Uniqueness among nonnegative eigenvectors (claim 5). Suppose $\mathbf{y}_1$ and $\mathbf{y}_2$ are two nonnegative vectors with eigenvalue λ_*.

$$A\mathbf{y}_1 = \lambda_*\mathbf{y}_1,$$
$$A\mathbf{y}_2 = \lambda_*\mathbf{y}_2.$$

Iterating K times, we get

$$A^K\mathbf{y}_1 = \lambda_*^K\mathbf{y}_1, \tag{2}$$
$$A^K\mathbf{y}_2 = \lambda_*^K\mathbf{y}_2. \tag{3}$$

We assume for contradiction that $\mathbf{y}_2$ is not a scalar multiple of $\mathbf{y}_2$. Then, by rescaling the vectors, we can assume that neither $\mathbf{y}_1 \leq \mathbf{y}_2$ nor $\mathbf{y}_1 \geq \mathbf{y}_2$ holds. The elementwise maximum $\hat{\mathbf{y}} := \max(\mathbf{y}_1, \mathbf{y}_2)$ is therefore a vector different from $\mathbf{y}_1$ and $\mathbf{y}_2$. Lemma 1 implies

$$A^K\hat{\mathbf{y}} > A^K\mathbf{y}_1 = \lambda_*^K\mathbf{y}_1,$$
$$A^K\hat{\mathbf{y}} > A^K\mathbf{y}_2 = \lambda_*^K\mathbf{y}_2,$$

and therefore

$$A^K\hat{\mathbf{y}} > \max(\lambda_*^K\mathbf{y}_1, \lambda_*^K\mathbf{y}_2) = \lambda_*^K\hat{\mathbf{y}}.$$

On the other hand, by applying the previous arguments to A^K instead of A, We conclude from (2) and (3) that λ_*^K is the unique eigenvalue for A^K with a positive eigenvector. Therefore, the Collatz-Wielandt inequality for A^K implies that $A^K\hat{\mathbf{y}} > \lambda_*^K\hat{\mathbf{y}}$ is impossible. □

We denote the unique nonnegative eigenvector of A by $\mathbf{y}_*$.

Uniqueness among arbitrary (complex) eigenvectors (claims 3 and 4). Let $\mathbf{z} \in \mathbb{C}^n$ be an eigenvector with eigenvalue λ, or in other words, for every i, $\lambda z_i = (A\mathbf{z})_i = \sum_{j=1}^n a_{ij}z_j$. Taking absolute values, we get

$$|\lambda| \cdot |z_i| = \left|\sum_{j=1}^n a_{ij}z_j\right| \leq \sum_{j=1}^n |a_{ij}z_j| = \sum_{j=1}^n a_{ij}|z_j|. \tag{4}$$

In other words, the nonnegative vector $\mathbf{x}$ with $x_i = |z_i|$ fulfills $A\mathbf{x} \geq |\lambda|\mathbf{x}$, and it follows from Lemma 3 that $|\lambda| \leq \lambda_*$.

Let us discuss the case of equality, $|\lambda| = \lambda_*$. Then $A\mathbf{x} \geq |\lambda_*|\mathbf{x}$. This can only hold with equality. (Otherwise, Lemma 1 would imply that $A(A^K\mathbf{x}) > |\lambda_*|A^K\mathbf{x}$, contradicting Corollary 4.) Therefore $\mathbf{x}$ must be a multiple of the unique nonnegative eigenvector $\mathbf{y}_*$. In the application of the triangle inequality in (4), equality holds only if $\mathbf{z}$ is a (complex) multiple of $\mathbf{x}$. It follows that there is no eigenvector with eigenvalue λ_* except $\mathbf{y}_*$ and its scalar multiples.

Ratio convergence (claim 2). The two vectors involved in the quotient are $A^{k+1}\mathbf{x}^{(0)} = \mathbf{x}^{(k+1)} = A\mathbf{x}^{(k)}$ and $A^k\mathbf{x}^{(0)} = \mathbf{x}^{(k)}$. Nothing is changed if we replace these vectors by the scaled vectors $A\mathbf{y}^{(k)}$ and $\mathbf{y}^{(k)}$. Since $\mathbf{y}^{(k)}$ converges to the eigenvector $\mathbf{y}_*$, the "elementwise ratio" between $\mathbf{y}^{(k)}$ and $A\mathbf{y}^{(k)}$ converges to λ_*.

It does not matter if the start vector $\mathbf{x}^{(0)}$ is not positive. The vectors $\mathbf{x}^{(k)}$ will be positive after at most K steps, and by cutting out the first K steps, we obtain the same iteration with a positive start vector.

Root convergence of the individual entries, $\lim_{k\to\infty} \sqrt[k]{(A^k\mathbf{x}^{(0)})_i} = \lambda_*$, is a direct consequence, and thus it holds also for the norm (claim 1).

Alternative constructions of the eigenvector. We have in fact shown (Lemma 3) that $\mathbf{y}_*$ is the vector that maximizes $\lambda_{\min}(\mathbf{x})$ among all nonnegative vectors. Alternatively, it can be defined as the vector that minimizes $\lambda_{\max}(\mathbf{x})$ among all nonnegative vectors.