THE PERRON-FROBENIUS THEOREM

We state and prove here a simplified version of the Perron-Frobenius Theorem, that has manifold applications.

Theorem 0.1. *Let A be an $n \times n$ matrix with strictly positive matrix elements. There exists an eigenvalue $\lambda_{\max} > 0$ which is not degenerate and whose eigenvector $x_{\max}$ has strictly positive components. Moreover, any other eigenvector x of A with non-negative entries is equal to $x_{\max}$. Further, if λ is any other eigenvalue of A (which may be complex), then $|\lambda| \le \lambda_{\max}$.*

Let Q be the positive orthant in $\mathbb{R}^n$. Let $x \in Q$, $x \neq 0$ and set

$$E(x) = \min_{i, x_i \neq 0} \frac{(Ax)_i}{x_i} \ .$$

We first start with a lemma.

Lemma 0.2. *The function $E(x)$ is bounded, in fact*

$$E(x) \le \max_k \sum_i a_{ki} \ .$$

Further, for any $x \in Q, x \neq 0$ we have that

$$E(Ax) \ge E(x)$$

with equality only if x satisfies $Ax = E(x)x$, i.e., x is an eigenvector.

Proof. The definition of $E(x)$ is equivalent with the statement that $E(x)$ is the largest number such that

$$(Ax)_i - E(x)x_i \ge 0 \ , i = 1, \dots, n \ .$$

To see that

$$E(Ax) \ge E(x) \ ,$$

simply note that

$$[A(Ax - E(x)x)]_i \ge 0 \ , i = 1, \dots, n$$

since $(Ax)_i - E(x)x_i \ge 0, i = 1, \dots, n$. This means that

$$E(Ax) = \min_i \frac{(A^2x)_i}{(Ax)_i} \ge E(x) \ .$$

Note that since A has strictly positive elements and $x \in Q$ is not the zero vector we have that Ax has strictly positive components. Now, suppose that $x \in Q$ is not an eigenvector of A. Then, as noted above, $(Ax)_i - E(x)x_i \ge 0, i = 1, \dots, n$ and not all of them are equal to zero. Hence

$$[A(Ax - E(x)x)]_i > 0 \ , i = 1, \dots, n$$

with a strict inequality. Thus $E(Ax) > E(x)$.

Finally to see that $E(x)$ is bounded, note that for any given $x \in Q, x \neq 0$ we have that

$$\min_{i, x_i \neq 0} \frac{(Ax)_i}{x_i} \le \frac{(Ax)_k}{x_k}$$

where $x_k = \max_i x_i$. But then $(Ax)_k = \sum_i a_{ki} x_i \le (\sum_i a_{ki}) x_k$ so that

$$\frac{(Ax)_k}{x_k} \le (\sum_i a_{ki}) \ .$$

□

To gather some more information about $E(x)$ we have to introduce the set Σ which is the intersection of Q and the unit sphere in $\mathbb{R}^n$. The next step is the analytical part of the proof. The key fact from analysis is that if a function is continuous on a compact set, then it attains its maximum and minimum on that set. The problem is that $E(x)$ may not be continuous on the set Σ. Since the ratio $\frac{(Ax)_i}{x_i}$ is only defined for i with x_i non-zero. What can happen is that $\min_i \frac{(Ax)_i}{x_i}$ could jump as one varies the vector $x \in Q$ and one or more components vanish during that process. Thus E is not continuous on Σ, but it is on $A(\Sigma)$ as we shall show next.

Lemma 0.3. *Consider the set $A(\Sigma)$, i.e., the image of the set Σ under A. This set is closed and bounded. Moreover, on this set $A(\Sigma)$ the function $E(x)$ is continuous.*

Proof. The set Σ is closed and bounded and hence compact. Multiplication by a matrix is a continuous operation and hence $A(\Sigma)$ is also compact. Since A has strictly positive matrix elements, we have that

$$\min_{x \in \Sigma} (Ax)_i \ge m_i > 0, \ i = 1, \dots, n \ .$$

Hence E is continuous on $A(\Sigma)$. □

Proof of the Theorem. The function E being continuous on $A(\Sigma)$ and $A(\Sigma)$ being compact, attains its maximum on $A(\Sigma)$. Moreover, since $E(Ax) \ge E(x)$ it also attains its maximum on Σ. Lets denote a vector where the maximum attained by z. We have that $E(Az) \ge E(z)$ and since $E(z)$ is the maximum, we have that $E(Az) = E(z)$. By the first lemma the vector z must be an eigenvector. Moreover z has strictly positive components because $Az = E(z)z$ and the components of Az are strictly positive. We set $\lambda_{\max} := E(z)$. Note that $\lambda_{\max}$ is strictly positive, for otherwise A would be the zero matrix.

