Math 346 Lecture #38
12.8 The Perron-Frobenius Theorem

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.

12.8.1 Perron's Theorem

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

Remark 12.8.2. We use the notation $B \succeq A$ to mean $B - A \succeq 0$, and the notation $B \succ A$ 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. By Lemma 12.6.3, this implies that the spectral radius of a positive matrix is positive, i.e., if $A \succ 0$ then $r(A) > 0$. Recall that

$$r(A) = \sup\{|\lambda| : \lambda \in \sigma(A)\} \geq 0.$$

Theorem 12.8.4. If $A \succeq 0$, then there exists $\lambda \in \sigma(A)$ such that $\lambda = r(A) \geq 0$, and associated to λ is an eigenvector all of whose entries are nonnegative.

Proof. If $r(A) = 0$, there is nothing to show. So suppose that $r(A) > 0$.

By Theorems 12.3.8 and 12.3.14 we have the Laurent series for the resolvent

$$R_A(z) = \sum_{k=0}^{\infty} \frac{A^k}{z^{k+1}}$$

converges on the annulus $|z| > r(A)$.

By way of contradiction, suppose that $\lambda = r(A) > 0$ is not an eigenvalue of A.

Then $R_A(z)$ is continuous at λ so that

$$R_A(\lambda) = \lim_{z\to\lambda} R_A(z) = \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}}.$$

Since $A \succeq 0$ then $A^k \succeq 0$ for all $k \in \mathbb{N}$, implying that each entry in the power series is a series of nonnegative terms, and hence is absolutely convergent, i.e., for A^k_{ij} the $(i,j)^{\text{th}}$ entry of A^k, there holds for all $i, j \in \{1, \ldots, n\}$,

$$\sum_{k=1}^{\infty} \frac{A^k_{ij}}{\lambda^{k+1}} < \infty.$$

Under the change of variables $z = 1/w$, the power series

$$\sum_{k=0}^{\infty} A^k w^{k+1}$$

converges uniformly on the compact disk $\overline{B(0,1/\lambda)}$ because for any $w \in \mathbb{C}$ with $|w| = 1/\lambda$ there holds

$$\sum_{k=0}^{\infty} A_{ij}^k |w|^{k+1} \leq \sum_{k=0}^{\infty} \frac{A_{ij}^k}{\lambda^{k+1}} < \infty.$$

This implies that the Laurent series for the resolvent $R_A(z)$ converges for all $z \in \mathbb{C}$ on $|z| = \lambda$, whence that $r(A) < \lambda$.

This contradiction shows that $\lambda = r(A)$ is an eigenvalue of A.

To show that λ has an eigenvector with nonnegative entries, we multiply the form of the resolvent, given in the Spectral Decomposition Theorem, by the power of the eigennilpotent $D_\lambda^{m_\lambda - 1}$.

When $m_\lambda > 1$ this gives, and using the properties of the eigenprojections and eigennilpotents, that

$$\begin{aligned} R_A(z) D_\lambda^{m_\lambda - 1} &= \sum_{\mu \in \sigma(A)} \left[\frac{P_\mu}{z - \mu} + \sum_{k=1}^{m_\mu - 1} \frac{D_\mu^k}{(z-\mu)^{k+1}} \right] D_\lambda^{m_\lambda - 1} \\ &= \frac{D_\lambda^{m_\lambda - 1}}{z - \lambda}. \end{aligned}$$

When $m_\lambda = 1$ we set $D_\lambda^{m_\lambda - 1} = P_\lambda$.

In either case we obtain

$$(z - \lambda) D_\lambda^{m_\lambda - 1} = R_A(z)^{-1} D_\lambda^{m_\lambda - 1} = (zI - A) D_\lambda^{m_\lambda - 1}.$$

Cancelling the common z from both sides results in

$$A D_\lambda^{m_\lambda - 1} = \lambda D_\lambda^{m_\lambda - 1}.$$

This says that any nonzero column of $D_\lambda^{m_\lambda - 1}$ is an eigenvector of A corresponding to λ.

It remains to show that $D_\lambda^{m_\lambda - 1}$ has a nonzero column, and a nonzero column with all nonnegative entries.

