1 The Perron-Frobenius Theorem.

We say that a real matrix T is **non-negative** (or **positive**) if all the entries of T are non-negative (or positive). We write $T \geq 0$ or $T > 0$. We will use these definitions primarily for square ($n \times n$) matrices and for column vectors ($n \times 1$ matrices). We let

$$Q := \{x \in \mathbf{R}^n : x \geq 0, \quad x \neq 0\}$$

so Q is the non-negative "orthant" excluding the origin. Also let

$$C := \{x \geq 0 : \|x\| = 1\}.$$

So C is the intersection of the orthant with the unit sphere.

A non-negative matrix square T is called **primitive** if there is a k such that all the entries of T^k are positive. It is called **irreducible** if for any i, j there is a $k = k(i,j)$ such that $(T^k)_{ij} > 0$. If T is irreducible then $I + T$ is primitive. Until further notice in this section we will assume that T is non-negative and irreducible.

Theorem 1 Perron-Frobenius,1. *T has a positive (real) eigenvalue $\lambda_{\max}$ such that all other eigenvalues of T satisfy*

$$|\lambda| \leq \lambda_{\max}.$$

Furthermore $\lambda_{\max}$ has algebraic and geometric multiplicity one, and has an eigenvector x with $X > 0$. Finally any non-negative eigenvector is a multiple of x. More generally, if $y \geq 0,\ y \neq 0$ is a vector and μ is a number such that

$$Ty \leq \mu y$$

then

$$y > 0, \quad \text{and} \quad \mu \geq \lambda_{\max}$$

with $\mu = \lambda_{\max}$ if and only if y is a multiple of x.

If $0 \leq S \leq T,\ S \neq T$ then every eigenvalue σ of S satisfies $|\sigma| < \lambda_{\max}$. In particular, all the diagonal minors T_i obtained from T by deleting the i-th row and column have eigenvalues all of which have absolute value $< \lambda_{\max}$.

Proof. Let

$$P := (I + T)^{n-1}$$

and for any $z \in Q$ let

$$L(z) := \max\{s : sz \leq Tz\} = \min_{1 \leq i \leq n, z_i \neq 0} \frac{(Tz)_i}{z_i}.$$

By definition $L(rz) = L(z)$ for any $r > 0$, so $L(z)$ depends only on the ray through z. If $z \leq y,\ z \neq y$ we have $Pz < Py$. Also $PT = TP$. So if $sz \leq Tz$ then

$$sPz \leq PTz = TPz$$

so

$$L(Pz) \geq L(z).$$

Furthermore, if $L(z)z \neq Tz$ then $L(z)Pz < TPz$. So $L(Pz) > L(z)$ unless z is an eigenvector of T. Consider the image of C under P. It is compact (being the image of a compact set under a continuous map) and all of the elements of $P(C)$ have all their components strictly positive (since P is positive). Hence the function L is continuous on $P(C)$. Thus L achieves a maximum value, $L_{\max}$ on $P(C)$. Since $L(z) \leq L(Pz)$ this is in fact the maximum value of L on all of Q, and since $L(Pz) > L(z)$ unless z is an eigenvector of T, we conclude that $L_{\max}$ is achieved at an eigenvector, call it x of of T and $x > 0$ with $L_{\max}$ the eigenvalue. Since $Tx > 0$ and $Tx = L_{\max}x$ we have $L_{max} > 0$.

We will now show that this is in fact the maximum eigenvalue in the sense of the theorem. So let y be any eigenvector with eigenvalue λ, and let $|y|$ denote the vector whose components are $|y_j|$, the absolute values of the components of y. We have $|y| \in Q$ and from

$$Ty = \lambda y$$

and the triangle inequality we conclude that

$$|\lambda||y| \leq T|y|.$$

Hence $|\lambda| \leq L(|y|) \leq L_{\max}$. So we may use the notation

$$\lambda_{\max} := L_{\max}$$

since we have proved that

$$|\lambda| \leq \lambda_{\max}.$$

Suppose that $0 \leq S \leq T$. Then $sz \leq Sz$ and $Sz \leq Tz$ implies that $sz \leq Tz$ so $L_S(z) \leq L_T(z)$ for all z and hence

$$L_{\max}(S) \leq L_{\max}(T).$$

We may apply the same argument to $T^\dagger$ to conclude that it also has a positive maximum eigenvalue. Let us call it η. (We shall soon show that $\eta = \lambda_m ax$.) This means that there is a vector $w > 0$ such that

$$w^\dagger T = \eta w.$$

We have

$$w^\dagger Tx = \eta w^\dagger x = \lambda_{\max} w^\dagger x$$

implying that $\eta = \lambda_{max}$ since $w^\dagger x > 0$.