Now let y be any eigenvector of A with eigenvalue λ which could be complex. Then taking absolute values in the equation $\lambda y_i = \sum_j a_{ij} y_j$ we get using the triangle inequality

$$|\lambda| |y_i| \le \sum_j a_{ij} |y_j| \ .$$

Hence

$$|\lambda| \le \min_{i, |y_i| \ne 0} \frac{\sum_j a_{ij} |y_j|}{|y_i|} = E(|y|) \le \lambda_{\max}$$

where we denote by $|y|$ the vector that has the components $|y_i|, i = 1, \dots, n$. This proves that any eigenvalue λ must satisfy $|\lambda| \le \lambda_{\max}$.

We have to show that $\lambda_{\max}$ has geometric multiplicity one. Suppose that the exists $x \in Q, x \ne 0$ such that $Ax = \lambda_{\max} x$. Then x must have strictly positive components. If x and z are not proportional they span a two dimensional space and hence there are number a, b so that the vector $y := az + bx$ has a zero component. Then $Ay = \lambda_{\max} y$ and as before taking magnitudes

$$\lambda_{\max} |y_i| \le \sum_j a_{ij} |y_j|$$

and we conclude as before that

$$\lambda_{\max} \le E(|y|) \le \lambda_{\max}$$

and hence there must be equality. Because $\lambda_{\max} \ge E(A|y|) \ge E(|y|) = \lambda_{\max}$ the vector $|y|$ must be an eigenvector, i.e., $A|y| = \lambda_{\max}|y|$. But, this means that all the components of $|y|$ must be strictly positive contradicting the fact that y has a zero component.

What is left is to show that if $Ax = \lambda x$ and x has non-negative components, then $\lambda = \lambda_{\max}$ and x is a positive multiple of z. To prove this, we consider the transpose A^T which also has strictly positive matrix elements. Hence we may apply the same reasoning and find a vector w with strictly positive entries such that $A^T w = \mu w, \mu > 0$. The claim is that $\mu = \lambda_{\max}$. To see this we compute

$$\mu w^T z = (A^T w)^T z = w^T A z = \lambda_{\max} w^T z$$

and since $w^T z > 0$ we have that $\mu = \lambda_{\max}$. Let x be any eigenvector of A with non-negative entries, i.e., $(Ax)_i = \lambda x_i, i = 1, \dots, n$. Then by the same reasoning, using that w has strictly positive components, we find that $w^T x > 0$ and hence

$$\lambda_{\max} = \lambda \ .$$

Since the eigenvalue $\lambda_{\max}$ has geometric multiplicity one, the vector x must be a positive multiple of z. □

We assumed that the matrix elements a_{ij} are strictly positive and we can say a bit more.

Theorem 0.4. *Suppose that $Ay = \lambda y$ and $|\lambda| = \lambda_{\max}$. Here λ may be complex and and y a complex vector. Then $y = cz$ where $c \neq 0$ is in general a complex number. In other words, if $Ay = \lambda y$ and y is not proportional to z then $|\lambda| < \lambda_{\max}$.*

Proof. The reasoning is as before. We have

$$\lambda_{\max}|y_i| = |\lambda||y_i| \le \sum_j a_{ij}|y_j|$$

from which we conclude as before that the vector $|y|$ having the components $|y_i|$ is an eigenvector with non-negative entries with eigenvalue $\lambda_{\max}$ and hence proportional to z. Hence we must have the equality

$$|\sum_j a_{ij} y_j| = \sum_j a_{ij}|y_j| \ , i = 1, \dots, n \ .$$

The rest follows by an inductive application of the simple lemma below. □

Lemma 0.5. *Let $a, b > 0$. Then*

$$|a + e^{i\phi} b| = |a + b|$$

implies that $e^{i\phi} = 1$.

Proof. We compute

$$|a + e^{i\phi} b|^2 = a^2 + b^2 + 2ab\Re e^{i\phi} = a^2 + b^2 + 2ab$$

from which we get that $\Re e^{i\phi} = 1$. Since $|e^{i\phi}| = 1$, we have that $e^{i\phi} = 1$.

□