The Perron-Frobenius theorem

The theorem we discuss here about matrices with non-negative entries was proved first for matrices with positive entries by Perron in 1907 and extended by Frobenius to irreducible matrices with non-negative entries in 1912. This theorem has miriads of applications; in particular for ranking the "importance" of URL's on the web (the Google ranking).

1 Non-negative and positive matrices

We say that a real vector $x = (x_i)_{i\in[n]} \in \mathbb{R}^n$ is *non-negative*, (resp. *positive*) if all the entries x_i are non-negative (resp. positive) ; we write $x \geqslant 0$ (resp. $x > 0$). We also use these definitions for real matrices.

A non-negative matrix square A is *irréductible* if

$$\forall i, j \in [n]^2, \exists t \in \mathbb{N}, \ : \ A^t_{i\,j} > 0.$$

It is called *primitive* if

$$\exists t \in \mathbb{N}, \forall i, j \in [n]^2, \ : \ A^t_{i\,j} > 0.$$

The *graph associated to the non-negative square matrix* A of size $n \times n$ is the directed graph $G(A)$ with the set of vertices equal to $[n]$ and a set of directed edges defined by

$$(i, j) \in E \Leftrightarrow A_{i\,j} > 0.$$

We easily check that $A^t_{i\,j} > 0$ if and only if there exists a path from i to j of length t in $G(A)$. Hence A is irréductible if and only if $G(A)$ is strongly connected.

Exercise 1. *Let A be a square positive matrix. Show that if A is nilpotent, then A is not irreducible.*

Proposition 2. *Let A be a square positive matrix. If A is irreducible, then $I + A$ is primitive.*

Proof. The binomial expansion

$$(I + A)^n = \sum_{k=0}^{n} \binom{n}{k} A^k$$

has positive entries since A is irreducible. □

2 Perron-Frobenius theorem

Theorem 3. *Let A be an irreducible matrix.*

1. *The spectral radius ρ_A of A is a positive eigenvalue of A.*
2. *Furthermore ρ_A has algebraic and geometric multiplicity one with a positive eigenvector x.*

3. If A is primitive, then each other eigenvalue λ of A satisfies

$$|\lambda| < \rho_A.$$

We now embark on the proof of this important theorem. Let

$$P = (I + A)^n.$$

Since P is positive, then for any non-negative and non-null vector v we have $Pv > 0$.

We let

$$Q = \{x \in \mathbb{R}^n \ : \ x \geqslant 0, \ x \neq 0\}$$

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

$$C = \{x \in \mathbb{R}^n \ : \ x \geqslant 0, \ \|x\| = 1\}$$

where $\|.\|$ is any norm on $\mathbb{R}^n$. Clearly C is a compact set.

For any $z \in Q$, let us introduce

$$L(z) = \max\{s \in \mathbb{R} \ : \ sz \leqslant Pz\} = \min_{1 \leqslant i \leqslant n, z_i \neq 0} \frac{(Az)_i}{z_i}.$$

We now give some basic properties of L.

1. By definition $L(rz) = L(z)$ for any $r > 0$.
2. If z is an eigenvector of A for the eigenvalue λ, then $L(z) = \lambda$.
3. If $sz \leqslant Az$, then

   $$sPz \leqslant PAz = APz,$$

   and so

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

   Furthermore, if z is not an eigenvector of A, then $sz \neq Az$ for any s and $sPz < APz$. From the second expression of $L(z)$, it follows that $L(z) < L(Pz)$.

That suggests a plan for the proof of the Perron-Frobenius theorem: we look for a positive vector which maximizes L, show that it is the eigenvector we want in the theorem, and establishes the properties stated in the theorem.

Proof of the Perron-Frobenius theorem

1. Finding a positive eigenvector.

Consider the image of C under P: it is a compact set and all the vectors in $P(C)$ are positive. Hence by the second expression of $L(z)$, we obtain that L is continuous on $P(C)$. Thus L achieves its maximum value on $P(C)$, i.e., there exists $x \in P(C)$ such that

$$L(x) = \sup_{z \in C} L(Pz).$$

Since $L(z) \leqslant L(Pz)$, in fact x realizes the maximum value $L_{\max}$ of L on Q. Hence

$$L_{\max} = L(x) \leqslant L(Px) \leqslant L_{\max}.$$

From the third property of L, it follows that x is an eigenvector of A with the eigenvalue $L_{\max}$. Since $x \in P(C)$, x is a positive vector.

