4.3 Perron Theorem

Say a real $n \times n$ matrix M is a Perron matrix if it is entrywise positive, namely $M_{ij} > 0$ for all i, j.

Theorem 42 (Perron [50]). *A Perron matrix M has a "Perron eigenvalue" $\bar{\lambda}$ and a left "Perron eigenvector" v with the following properties:*

1. *$\bar{\lambda} > 0$ and has a positive eigenvector v, i.e., $v_i > 0$ for all i.*
2. *$\bar{\lambda}$ is algebraically simple (which is stronger than saying it is geometrically simple). That is, $\bar{\lambda}$ is a simple root of the characteristic polynomial of M.*
3. *M has no other left eigenvector with nonnegative entries.*
4. *For any other eigenvalue κ of M, $|\kappa| < \bar{\lambda}$.*

Of course, the same assertions hold with respect to right eigenvectors.

We follow an argument credited to H. Bohnenblust.

Proof. For nonzero vector $x \geq 0$, let

$$\lambda(x) = \sup\{\lambda : xM \geq \lambda x\}.$$

where a vector inequality means that the inequality holds in every coordinate.

Let $\Delta = \{x : x \geq 0, \sum x_i = 1\}$. Observe that $\lambda(x)$ is invariant on rays from the origin, so it achieves all its values on Δ. Moreover (exercise) it is continuous on Δ. Let $L(M) = \text{image}\,\lambda = \{\lambda(x) : x \in \Delta\}$.

Lemma 43. *For M Perron, $L(M)$ is a closed, nonempty, bounded subinterval of $(0, \infty)$.*

Proof. Since λ is continuous on compact Δ, the image is closed, bounded and connected.

For the lower bound, note that for nonzero $x \geq 0$, $(xM)_j \geq (\max x_i)(\min M_{ij}) \geq x_j(\min M_{ij})$, so image $\lambda \subseteq [\min_{i,j} M_{ij}, \infty)$.

For a specific upper bound, let $b = \max_i (M\mathbb{1})_i$, where $\mathbb{1}$ is the all-ones column vector. Then

$$bx\mathbb{1} \geq xM\mathbb{1} \geq \lambda_x x\mathbb{1}$$

so $\sup L(M) \leq b$. □

Let $\bar{\lambda} = \max L(M)$. As noted by compactness there is a $v \in \Delta$ achieving $vM \geq \bar{\lambda}v$. We claim v is an eigenvector with eigenvalue $\bar{\lambda}$. This is a corollary of:

Lemma 44. *For $x \in \Delta$, $\lambda(xM) \geq \lambda(x)$, with equality only if x is an eigenvector.*

Proof. Suppose x is not an eigenvector, then there is a j s.t. $(xM)_j > \lambda(x)x_j$. Then for sufficiently small $\varepsilon > 0$, and with e_j denoting the singleton vector,

$$xM \geq \lambda(x)x + \varepsilon e_j.$$

Then (and using that any entry of x is ≤ 1):

$$\begin{aligned} ((x+\varepsilon e_j)M)_j &\geq (xM)_j && \geq \lambda(x)x_j + \varepsilon && \geq (\lambda(x)+\varepsilon)x_j \\ ((x+\varepsilon e_j)M)_k &\geq \lambda(x)x_k + \varepsilon M_{jk} && \geq (\lambda(x)+\varepsilon M_{jk})x_k && \text{for any } k \neq j \end{aligned}$$

Consequently $\lambda(x+\varepsilon e_j) \geq \lambda(x) + \varepsilon \min_k M_{jk} > \lambda(x)$. □

Lemma 45. *A nonnegative eigenvector u of M with positive eigenvalue λ is actually positive.*

Proof. $u_k = \frac{1}{\lambda}(uM)_k = \frac{1}{\lambda}\sum u_i M_{ik} > 0$, because all $M_{ik} > 0$. □

Consequently, v is positive. We have shown part 1 of the theorem with the specified $\bar{\lambda}$ and v.

Now for part 2: if this is false then there is a vector y, linearly independent of v, s.t. for some c,

$$yM = \bar{\lambda}y + cv \tag{4.2}$$

Since $\bar{\lambda}$ is real, both y and c can be taken real.

