A short proof of Perron's theorem.

Hannah Cairns, April 25, 2014.

A matrix A or a vector ψ is said to be *positive* if every component is a positive real number. $A \geq B$ means that every component of A is greater than or equal to the corresponding component of B. In the same way, $A > B$ means that every component of A is greater than the corresponding component of B.

DEFINITION. The *spectral radius* $\rho(A)$ of a square matrix A is the maximum of the absolute values of the eigenvalues of A.

THEOREM. (*Perron's Theorem.*) Let A be a positive square matrix. Then:

   a) $\rho(A)$ is an eigenvalue, and it has a positive eigenvector.
   b) $\rho(A)$ is the only eigenvalue on the disc $|\lambda| = \rho(A)$.
   c) $\rho(A)$ has geometric multiplicity 1.
   d) $\rho(A)$ has algebraic multiplicity 1.

PRELIMINARIES. The proof of the Perron-Frobenius theorem will rely on the following simple positivity trick.

TRICK. Let A be a positive square matrix. If θ and ψ are two nonequal column vectors with $\theta \geq \psi$, then $A\theta > A\psi$. There is some positive $\varepsilon > 0$ with $A\theta > (1 + \varepsilon)A\psi$.

PROOF. The vector $A(\theta - \psi) = \sum A_{ij}(\theta_j - \psi_j)$, so

$$(A[\theta - \psi])_i \geq \sum_j \min_{ij}(A_{ij})(\theta_j - \psi_j) = \min_{ij}(A_{ij}) \sum(\theta_j - \psi_j) > 0.$$

So $(A\theta)_i - (A\psi)_i > 0$, and by definition $A\theta > A\psi$. This proves the first statement. If we change the vector $A(\theta - \psi)$ by a small amount then it will still be positive, so there is some $\varepsilon > 0$ with $A(\theta - \psi) - \varepsilon A\psi > 0$, or $A\theta > (1 + \varepsilon)A\psi$. That proves the second statement. $\square$

We also need the idea of the spectral radius of a matrix, together with a theorem from linear algebra.

THEOREM. (Gelfand's formula) The spectral radius of a matrix A can be written in terms of the norms of its powers:

$$\rho(A) = \lim(||A^n||)^{1/n}.$$

We can now begin the proof of the Perron-Frobenius theorem.

PROOF.

(a) *A has λ as an eigenvalue, and λ has a positive eigenvector.*

There is some eigenvalue λ with $|\lambda| = \rho(A)$. Let ψ be an eigenvector. Let Ψ be the vector with $\Psi_j = |\psi_j|$. Then

$$(A\Psi)_i = \sum A_{ij}|\psi_j| \geq \left|\sum A_{ij}\psi_j\right| = |\lambda\psi_i| = \rho(A)\Psi_i,$$

so $A\Psi \geq \rho(A)\Psi$.

If they are not equal, then by the positivity trick, $A^2\Psi > \rho(A)A\Psi$, and there is some positive $\varepsilon > 0$ with $A^2\Psi \geq (1+\varepsilon)\rho(A)A\Psi$.

The matrix A is nonnegative, and so are all its powers A^n, so multiplying both sides of the equation by A^n preserves the inequality:

$$A^{n+2}\Psi \geq (1+\varepsilon)\rho(A)A^{n+1}\Psi.$$

This is true for every n, so

$$A^{n+1}\Psi \geq (1+\varepsilon)\rho(A)A^n\Psi$$
$$\geq ((1+\varepsilon)\rho(A))^2 A^{n-1}\Psi$$
$$\vdots$$
$$\geq ((1+\varepsilon)\rho(A))^n A\Psi.$$

Gelfand's formula tells us that

$$\rho(A) = \lim_n ||A^n||^{1/n} \geq (1+\varepsilon)\rho(A)$$

which is a contradiction, so our assumption must have been mistaken. This means that the two vectors were equal: $A\Psi = \rho(A)\Psi$.