2. Showing that $L_{\max}$ is the spectral radius.

Let $z \in \mathbb{C}^n$ be an eigenvector of A with the eigenvalue $\lambda \in \mathbb{C}$, and let $|z|$ the vector in $\mathbb{R}^n$ whose entries are $|z_i|$. We have $|z| \in Q$, and from $Az = \lambda z$ which says that

$$\lambda z_i = \sum_{k=1}^{n} A_{i\,k} z_k$$

and the fact that $A_{i\,k} \geqslant 0$ we conclude that

$$|\lambda|\,|z_i| \leqslant \sum_{k=1}^{n} A_{i\,k} |z_k|$$

which we write for short as

$$|\lambda|\,|z| \leqslant A|z|.$$

By definition of L, it follows that

$$|\lambda| \leqslant L(|z|).$$

Hence $|\lambda| \leqslant L_{\max}$ which proves that

$$\rho \leqslant L_{\max}$$

where ρ is the spectral radius of A. Conversely from what we have just proved, we deduce that

$$L_{\max} \leqslant \rho.$$

That achieves the proof of item 1 in the theorem.

3. Showing that $L(z) = L_{\max} \;\Rightarrow Az = L_{\max}\, z \;\wedge\; z > 0$

Observe that the above proof shows that if $L(z) = L_{\max}$, then

$$L(z) = L(Pz).$$

Thus z is an eigenvector of A for the eigenvalue $L_{\max}$. It follows that z is also an eigenvector of P, i.e., $Pz = \lambda z$. Since P is positive, we have $Pz > 0$. So z is positive.

4. Showing that $0 \leqslant B \leqslant A$, $B \neq A \;\Rightarrow\; \rho_B < \rho_A$.

First let us stress on the fact that, contrary to A, the matrix B is not supposed to be irreducible.

Suppose that $Bz = \lambda z$ with $z \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$. Then

$$|\lambda|\,|z| \leqslant B|z| \leqslant A|z|.$$

It follows that

$$|\lambda| \leqslant L_A(|z|) \leqslant \rho_A.$$

Therefore

$$\rho_B \leqslant \rho_A.$$

Suppose that $|\lambda|= \rho_A$. Then from the above inequalities, we derive that $L_A(z) = \rho_A$. Using the above remark, we obtain that $|z|$ is an eigenvector of A for the eigenvalue ρ_A and z positive. Hence $B|z|= A|z|$ with $z > 0$ which is impossible unless $A = B$.

Replacing the i-th row and column of A by zeros gives a non-negative matrix A_i such that $0 \leqslant A_i \leqslant A$. Moreover $A_i \neq A$ since the irreducibility of A precludes all the entries in a row being zeros. This proves that for each matrix $A_{(i)}$ obtained by eliminating the i-th row and the i-th column of A, the eigenvalues of $A_{(i)}$ are all less than ρ_A.

5. A basic lemma in linear algebra

Let A be a square marix of size n and Δ the diagonal matrix with entries $X^T = (x_1, \cdots, x_n) \in \mathbb{R}^n$ along the diagonal. Expanding $\det(\Delta - A)$ along the i-th row shows that

$$\frac{\partial}{\partial x_i} \det(\Delta - A) = \det(\Delta_{(i)} - A_{(i)}).$$

So

$$\frac{\mathrm{d}}{\mathrm{d}\,x} \det(xI - A) = \sum_{i=1}^{n} \det(xI - A_{(i)}).$$

6. Showing that ρ_A has algebraic multiplicity one

First observe that

$$\det(xI - A_i) = x \det(xI - A_{(i)}).$$

By what we have just proved

$$\det(\rho_A I - A_{(i)}) > 0.$$

This shows that the derivative of the characteristic polynomial of A is not zero at ρ_A, and therefore the algebraic (and so geometric) multiplicity of ρ_A is one. That completes the proof of item 2.

