12.8 The Perron-Frobenius Theorem

April 13, 2020

We identify two classes of real linear operators on finite dimensional vector spaces each of which have a simple eigenvalue equal to their spectral radius.

These two classes are

- the positive matrices, and
- the irreducible nonnegative matrices.

These classes of matrices have applications in Markov chains in probability theory, compartmental models in differential equations, and Google's PageRank algorithm for information retrieval.

One should keep in mind that it is often quite difficult to say anything definitive about the spectra for real linear operators on finite dimensional vector spaces of large dimension, since it is nearly impossible to factor polynomials of large degree.

Definition 12.8.1. A matrix $A \in M_n(\mathbb{R})$ is called

(1) nonnegative, denoted $A \succeq 0$, if every entry of A is nonnegative, and

(2) positive, denoted by $A \succ 0$, if every entry of A is positive.

Remark 12.8.2. We use the notation

$$B \succeq A \text{ to mean } B - A \succeq 0,$$

and the notation

$$B \succ A \text{ to mean } B - A \succ 0.$$

Remark 12.8.3. If $A \succ 0$, then $A^k \succ 0$ for all $k \in \mathbb{N}$ because each entry of

$$A^{k-1}A$$

is the sum of products of strictly positive real numbers (by induction).

This implies that $A^k \neq 0$ for all $k \in \mathbb{N}$, so that A is not nilpotent.

Recall the spectral radius and an upper bound of it,

$$r(A) = \sup\{|\lambda| : \lambda \in \sigma(A)\} \leq \|A\|$$

for any matrix norm $\|\cdot\|$.

Note. Recall Lemma 12.6.3:

$$B \text{ is nilpotent } \Leftrightarrow r(B) = 0.$$

Since a positive matrix is not nilpotent we have

$$A \succ 0 \;\Leftrightarrow\; r(A) > 0.$$

This means that a positive matrix has a nonzero eigenvalue!

Can we reach the same conclusion for a nonnegative matrix? If $A \succeq 0$ then $r(A) > 0$?

Unfortunately, the answer is no as the zero matrix is a nonnegative matrix whose spectral radius is 0.

For any nonnegative nilpotent matrix we have the same issue.

An extra condition on a nonnegative matrix shows it is not nilpotent and hence its spectral radius is positive.

Lemma 12.8.5. If $B \succeq 0$ has a positive entry on the diagonal, then B is not nilpotent.

Proof. Let $B, C \in M_n(\mathbb{R})$ with $B, C \succeq 0$ and $b_{kk} > 0$, $c_{kk} > 0$ for some $k \in \{1, \ldots, n\}$.

The $(k,k)^{\text{th}}$ entry of BC is

$$\sum_{i=1}^{n} b_{ki} c_{ik} = b_{kk} c_{kk} + \sum_{i \neq k} b_{ki} c_{ik}.$$

Here the term $b_{kk} c_{kk} > 0$ and the rest of the sum is nonnegative, meaning the $(k,k)^{\text{th}}$ entry of BC is positive.

By induction with $C = B^{m-1}$ we obtain that B^m has a positive value in the $(k,k)^{\text{th}}$ entry for all $m \in \mathbb{N}$.

Thus B is not nilpotent. $\square$

What questions do you have?

Examples. (i) The matrix

$$A = \begin{bmatrix} 0.80 & 0.05 & 0.05 \\ 0.10 & 0.90 & 0.10 \\ 0.10 & 0.05 & 0.85 \end{bmatrix}$$

is positive, so that in the 1-norm we have

$$0 < r(A) \le \|A\|_1 = 1.$$

A right eigenvector for A is $\begin{bmatrix} 2 & 5 & 3 \end{bmatrix}^{\mathrm{T}}$, i.e.,

$$\begin{bmatrix} 0.80 & 0.05 & 0.05 \\ 0.10 & 0.90 & 0.10 \\ 0.10 & 0.05 & 0.85 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix}.$$

This implies that $r(A) \ge 1$, whence $r(A) = 1$; moreover

$$r(A) = 1 \in \sigma(A).$$

[We will see this matrix again.]

