25 Proof of the Frobenius-Perron Theorem

We now embark on the proof of the Frobenius-Perron Theorem, starting with the following lemma.

Lemma 25.1. *Let A be a positive $n \times n$ matrix. Then A has a positive eigenvalue with a positive corresponding eigenvector.*

Proof. Let A be a positive matrix in $M_n(\mathbb{R})$.
Let S denote the set of vectors x in $\mathbb{R}^n$ that have all entries non-negative and satisfy $||x|| = 1$. This is the intersection of the sphere S^{n-1} with the set of non-negative vectors in $\mathbb{R}^n$. It is a compact subset of $\mathbb{R}^n$.

If $x \in S$, then all entries of Ax are positive. To see this, just think about any entry of Ax. The non-zero entries of x all contribute positively to this, and there are no negative contributions. We define a function $L : S \to \mathbb{R}_{>0}$ as follows. For $x \in S$,

$$L(x) = \min \left\{ \frac{(Ax)_i}{x_i} : x_i \neq 0 \right\}.$$

To understand what the function L does, start with $x \in S$. Compare the vectors x and Ax entry by entry. Look at those positions i in which x has a positive entry x_i. For each of these i, the ith entry of Ax is also positive, so it is x_i multiplied by some positive scaling factor α_i. The least of these α_i is what we are calling $L(x)$. It is a positive real number.

That is how $L(x)$ is defined for a particular $x \in S$, and L is a continuous function from S to the set of positive real numbers. Since S is compact, this means that L has a maximum value on S. Call this ρ, and let $v \in S$ be a vector for which $L(v) = \rho$.

We will show that ρ is an eigenvalue of A and that v is a corresponding eigenvector, and that $v > 0$. There are two steps.

1. First we show that $Av = \rho v$. We know that $Av \geqslant \rho v$ since $L(v) = \rho$, this means that $(Av)_i \geqslant \rho v_i$ for all i. Thus $Av - \rho v \geqslant 0$. This means that $A(Av - \rho v)$ is a positive vector and so we can choose $\epsilon > 0$ small enough that $A(Av - \rho v) > \epsilon Av$. The vector Av may not belong to S, but there is a positive real number c for which $cAv \in S$.

$$A(Av) > (\rho + \epsilon)Av \Longrightarrow A(cAv) \geqslant (\rho + \epsilon)cAv \Longrightarrow L(cAv) \geqslant (\rho + \epsilon).$$

   This contradicts the choice of ρ as the maximum value of L on S, and we conclude that $Av = \rho v$.

2. Secondly, we know that $v \geqslant 0$ since $v \in S$. It follows that $Av > 0$ - no entry of Av can be equal to zero since v is a non-negative non-zero vector and A is positive. Hence ρv is strictly positive and so v is strictly positive also.

□

Lemma 25.2. *The spectral radius of A is ρ.*

Proof. Let μ be any eigenvalue of A, and let y be a corresponding eigenvector, with $||y|| = 1$. Bear in mind that μ, and the entries of y, need not be real numbers. Now look at entry i of Ay and μy.

$$\begin{aligned}
\mu y &= Ay \\
\Longrightarrow \mu y_i &= \sum_{j=1}^{n} A_{ij} y_j \\
\Longrightarrow |\mu y_i| &\leqslant \sum_{j=1}^{n} |A_{ij} y_j| \\
\Longrightarrow |\mu|\,|y_i| &\leqslant \sum_{j=1}^{n} A_{ij} |y_j|.
\end{aligned}$$

Let $|y|$ denote the vector whose entries are the moduli of the entries of y. Then $|y| \in S$, and the last statement above says that each entry of the vector $A|y|$ is at least equal to $|\mu|$ multiplied by the corresponding element of $|y|$. This means exactly that $L(|y|) \geqslant |\mu|$. Since ρ is the maximum value of L on S, it follows that $|\mu| \leqslant \rho$. Thus ρ is the spectral radius of A. □

We have now proved parts 1. and 3. of Theorem 24.2, but we have not yet fully proved any of the other parts.

Lemma 25.3. ρ *has geometric multiplicity 1 as an eigenvalue of* A.

Proof. We know that v is a positive eigenvector of A corresponding to ρ. Suppose, anticipating contradiction, that u is an eigenvector of A corresponding to ρ, and that u is independent of v over $\mathbb{C}$.