From the Laurent series (12.31) of the resolvent $R_A(z)$ about λ,

$$R_A(z) = \frac{P_\lambda}{z - \lambda} + \sum_{k=1}^{m_\lambda - 1} \frac{D_\lambda^k}{(z-\lambda)^{k+1}} + \sum_{k=0}^{\infty} (-1)^k (z - \lambda)^k S_\lambda^{k+1},$$

we obtain

$$D_\lambda^{m_\lambda - 1} = \lim_{z \to \lambda} (z - \lambda)^{m_\lambda} R_A(z).$$

On the other hand, we have another Laurent series for $R_A(z)$ about λ,

$$R_A(z) = \sum_{k=0}^{\infty} \frac{A^k}{z^{k+1}}, \ |z| > r(A).$$

With $A \succeq 0$ and taking $z \in \mathbb{R}$ approaching λ from the right we obtain

$$D_\lambda^{m_\lambda - 1} = \lim_{z \to \lambda^+} (z-\lambda)^{m_\lambda} \sum_{k=0}^{\infty} \frac{A^k}{z^{k+1}} \succeq 0.$$

This shows that every entry of $D_\lambda^{m_\lambda-1}$ is nonnegative.

Since $D_\lambda^{m_\lambda-1} \neq 0$, it has at least one nonzero column, and this is the desired eigenvector with nonnegative entries. □

Lemma 12.8.5. If $B \in M_n(\mathbb{R})$ is nonnegative and has a positive entry on 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. □

Remark. We are now in a position to state Perron's Theorem about positive matrices. It gives the existence of a simple dominant eigenvalue equal to the spectral radius which can be estimated using any matrix norm $\|\cdot\|$ in the limit $r(A) = \lim_{k\to\infty} \|A^k\|^{1/k}$. Perron's Theorem also guarantees the existence of an eigenvector associated to the dominant simple eigenvalue all of whose entries are positive. We can apply the power method to find this eigenvector if this eigenvector is not obvious to guess.

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 right eigenvector all of whose entries are positive.

Proof. Suppose $A \succ 0$. Then $r(A) > 0$

Then $A \succeq 0$ so by Theorem 12.8.4 the spectral radius $r(A)$ is an eigenvalue λ of A and associated to this eigenvalue is an eigenvector v_1 all of whose entries are nonnegative.

We show that any eigenvector for λ with nonnegative entries must have positive entries.

Since $A \succ 0$ and $\mathrm{v}_1 \succeq 0$ it follows that $A\mathrm{v}_1 \succ 0$ (it is impossible for any entry of $A\mathrm{v}_1$ to be 0 as any positive entry of v_1 gets multiplied by a positive entry of A and to this is added no worse than a nonnegative quantity).

Since $A\mathrm{v}_1 = \lambda \mathrm{v}_1$, $\lambda = r(A) > 0$, and $A\mathrm{v}_1 \succ 0$, it follows that $\mathrm{v}_1 \succ 0$, i.e.,

$$\mathrm{v}_1 = (1/\lambda)A\mathrm{v}_1 \succ 0.$$

This shows that any eigenvector for λ with nonnegative entries must have positive entries.

Next we show that the eigenvalue $\lambda = r(A)$ is semisimple.

In the proof of Theorem 12.8.4 we showed that the columns of $D_\lambda^{m_\lambda - 1}$ are eigenvectors if $m_\lambda > 1$.

The eigenvector v_1 is one of these columns and has every entry positive, implying by Lemma 12.8.5 that D_λ is not nilpotent.

Thus $D_\lambda = 0$, i.e., $m_\lambda = 1$, so that λ is a semisimple eigenvalue of A.

Next we show that $\lambda = r(A)$ is a simple eigenvalue.

Since A and λ are both real, the eigenspace $\mathscr{E}_\lambda$ has a basis of real vectors.

Suppose that $\{\mathrm{v}_1, \mathrm{v}_2\}$ is a linearly independent subset in $\mathscr{E}_\lambda$.

We can assume WLOG that at least one entry of v_2 is negative (just multiply v_2 by -1 if every entry of v_2 is positive).

For every $t \in \mathbb{R}$, the vector $t\mathrm{v}_1 + \mathrm{v}_2$ is an eigenvector of A corresponding to λ, i.e.,

$$A(t\mathrm{v}_1 + \mathrm{v}_2) = \lambda(t\mathrm{v}_1 + \mathrm{v}_2) = t\lambda\mathrm{v}_1 + \lambda\mathrm{v}_2.$$

Since $\mathrm{v}_1 \succ 0$ there exists $t \in \mathbb{R}$ such that $t\mathrm{v}_1 + \mathrm{v}_2 \succeq 0$ with at least one entry equal to 0.

But as we showed above, a nonnegative eigenvector must be a positive eigenvector.