(ii) The matrix

$$A = \begin{bmatrix} 0.5 & 0 & 0.3 \\ 0 & 0.4 & 0.7 \\ 0.5 & 0.6 & 0 \end{bmatrix}$$

is nonnegative and has a positive diagonal entry, so that in the 1-norm we have

$$0 < r(A) \le \|A\|_1 = 1.$$

A right eigenvector for A is $\begin{bmatrix} 18 & 35 & 30 \end{bmatrix}^{\mathrm{T}}$, i.e.,

$$\begin{bmatrix} 0.5 & 0 & 0.3 \\ 0 & 0.4 & 0.7 \\ 0.5 & 0.6 & 0 \end{bmatrix} \begin{bmatrix} 18 \\ 35 \\ 30 \end{bmatrix} = \begin{bmatrix} 18 \\ 35 \\ 30 \end{bmatrix}.$$

This implies that $r(A) \ge 1$, whence $r(A) = 1$; moreover

$$r(A) = 1 \in \sigma(A).$$

[We will see this matrix again.]

Both of these examples are illustrative of the next result.

Theorem 12.8.4. If $A \succeq 0$, then $r(A) \in \sigma(A)$ and associated to $r(A)$ is a right eigenvector all of whose entries are nonnegative.

Definition. An eigenvalue $\lambda \in \sigma(A)$ is called simple for $A \in M_n(\mathbb{C})$ if its algebraic multiplicity is 1.

We are now in a position to state Perron's Theorem about positive matrices.

Theorem 12.8.6 (Perron). A *positive* matrix $A \in M_n(\mathbb{R})$ has a *simple* eigenvalue equal to $r(A)$, and all the other eigenvalues of A are smaller in modulus than $r(A)$. Additional, associated to the eigenvalue $r(A)$ is a *positive* right eigenvector (i.e., all of whose entries are positive).

Remark. The spectral radius is often estimated using

$$r(A) = \lim_{k\to\infty} \|A^k\|^{1/k}$$

for any matrix norm $\|\cdot\|$, and the positive right eigenvector is often estimated by the power method.

What questions do you have?

We extract a part of the proof of Perron's Theorem that is important for nonnegative matrices and positive matrices.

Proposition. Suppose $A \succeq 0$. If the diagonal entries of A are all positive, and a is the value of the smallest diagonal entry, then

$$\sigma(A) \subset \overline{B(a, r(A) - a)},$$

and consequently, $r(A)$ is the only eigenvalue on the circle $|z| = r(A)$. [Draw the picture.]

Proof. With $A \succeq 0$, by Theorem 12.8.4 the spectral radius of A is an eigenvalue, i.e., $r(A) \in \sigma(A)$.

With a being the smallest of the diagonal entries, we have $A - aI \succeq 0$.

By Theorem 12.8.4, the spectral radius of $A - aI$ is an eigenvalue, $r(A - aI) \in \sigma(A - aI)$.

For the entire function $f(z) = z - a$ we have $f(A) = A - aI$.

By the Spectral Mapping Theorem we have $\sigma(f(A)) = f(\sigma(A))$.

This implies that

$$\sigma(A - aI) = \{\lambda - a : \lambda \in \sigma(A)\}.$$

The holomorphic function f has shifted the spectrum of A to the left by a to give the spectrum of $A - aI$: we write this in unshifted form,

$$\sigma(A) = \sigma(A - aI) + a.$$

The largest positive eigenvalue of A is $r(A)$.

With every eigenvalue of $\sigma(A)$ being shifted to the left, the eigenvalue $r(A) - a$ of $A - aI$ is the largest positive eigenvalue of $A - aI$.

On the other hand, the largest positive eigenvalue of $A - aI$ is $r(A - aI)$, and so these two must agree:

$$r(A - aI) = r(A) - a.$$

With $r(A - aI)$ being the spectral radius of $A - aI$ we have

$$\sigma(A - aI) \subset \overline{B(0, r(A - aI))}.$$

Since $r(A - aI) = r(A) - a$ we obtain

$$\sigma(A - aI) \subset \overline{B(0, r(A) - a)}.$$

Since $\sigma(A) = \sigma(A - aI) + a$ we obtain

$$\sigma(A) = \sigma(A - aI) + a \subset a + \overline{B(0, r(A) - a)} = \overline{B(a, r(A) - a)}$$

where the $a+$ has shifted the center of the ball.

Note. This Proposition for nonnegative matrices with positive diagonal entries does NOT say that $r(A)$ is a simple eigenvalue.