We may assume that the entries of u are real, since ρ is real. If u has entries that are non-real complex numbers, then the real and imaginary part of u would separately be eigenvectors of A and at least one of them would be independent of v.

Now, according to this hypothesis, every element of the 2-dimensional space spanned by u and v (over $\mathbb{C}$ or $\mathbb{R}$) is an eigenvector of A corresponding to ρ. Since $v > 0$, there is a real number ϵ with the property that $u' = v + \epsilon u$ is a non-negative vector with at least one entry equal to zero. However $u' \neq 0$ since u and v are independent.

This is the required contradiction, since Au' would be positive in this case and could not be a scalar multiple of u'. □

Lemma 25.4. *The algebraic multiplicity of* ρ *as an eigenvalue of* A *is 1.*

Proof. The key to this step is to show that A is similar to a (real) matrix A' that has the entry ρ in the $(1,1)$ position and zeros throughout the rest of Row 1 and Column 1.

Since A and its transpose have the same characteristic polynomial and hence the same spectrum, the spectral radius of A^T is ρ. Our proof of Lemma 25.1 shows that there is a positive column vector w that is an eigenvector of A^T corresponding to ρ. Thus $A^T w = \rho w$ and the row vector w^T satisifies

$$w^T A = \rho w^T.$$

Now let U be the $(n-1)$-dimensional orthogonal complement of w with respect to the ordinary scalar product on $\mathbb{R}^n$:

$$U = \{u \in \mathbb{R}^n : w^T u = 0\}.$$

Let $u \in U$, and consider the vector $Au \in \mathbb{R}^n$. Note that

$$w^T A u = \rho w^T u = 0,$$

so $Au \in U$ whenever $u \in U$. This means that the subspace U of $\mathbb{R}^n$ is A-*invariant*. This is because U is the orthogonal complement in $\mathbb{R}^n$ of a left eigenvector of A, it has nothing to do with the positivity of A or the special properties of ρ and w. However these special properties give us an important extra piece of information.

Let v be the positive eigenvector of A corresponding to ρ, whose existence was shown in Lemma 25.1. Then $v \notin U$ since $w \cdot v = w^T v > 0$, because w and v are both positive. Let $\{b_1, \dots, b_{n-1}\}$ be a basis of U. Then $\mathcal{B} = \{v, b_1, \dots, b_{n-1}\}$ is a basis of $\mathbb{R}^n$.

Now the matrix A' that describes the linear transformation of $\mathbb{R}^n$ determined by left multiplication by A, with respect to the basis $\mathcal{B}$, has the following form:

$$A' = \left(\begin{array}{c|ccc} \rho & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & B_{(n-1)\times(n-1)} & \\ 0 & & & \end{array}\right)$$

Here B is $n \times n$ matrix with real entries. Since A and A' are similar, ρ occurs as an eigenvalue of both, with the same algebraic multiplicity and with geometric multiplicity 1 in each case (by

Lemmas 2.2.3 and 25.3). The characteristic polynomial of A' is $(x-\rho)p_B(x)$, where $p_B(x)$ is the characteristic polynomial of B. If the algebraic multiplicity of ρ as an eigenvalue of A' exceeds 1, then ρ is an eigenvalue of B with a corresponding eigenvector $v_B \in \mathbb{R}^{n-1}$. This means that the vector in $\mathbb{R}^n$ obtained by preceding v with a zero entry is an eigenvector of A' corresponding to ρ. Since e_1 is also an eigenvector of A' corresponding to ρ, this means that ρ has geometric multiplicity at least 2 as an eigenvector of A', and hence also as an eigenvector of A. This contradiction to Lemma 25.3 completes the proof, and we conclude that ρ occurs once as a root of the characteristic polynomial of A. ☐

Now we come to part 5., which is easy at this stage.