So $|\psi|$ is also an eigenvector with eigenvalue $\rho(A)$. It is not only nonnegative but positive, because $A|\psi| = \rho(A)|\psi|$ is positive.

(b) *The only eigenvalue on the circle $|\lambda| = \rho(A)$ is $\rho(A)$.*

Suppose $\lambda \neq \rho(A)$ is some other eigenvalue on the circle $|\lambda| = \rho(A)$. We will reach a contradiction. Repeat the reasoning from part (a) again: if $\Psi = |\psi|$, then we showed that $A\Psi = \rho(A)\Psi$, or

$$\sum A_{ij}|\psi_j| = \left|\sum A_{ij}\psi_j\right|$$

for every $1 \leq j \leq n$.

LEMMA. Let A be a positive $n \times n$ matrix. Let ψ be a vector in $\mathbb{C}^n$. Choose an index j. If

$$\sum A_{ij}|\psi_j| = \left|\sum A_{ij}\psi_j\right|,$$

then there is some $c \in \mathbb{C}$ with $c \neq 0$ so that the product $c\psi$ is a nonnegative vector.

PROOF. If $\psi = 0$, then $c = 1$ works. Suppose ψ is not zero. Square both sides and write $|\sum A_{ij}\psi_j|^2 = (\sum A_{ij}\psi_j)\overline{(\sum A_{k\ell}\psi_\ell)}$.

We get the equality

$$\sum A_{ij}A_{k\ell}|\psi_j||\psi_\ell| = \sum A_{ij}A_{k\ell}\psi_j\overline{\psi_\ell}.$$

Subtract the right-hand side from the left:

$$\sum A_{ij}A_{k\ell}\big(|\psi_j||\psi_\ell| - \psi_j\overline{\psi_\ell}\big) = 0$$

All of the coefficients $A_{ij}A_{k\ell}$ are positive real numbers, and the term in parentheses has nonnegative real part, and positive real part unless $\psi_j\overline{\psi_\ell} \geq 0$ for every pair j, ℓ.

Then we can take $c = \overline{\psi_\ell}$ for some index ℓ where $\psi_\ell \neq 0$. $\qquad\square$

The conditions of the lemma hold, so $c\psi > 0$ for some complex number $c \neq 0$. But this means that

$$\lambda(c\psi) = c(\lambda\psi) = c(A\psi) = A(c\psi) \geq 0$$

is positive, and $c\psi > 0$, so λ must be nonnegative.

But $\rho(A)$ is the only nonnegative number on the circle $|\lambda| = \rho(A)$.

(c) $\rho(A)$ *has geometric multiplicity 1.*

Suppose ψ is the positive eigenvector above and ψ' is a linearly independent eigenvector of the eigenvalue $\rho(A)$.

We can assume that ψ' is real; otherwise we take real and imaginary parts, and the parts are still eigenvectors, because A and $\rho(A)$ are real. One of them must be linearly independent of ψ.

Let $c > 0$ be chosen so that $\psi - c\psi'$ is nonnegative and at least one entry is zero. It is not the zero vector, because ψ, ψ' are linearly independent. But

$$\psi - c\psi' = \frac{A[\psi - c\psi']}{\rho(A)} > 0.$$

We chose c so that at least one entry was zero, so this is a contradiction. Therefore, there cannot be two linearly independent eigenvectors, so $\rho(A)$ has geometric multiplicity 1.

(d) $\rho(A)$ *has algebraic multiplicity 1.*

As in previous parts, let ψ be a right eigenvector of $\rho(A)$. We now know that there is only one eigenvector and that it is positive.

Let π be a positive left eigenvector of $\rho(A)$. We can get such an eigenvector by applying (a) to A^T.

This pair of eigenvectors allows us to decompose $\mathbb{R}^n$ as a direct sum.

LEMMA. *The space $\pi^\perp = \{x \in \mathbb{R}^n : \pi x = 0\}$ is invariant under A.*

PROOF. If $\pi x = 0$, then $\pi A x = (\pi A)x = \pi x = 0$. □

This space has dimension $n - 1$, and ψ is a vector that is not in $\pi^\perp$, because

$$\pi\psi = \sum \pi_j \psi_j > 0.$$

Therefore, $\mathbb{R}^n$ is the direct sum of span$\{\psi\}$ and $\pi^\perp$.