To get $r(A)$ as a simple eigenvalue of a nonnegative matrix requires more than positive diagonal entries, as we will see in the Perron-Frobenius Theorem.

Definition. For $A \in M_n(\mathbb{R})$ satisfying $A \succeq 0$ and having positive diagonal entries, the eigenvalue $r(A) \in \sigma(A)$ is called the Perron root or the Perron-Frobenius eigenvalue of A.

What questions do you have?

Example (in lieu of 12.8.8). Find the Perron root of the positive matrix

$$A = \begin{bmatrix} 0.80 & 0.10 & 0.05 & 0.05 \\ 0.10 & 0.80 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.80 & 0.10 \\ 0.05 & 0.05 & 0.10 & 0.80 \end{bmatrix}.$$

The spectral radius $r(A)$ is bounded above by $\|A\|$ for any matrix norm $\|\cdot\|$.

In terms of the induced matrix ∞-norm we have

$$\|A\|_\infty = 1,$$

so that

$$r(A) \leq 1.$$

Can you guess an eigenvector of A?

Notice that each row sums to 1?

Thus

$$\begin{bmatrix} 0.80 & 0.10 & 0.05 & 0.05 \\ 0.10 & 0.80 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.80 & 0.10 \\ 0.05 & 0.05 & 0.10 & 0.80 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

This says that 1 is an eigenvalue of A, so that $r(A) = 1$ is the Perron root of A.

Notice also that each column of A sums to 1 as well, so that A has a left eigenvector

$$\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$$

corresponding to eigenvalue 1.

A nonnegative matrix in which all row sums and columns sums equal 1 is called a doubly stochastic matrix.

[You are NOT responsible for the notion of doubly stochastic matrix.]

What questions do you have?

Example. Find the Perron root and a corresponding positive right eigenvector for the positive matrix

$$A = \begin{bmatrix} 0.80 & 0.05 & 0.05 \\ 0.10 & 0.90 & 0.10 \\ 0.10 & 0.05 & 0.85 \end{bmatrix}.$$

Notice here that each row does not sum to 1, but that each column does sum to 1.

A nonnegative matrix in which each column sums to 1 is called a stochastic matrix. [You are NOT responsible for the notion of stochastic matrix.]

Thus A has a left eigenvector $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ with eigenvalue 1 because

$$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0.80 & 0.05 & 0.05 \\ 0.10 & 0.90 & 0.10 \\ 0.10 & 0.05 & 0.85 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

The spectral radius $r(A)$ is bounded above by $\|A\|_1 = 1$.

Since $1 \in \sigma(A)$, the Perron root is $r(A) = 1$, but the corresponding positive right eigenvector is not so obvious.

We use the Power Method to find a positive right eigenvector.

For the initial guess we use

$$\mathrm{v} = \begin{bmatrix} 0.6 \\ 0.25 \\ 0.15 \end{bmatrix},$$

a probability vector.

We then have

$$A\mathrm{v} = \begin{bmatrix} 0.5 \\ 0.3 \\ 0.2 \end{bmatrix}, \ A^2\mathrm{v} = \begin{bmatrix} 0.425 \\ 0.34 \\ 0.235 \end{bmatrix}, \dots, A^{20}\mathrm{v} = \begin{bmatrix} 0.2012 \\ 0.4971 \\ 0.3016 \end{bmatrix}, \dots,$$

$$A^{40}\mathrm{v} = \begin{bmatrix} 0.2000 \\ 0.4999 \\ 0.3000 \end{bmatrix}.$$

It appears that the sequence is converging to

$$\begin{bmatrix} 0.2 \\ 0.5 \\ 0.3 \end{bmatrix},$$

so that our initial guess was "good."

We verify the limiting vector as a positive right eigenvector for the Perron root of A:

$$\begin{bmatrix} 0.80 & 0.05 & 0.05 \\ 0.10 & 0.90 & 0.10 \\ 0.10 & 0.05 & 0.85 \end{bmatrix} \begin{bmatrix} 0.2 \\ 0.5 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.5 \\ 0.3 \end{bmatrix}.$$

Notice that the sum of the entries of the positive right eigenvector is one which is to say that this right eigenvector is a probability vector for the stochastic matrix A.

What questions do you have?