Now consider the vector $ay + v$ for a real:

$$(ay + v)M = \bar{\lambda}(ay + v) + acv$$

Choose a s.t. $ac \geq 0$ and $|a|$ is as large as possible while maintaining $ay + v \geq 0$. Then $\lambda(ay + v) \geq \bar{\lambda}$ which by Lemma 44 forces $ay + v$ to be an eigenvector; however, since $ay + v \not> 0$, this is contradiction to Lemma 45.

Now to part 3: Let u be any other eigenvector of M, and let κ be its eigenvalue; from part 2 we know that $\kappa \neq \bar{\lambda}$. By part 1 applied to M^*, M has a right eigenvector v' with eigenvalue $\bar{\lambda}$, and $v' > 0$. (As usual * denotes conjugate transpose (or what analysts call adjoint); here it is simply transpose.)

Next, we claim that the inner product uv' equals 0. But then since $v' > 0$, we cannot have $u \geq 0$. The claim comes from a simple but very useful lemma that we note for future reference:

Lemma 46. *If matrix A has left eigenvector y for eigenvalue η, and right eigenvector z for eigenvalue ζ, with $\eta \neq \zeta$, then $yz = 0$.*

Proof. $\eta yz = (yA)z = y(Az) = \zeta yz$ which is only possible if $yz = 0$. □

Part 4: Observe that if $b_1, \ldots, b_n \in \mathbb{C}$ then $|\sum b_i| \leq \sum |b_i|$ with equality holding only if all nonzero b_i share a common phase, i.e., for all nonzero b_i, b_j, b_i/b_j is a positive real.

Now consider any eigenvalue $\kappa \neq \bar{\lambda}$, and let u be a corresponding eigenvector. Of course, for any j, $|(uM)_j| = |\kappa| \cdot |u_j|$. From part 3 we know that u cannot be nonnegative, and this is equivalent (since scalar multiples of eigenvectors are eigenvectors) to saying that u must possess some nonzero coordinates $u_i, u_{i'}$ such that $u_i/u_{i'}$ is not positive. For any j then the same holds for the vector $(u_1 M_{1j}, \ldots, u_n M_{nj})$. Consequently $|\sum_i u_i M_{ij}| < \sum |u_i| M_{ij}$.

Also, since $(|u_1|, \ldots, |u_n|) \geq 0$, there is a j for which $\sum |u_i| M_{ij} \leq \bar{\lambda}|u_j|$. So for this j

$$|\kappa| = \frac{|(uM)_j|}{|u_j|} = \frac{|\sum u_i M_{ij}|}{|u_j|} < \frac{\sum |u_i| M_{ij}}{|u_j|} \leq \frac{\bar{\lambda}|u_j|}{|u_j|} = \bar{\lambda}.$$

□

4.3.1 Perron-Frobenius and Markov Chains

Let M be a square matrix with nonnegative entries. We say that M is *irreducible* if $\forall i, j\ \exists k : (M^k)_{ij} > 0$. We say that M is *aperiodic* if $\forall i, j : \gcd\{k : (M^k)_{ij} > 0\} = 1$. In exercise ps3-5 you are asked to show that if M is irreducible and aperiodic then there is a k such that M^k is entrywise positive.

Theorem 47 (Perron-Frobenius [21])**.** *If M is entrywise nonnegative, irreducible and aperiodic, then all the conclusions of Theorem 42 go through.*

(This is assigned as ps3-6.)

We say that a probability distribution π over the states of a Markov chain M is stationary if $\pi M = \pi$.

Corollary 48 (Fundamental Theorem of Markov Chains)**.** *If M is an irreducible, aperiodic Markov chain then $\lim_{t\to\infty} M^t = \mathbb{1} \otimes \pi$ for some $\pi \in \Delta$. In particular π is the unique stationary distribution of M.*

It is worth commenting on what happens if M is irreducible but periodic with period k. Then M^k is reducible and decomposes into k irreducible aperiodic components. Let π be the unique stationary distribution of M^k on one of those components. Then M has a unique stationary distribution which equals $\frac{1}{k}\sum_{i=0}^{k-1} \pi M^i$. However, $\lim_{t\to\infty} M^t$ does not exist.