Let $\psi_2, \ldots, \psi_n$ be a basis for the space $\pi^\perp = \{x \in \mathbb{R}^n : \pi x = 0\}$. Let X be the matrix

$$X = \begin{bmatrix} \psi \ \psi_2 \ \cdots \ \psi_n \end{bmatrix}.$$

Then XAX^{-1} leaves invariant the spaces $X^{-1}\text{span}\{\pi\} = \text{span}\{e_1\}$ and $X^{-1}\pi^\perp = \text{span}\{e_2, \ldots, e_n\}$, so this change of basis turns A into a block diagonal matrix:

$$X^{-1}AX = \left[\begin{array}{c|c} \rho(A) & 0 \\ \hline 0 & Y \end{array}\right]$$

for some matrix Y.

Suppose λ has algebraic multiplicity more than one. Then it must also be an eigenvalue of Y. But then Y must have an eigenvector for $\rho(A)$ too, which makes two eigenvectors, which contradicts (b).

Therefore, $\rho(A)$ has algebraic multiplicity one. This proves part (d).

Finally, we can easily extend the theorem to the case where A is nonnegative and has a positive power A^m.

THEOREM. *(Perron-Frobenius theorem.)* The statements (a), (b), (c), (d) are also true for nonnegative matrices A so that some power A^m is positive.

PROOF. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A, counted with algebraic multiplicity. Suppose λ_1 has the largest absolute value.

Then the eigenvalues of A^m are $\lambda_1^m, \ldots, \lambda_n^m$, where λ_1^m is the largest in absolute value. Perron's Theorem tells us that λ_1^m is positive, and it has

a positive eigenvector ψ. It also tells us that all the other eigenvalues of A^m are smaller.

But if $|\lambda_1|^m > |\lambda_j|^m$ for $j \geq 2$, then $|\lambda_1| > |\lambda_j|$ for $j \geq 2$. Therefore, λ_1 is the largest eigenvalue in absolute value. It has algebraic multiplicity one. Its eigenvector must be ψ, because no other eigenvector of A has the right eigenvalue.

So we know (b), (c), (d), and (a) except for the positivity of λ_1. This follows from the fact that ψ is positive, and $\lambda_1 \psi = A\psi \geq 0$.

It turns out that this is equivalent to a condition on the directed graph $G(A)$ on $\{1, \ldots, n\}$ with an edge from i to j if $A_{ij} > 0$.

There is some m with $A^m > 0$ if and only if

    a) every vertex in $G(A)$ can be reached from every other vertex

    b) the gcd of the lengths of the loops in $G(A)$ is one.

This can be proved by the methods we saw in class. We sketch a proof.

SKETCH. We observe that there is a path of length m from i to j in $G(A)$ if and only if $(A^m)_{ij} > 0$. The first condition guarantees that for every i, j there are infinitely many m with $(A^m)_{ij} > 0$, and the second condition tells us that there exist loops at any vertex of any sufficiently long length (which we can obtain by chaining the shorter loops together).

Suppose that we can get from any vertex i to any other vertex j in time at most s, and that there are loops at every vertex of every length at least r. Let $m = s + r$. Then for every pair i, j, we can go from i to j in time $t \leq s$ and then pick a loop at j which has length $(m - t) \geq r$. This is a path of length m from i to j, so $(A^m)_{ij} > 0$.

References:

Roger A. Horn and Charles R. Johnson. *Matrix analysis.* Cambridge, 1985. (See chapter 8.)

Laurent Saloff-Coste. *Lectures on finite Markov chains.* In *Lectures on Probability Theory and Statistics,* volume 1665 of *Lecture Notes in Mathematics*, pages 301–413. http://www.math.cornell.edu/~lsc/math778-only/stf.pdf