What conditions on a nonnegative matrix imply its spectral radius is a simple eigenvalue? Positive diagonal entries is not enough.

Definition. We say $A \succeq 0$ is irreducible if for each $i, j \in \{1, \ldots, n\}$ there exists $k \in \mathbb{N}$ such that the $(i,j)^{\text{th}}$ entry of A^k is positive.

Definition. We say $A \succeq 0$ is primitive if there exists $k \in \mathbb{N}$ such that $A^k \succ 0$.

Second Reading Quiz Question. A nonnegative matrix A being primitive implies that A is irreducible.

True

Proposition 12.8.10. If $A \succeq 0$ is irreducible, then $I + A$ is primitive.

The proof of this is HW (Exercise 12.38. Hint: use the Binomial Theorem on $(I + A)^K$ for an appropriately chosen $K \in \mathbb{N}$).

Theorem 12.8.11 (Perron-Frobenius). A *nonnegative irreducible* matrix $A \in M_n(\mathbb{R})$ has a *simple* eigenvalue equal to $r(A)$ and associated to $r(A)$ is a *positive* right eigenvector.

First Reading Quiz Question. What are the similarities and differences between the Perron and the Perron-Frobenius Theorems?

For Perron, the hypothesis is

- $A \succ 0$

and the conclusions are

- $r(A) \in \sigma(A)$ is simple and dominant, and
- there exists a positive right eigenvector.

For Perron-Frobenius, the hypotheses are

- $A \succeq 0$, and
- A is irreducible,

and the conclusion are

- $r(A) \in \sigma(A)$ is simple, and
- there exists a positive right eigenvector.

In general, the power method cannot be used to find the positive right eigenvector for the Perron-Frobenius Theorem.

Examples. (i) The nonnegative matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

has positive diagonal entries, but it is not irreducible and is not primitive because

$$A^2 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & 2 & 2 & 0 \\ 0 & 2 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix},$$

$$A^3 = \begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & 2 & 2 & 0 \\ 0 & 2 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 & 4 \\ 0 & 4 & 4 & 0 \\ 0 & 4 & 4 & 0 \\ 4 & 0 & 0 & 4 \end{bmatrix},$$

etc., so that by induction we obtain

$$A^k = \begin{bmatrix} 2^{k-1} & 0 & 0 & 2^{k-1} \\ 0 & 2^{k-1} & 2^{k-1} & 0 \\ 0 & 2^{k-1} & 2^{k-1} & 0 \\ 2^{k-1} & 0 & 0 & 2^{k-1} \end{bmatrix},$$

which says that certain entries of A^k are always zero no matter the value of k.

But because $A \succeq 0$ we can apply Theorem 12.8.4 to get $r(A) \in \sigma(A)$.

Here $\sigma(A) = \{0, 2\}$ with each having algebraic multiplicity 2, so $r(A) = 2 \in \sigma(A)$ (the Perron root) but it is not simple.

(ii) The stochastic matrix

$$A = \begin{bmatrix} 0.5 & 0 & 0.3 \\ 0 & 0.4 & 0.7 \\ 0.5 & 0.6 & 0 \end{bmatrix}$$

is primitive because

$$A^2 = \begin{bmatrix} 0.4 & 0.18 & 0.15 \\ 0.35 & 0.58 & 0.28 \\ 0.25 & 0.24 & 0.57 \end{bmatrix}.$$

This also implies that A is irreducible.

The spectral radius $r(A)$ is 1 because $r(A)$ is bounded above by $\|A\|_1 = 1$ and $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ is a left eigenvector of A for eigenvalue 1.

By the Perron-Frobenius Theorem, the eigenvalue $r(A) = 1$ is simple, and it has a corresponding positive eigenvector.

Can we find this positive eigenvector by the power method? Because A is primitive, the answer is yes (see Exercise 12.39 part (i) which gives dominance of the simple eigenvalue $r(A) = 1$).

For the initial guess $\mathrm{v} = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}^{\mathrm{T}}$, the sequence of iterates $A^k\mathrm{v}$ converges to

$$\frac{1}{83}\begin{bmatrix} 18 \\ 35 \\ 30 \end{bmatrix}.$$

The initial guess and the limit vector are probability vectors: they are nonnegative vectors whose entries sum to 1, i.e., their 1-norms are 1.

What questions do you have?