Now suppose that $y \in Q$ and $Ty \leq \mu y$. Then

$$\lambda_{\max} w^\dagger y = w^\dagger Ty \leq \mu w^\dagger y$$

implying that $\lambda_{max} \leq \mu$, again using the fact that all the components of w are positive and some component of y is positive so $w^\dagger y > 0$. In particular, if $Ty = \mu y$ then then $\mu = \lambda_{\max}$.

Furthermore, if $y \in Q$ and $Ty \leq \mu y$ then

$$0 < Py = (I+T)^{n-1}y \leq (1+\mu)^{n-1}y$$

so

$$y > 0.$$

If $\mu = \lambda_{\max}$ then $w^\dagger(Ty - \lambda_{max}y) = 0$ but $Ty - \lambda_{\max} \leq 0$ and so $w^\dagger(Ty - \lambda_{max}y) = 0$ implies that $Ty = \lambda_{\max}y$.

Suppose that $0 \leq S \leq T$ and $Sz = \sigma z,\ z \neq 0$. Then

$$T|z| \geq S|z| \geq |\sigma||z|$$

so

$$|\sigma| \leq L_{\max}(T) = \lambda_{\max},$$

as we have already seen. But if $|\sigma| = \lambda_{\max}$ then $L(|z|) = L_{\max}(T)$ so $|z| > 0$ and $|z$ is also an eigenvector of T. with the same eigenvalue. But then $(T-S)|z| = 0$ and this is impossible unless $S = T$ since $|z| > 0$. Replacing the i-th row and column of T by zeros give an $S \geq 0$ with $S < T$ since the irreducibility of T precludes all the entries in a row being zero.

Now we use a fact from linear algebra:

$$\frac{d}{d\lambda}\det(\lambda I - T) = \sum_i \det(\lambda I - T_{(i)}). \tag{1}$$

(We will remind the reader of the proof of this fact at the end of this section). Each of the matrices $\lambda_{\max}I - T_{(i)}$ has strictly positive determinant by what we have just proved. This shows that the derivative of the characteristic polynomial of T is not zero at $\lambda_{\max}$, and therefore the algebraic multiplicity and hence the geometric multiplicity of $\lambda_{\max}$ is one. QED

Let us go back to one stage in the proof, where we started with an eigenvector y, so $Ty = \lambda y$ and we applied the triangle inequality to get

$$|\lambda||y| \leq T|y|$$

to conclude that $|\lambda| \leq \lambda_{\max}$. When do we have equality? This can happen only if all the entries of $\sum_j t_{ij}y_j$ have the same argument, meaning that all the y_j with $t_{ij} > 0$ have the same argument. If T is primitive, we may apply this same argument to T^k for which all the entries are positive, to conclude that all the entries of y have the same argument. So multiplying by a complex number of arrange value one we can arrange that $y \in Q$ and from $Ty = \lambda y$ that $\lambda > 0$ and hence $\lambda = \lambda_{\max}$ and hence that y is a multiple of x. In other words, if T is primitive then we have

$$|\lambda| < \lambda_{\max}$$

for all other eigenvalues.

The matrix of a cyclic permutation has all its eigenvalues on the unit circle, and all its entries zero or one. So without the primitivity condition this result is not true. But this example suggests how to proceed.

For any matrix S let $|S|$ denote the matrix all of whose entries are the absolute values of the entries of S. Suppose that $|S| \leq T$ and let $\lambda_{\max} = \lambda_{\max}(A)$, and suppose that $Sy = \sigma y$ for some $y \neq 0$,i.e. that σ is an eigenvalue of S. Then

$$|\sigma||y| = |\sigma y| = |Sy| \leq |S||y| \leq |S||y|$$

so

$$|\sigma| \leq \lambda_{\max} = L_{\max}(A).$$

Suppose we had equality. Then we conclude from the above proof that $|y| = x$, the eigenvector of T corresponding to $\lambda_{\max}$, and then from the above string of inequalities that $|B|x = Ax$ and since all the entries of x are positive that $|B| = A$. Define the complex numbers of absolute value one