This contradiction shows that $\dim(\mathscr{N}(A - \lambda I)) = 1$, so that $\lambda = r(A)$ a simple eigenvalue.

Finally we show that $\lambda = r(A)$ is the dominant eigenvalue of A.

Since $A \succ 0$ we can choose $\epsilon > 0$ so that $A - \epsilon I \succeq 0$.

The function $f(z) = z - \epsilon$ is entire, so by the Spectral Mapping Theorem we have $\sigma(A - \epsilon I) = \sigma(f(A)) = f(\sigma(A))$.

Since $r(A) \in \sigma(A)$, we have $r(A) - \epsilon = f(r(A)) \in \sigma(A - \epsilon I)$.

Since $r(A)$ is the largest positive eigenvalue of A, the quantity $r(A) - \epsilon$ is the largest positive eigenvalue of $A - \epsilon I$ (since the entire function $f(z) = z - \epsilon$ shifts all of the eigenvalues of A to the left by ϵ).

By Theorem 12.8.4, the matrix $A - \epsilon I \succeq 0$ has an eigenvalue equal to its spectral radius which must be the largest positive eigenvalue of $A - \epsilon I$, namely

$$r(A - \epsilon I) = r(A) - \epsilon = \lambda - \epsilon.$$

Because $f(z) = z - \epsilon$ shifts the $\sigma(A)$ by ϵ we have $\sigma(A) = \sigma(A - \epsilon I) + \epsilon$.

Since the $r(A-\epsilon) = \lambda - \epsilon$ we have that $\sigma(A - \epsilon I) \subset \overline{B(0, \lambda - \epsilon)}$, so that

$$\sigma(A) \subset (A - \epsilon I) + \epsilon \subset \overline{B(\epsilon, \lambda - \epsilon)}.$$

This implies that λ is the only eigenvalue of A on the circle $|z| = \lambda$, hence that λ is dominant. $\square$

Remark 12.8.7. For any $A \succeq 0$ whose smallest diagonal entry a is positive, we can apply the argument at the proof of Perron's Theorem to $A - \epsilon I$ for $\epsilon = a$ to show that

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

Thus, in the case of a nonnegative matrix with positive diagonal we still have the conclusion that the eigenvalue $\lambda = r(A)$ from Theorem 12.8.4 (often called the Perron root or Perron-Frobenius eigenvalue of A) is the only eigenvalue of A on the circle $|z| = r(A)$.

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.]

Example. Find the Perron root and a corresponding 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.]

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 right eigenvector is not so obvious.

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

For the initial guess we use $\mathrm{v} = \begin{bmatrix} 0.6 & 0.25 & 0.15 \end{bmatrix}^{\mathrm{T}}$.

We then have

$$A\mathrm{v} = \begin{bmatrix} 0.5 \\ 0.3 \\ 0.2 \end{bmatrix}, \ A^2\mathrm{v} = \begin{bmatrix} 0.425 \\ 0.34 \\ 0.235 \end{bmatrix}, \ldots, A^{20}\mathrm{v} = \begin{bmatrix} 0.2012 \\ 0.4971 \\ 0.3016 \end{bmatrix}, \ldots, A^{40}\mathrm{v} = \begin{bmatrix} 0.2000 \\ 0.4999 \\ 0.3000 \end{bmatrix}.$$

It appears that the sequence is converging to $\begin{bmatrix} 0.2 & 0.5 & 0.3 \end{bmatrix}^{\mathrm{T}}$, 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.

12.8.2 Perron-Frobenius Theorem

Frobenius extended the existence of a simple eigenvalue equal to the spectral radius (which eigenvalue may not be dominant) and an associated positive eigenvector of Perron's Theorem to the larger class of nonnegative matrices that are irreducible, a concept we now define.

Definition 12.8.9. Let $A \in M_n(\mathbb{R})$.

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 $(i,j)^{\mathrm{th}}$ entry of A^k is positive.

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

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 matrix $A \in M_n(\mathbb{R})$ has a simple eigenvalue λ equal to $r(A)$ and associated to λ is a positive eigenvector.

See the book for a proof.

Example. (i) The nonnegative matrix

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

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.

(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 (yes being primitive implies irreducibility!).

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

By the Perron-Frobenius Theorem, the eigenvalue $\lambda = r(A) = 1$ 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)).

For the initial guess $\mathrm{v} = \begin{bmatrix} 0.2 & 0.3 & 0.5 \end{bmatrix}$, 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.