Lemma 25.5. *Let u be a positive eigenvector of A. Then u is a real positive scalar multiple of v.*

Proof. Let μ be the eigenvalue of A to which u corresponds. Then, μ is real and $\mu > 0$, since A and u are positive and $Au = \mu u$. Thus $0 < \mu \leqslant \rho$. Choose ϵ small enough that $u' = v - \epsilon u$ is positive. For each i we have

$$(Au')_i = \rho v_i - \mu\epsilon u_i \geqslant \rho(v_i - \epsilon u_i) = \rho u'_i.$$

Thus $Au' \geqslant \rho u'$, which means that $Au' = \rho u'$ by the maximality of ρ as a value of the function L. This means that u' is a ρ-eigenvector of A, which means that u', hence u, is a scalar multiple of v and $\mu = \rho$. ☐

The only item remaining is Item 4.

Lemma 25.6. *Suppose that μ is an eigenvalue of A, $\mu \neq \rho$. Then $|\mu| < \rho$.*

Proof. Suppose, anticipating contradiction, that $|\mu| = \rho$, and let y be an eigenvector of A corresponding to μ, with $||y|| = 1$. Let $|y|$ denote the vector in $\mathbb{C}^n$ whose entries are the moduli of the entries of y. Then $|y| \in S$ and for each i we have

$$(A|y|)_i = \sum_j A_{ij}|y_j| = \sum |A_{ij}y_j| \geqslant \left|\sum_j A_{ij}y_j\right| = |\mu y_i| = \rho|y_i|.$$

Thus $A|y| \geqslant \rho|y|$ and by Lemmas 25.1 and 25.3 this means that $|y|$ is a ρ-eigenvector of A and $|y| = v$. Then equality holds in the triangle inequality above and we have for each i that

$$\sum_j |A_{ij}y_j| = \left|\sum_j A_{ij}y_j\right|.$$

So $A_{i1}y_1,\ A_{i2}y_2, \dots, A_{in}y_n$ are complex numbers with the property that the sum of their moduli is the modulus of their sum. This means that they all lie on the same ray in the complex plane (a ray is a half-line with its endpoint at 0). Since the numbers A_{ij} are all real and positive, this means that $y_1, \dots, y_n$ all lie on the same ray. Hence there is some θ for which $e^{i\theta}y$ is a positive vector. Thus y is a (complex) scalar multiple of a positive vector, and since ρ is the only eigenvalue of A to have a positive corresponding eigenvector, it follows that $\mu = \rho$. Thus the only eigenvalue of A to have modulus ρ is ρ itself, and every other eigenvalue has modulus strictly less than the spectral radius. ☐

This completes the proof of the Frobenius-Perron theorem.

The theorem was proved by Perron for positive matrices in 1907, and extended to a slightly broader class of non-negative matrices by Frobenius in 1912. The proof in our lecture notes is mostly due to Wielandt (1950).

We conclude with some slight extensions of the theorem. A non-negative square matrix is one whose entries are all non-negative real numbers (they can be zero).

Definition 25.7. *A non-negative square matrix A is called* primitive *if A^k is positive for some positive integer k. The Frobenius-Perron theorem as we have stated it holds for primitive non-negative matrices as well as positive matrices.*

A slightly weaker version of the Perron-Frobenius Theorem holds for *irreducible non-negative matrices*. The concept of irreducibility is most easily explained by reference to a graph. Associated to a non-negative $n \times n$ matrix A is the directed graph G on n vertices in which there is an arc from vertex i to vertex j if and only if the entry A_{ij} is non-zero (i.e. positive). The graph G is *strongly connected* if every vertex can be reached from every other by a path that follows the direction of the arcs. An example of a non-negative matrix that is *not* irreducible is $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

An example of a non-negative matrix that is irreducible but not primitive is

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

The corresponding directed graph has arcs $1 \to 2$, $2 \to 3$ and $3 \to 1$, but no power of the matrix is positive.

We have the following version of the Frobenius Perron theorem for irreducible matrices.

Theorem 25.8. *Let $A \in M_n(R), A \geqslant 0$. If A is irreducible, then*

1. *The spectral radiusρ of A is an eigenvalue, with a corresponding positive eigenvector v.*
2. *ρ has algebraic (and geometric) multiplicity 1 as an eigenvalue of A.*
3. *Every positive eigenvector of A is a scalar multiple of ρ.*
4. *It is not necessarily true that ρ is the only eigenvalue whose modulus is equal to ρ. The number of such eigenvalues is the greatest common divisor of the lengths of all closed paths in the directed graph of A, and they are evenly spaced around the circle $r = \rho$.*