$$e^{i\theta_k} := y_k/|y_k| = y_k/x_k$$

and let D denote the diagonal matrix with these numbers as diagonal entries, so that $y = Dx$. Also write $\sigma = e^{i\phi}\lambda_{\max}$. Then

$$\sigma y = e^{i\phi}\lambda_{\max} Dx = SDx$$

so

$$\lambda_{\max} x = e^{-i\phi}D^{-1}SDx = Tx.$$

Since all the entries of $e^{i\phi}D^{-1}SD$ have absolute values $\leq$ the corresponding entries of T, and since all the entries of X are positive, we must have $|e^{i\phi}D^{-1}SD| = T$ and all the rows have a common phase and in fact

$$S = e^{i\phi}DTD^{-1}.$$

In particular, we can apply this argument to $S = T$ to conclude that if $e^{i\phi}\lambda_{\max}$ for some ϕ then

$$T = e^{i\phi}DTD^{-1}.$$

Since DTD^{-1} has the same eigenvalues as T, this shows that rotation through angle ϕ carries *all* the eigenvalues of A into eigenvalues.

The subgroup of rotations in the complex plane with this property is a finite subgroup (hence a finite cyclic group) which acts transitively on the set of eigenvalues satisfying $|\sigma| = \lambda_{\max}$. It also must act faithfully on all non-zero eigenvalues, so the order of this cyclic group must be a divisor of the number of non-zero eigenvalues. If n is a prime and T has no zero eigenvalues then either all the eigenvalues have absolute value $\lambda_{\max}$ or $\lambda_{\max}$ has multiplicity one.

We first define the **period** p of a non-zero non-negative matrix as T follows: For each i consider the set of all positive integers s such that $T_{ii}^s > 0$ and let p_i denote the greatest common denominator of this set. We show that this does not depend on i. Indeed, for some other j, there is, by irreducibility an integer M such that $T_{ij}^M > 0$ and an integer N such that $T_{ji}^N > 0$. Since $T_{ii}^{M+N} > T_{ij}^M T_{ji}^N > 0$ we conclude that $p_i|(M+N)$ and similarly that $p_j|(M+N)$. Also, if $T^s_{ii} > 0$ then $T^{s+M+N}_{jj} > T^M_{ij} T^s_{ii} T^N_{ji} T^M_{ij} > 0$ so $p_j | s$ and so $p_j | p_i$ and the reverse. Thus $p_i = p_j$, and we call this common value p.

Using the arguments above we can be more precise. We claim that $T^s_{ii} = 0$ unless s is a multiple of the order of our cyclic group of rotations, so this order is precisely the period of T. Indeed, let k be the order of this cyclic group and $\phi = 2\pi/k$. We have

$$T = e^{i\phi} D T D^{-1}$$

and hence

$$T^s = e^{is\phi} D T^s D^{-1},$$

in particular

$$T^s_{ii} = e^{is\phi} T^s_{ii}.$$

Since $e^{is\phi} \neq 1$ if s is not a multiple of k we conclude that $k = p$. So we can supplement the Perron-Frobenius theorem as

Theorem 2 Perron-Frobenius 2. *If T is primitive, all eigenvalues satisfy $|\sigma| < \lambda_{\max}$. More generally, let p denote the period of T as defined above. Then there are exactly p eigenvalues of T satisfy $|\sigma| = \lambda_{\max}$ and the entire spectrum of T is invariant under the cyclic group of rotations of order p.*

Proof of (1). The characteristic polynomial $\det(\lambda I - A)$ is invariant under conjugacy, hence so are all its coefficients. In particular the coefficient of λ which is

$$-\sum_i \det A_{(i)}$$

is invariant. Applied to the matrix $(\lambda I - T)$ we see that the right hand side of (1) is a polynomial which is invariant under conjugacy of T. So is the left. Now it is enough to prove (1) for matrices with distinct eigenvalues, since such matirces are dense in the space of all matrices. A matrix with distinct eigenvalues is conjugate to a diagonal matrix, and since both sides of (1) are invariant under conjugation, it is enough to prove (1) for diagonal matrices. But if $\lambda_1, dots, \lambda_n$ are the eigenvalues, then (1) becomes

$$\frac{d}{d\lambda} \prod_j (\lambda - \lambda_j) = \sum_i \prod_{j \neq i} (\lambda - \lambda_j)$$

which follows for the rule of differentiation of a product. QED