7. Proof of the last assertion of the Perron-Frobenius theorem

The t-th powers of the eigenvalues of A are the eigenvalues of A^t. So if we want to show that there are no eigenvalues of a primitive matrix with absolute values[1] equal to ρ_A other than ρ_A, it is enough to prove this for a positive matrix.

Suppose that $Az = \lambda z$ with $z \in \mathbb{C}^n$, $\lambda \in \mathbb{C}$ and $|\lambda|= \rho_A$. Then

$$\rho_A\,|z|= |Az| \leqslant A|z|.$$

It follows that

$$\rho_A \leqslant L(|z|) \leqslant \rho_A,$$

which implies that $L(|z|) = \rho_A$. From the assertion (4) above, we deduce that $|z|$ is an eigenvector of A with the eigenvalue ρ_A. Moreover, we have

$$|Az|= A|z|.$$

[1] Attention : la traduction en francais, de "absolute value" est "module" (et non "valeur absolue").

In particular

$$|\sum_{i=1}^{n} A_{1\,i} z_i| = \sum_{i=1}^{n} A_{1\,i} |z_i|.$$

Since all the entries of A are positive, this implies that there exists $u \in \mathbb{C}$ (with an absolute value equal to 1) such that

$$\forall i \in [n],\ z_i = u|z_i|.$$

Hence z and $|z|$ are collinear eigenvectors of A. So the corresponding eigenvalues λ and ρ_A are equal, as required.

3 Asymptotic behavior of powers of a primitive matrix

Let A be a primitive matrix and ρ_A its spectral radius. Let x be a positive eigenvector of A with the eigenvalue ρ_A. The transpose of A has the same spectrum (with same algebraic multiplicities) as A. In particular, the spectral radius of A^T is also ρ_A, and ρ_A is an eigenvalue with algebraic multiplicity 1. Since A^T is also an irreducible positive matrix, we can apply the Perron-Frobenius theorem to A^T and derive that there exists a unique positive vector y up to scalar by a positive number such that

$$A^T(y) = y.$$

Let us choose y so that $x^T y = \sum_{i=1}^{n} x_i y_i = 1$.

We easily check that

$$\mathbb{R}^n = \mathbb{R}x \oplus \ker y^T$$

and $\mathbb{R}x$ as well as $\ker y^T$ are invariant under A. Moreover the matrix $H = xy^T$ is the projection whose image is $\mathbb{R}x$ and whose kernel is $\ker y^T$.

We set

$$P = \frac{1}{\rho_A} A$$

and we consider the restriction of P to $\ker y^T$ that we denote Q. Clearly 1 is an eigenvalue of P with algebraic multiplicity one, and the spectrum of Q is equal to

$$Sp(Q) = Sp(P) \setminus \{1\}.$$

Furthermore by item 3 in the Perron-Frobenius theorem, every eigenvalue λ of P different from 1 satisfies $|\lambda| < 1$. Hence

$$\rho_Q < 1.$$

Let

$$v = \mu x + z$$

with $z \in \ker y^T$; so $Hv = \mu x$. For any $t \in \mathbb{N}$,

$$P^t v = \mu x + Q^t z.$$

By Gelfand's theorem, we have

$$\|Q^t\| \sim (\rho_Q)^t$$

where $\|.\|$ is any matrix norm. Therefore

$$\lim_{t \to +\infty} P^t v = Hv.$$

So we have proved that

Theorem 4. *Let A be a non-negative matrix. If A is primitive, then*

$$\lim_{t \to +\infty} \left(\frac{1}{\rho_A} A\right)^t = xy^T$$

where x and y are positive eigenvectors of A and A^T for the eigenvalue ρ_A, and $x^T y = 1$.

4 Criteria for a matrix to be primitive

The *cyclicity* of an irreducible non-negative matrix A is the greatest common divisor of the lengths of the cycles in the associated graph. Let $\mathbf{N}_{i,j}$ be the subset of integers defined by:

$$\mathbf{N}_{i,j} = \{t \in \mathbb{N} \mid \left(A^t\right) > 0\}.$$

Let us denote the cyclicity of A by γ; let $\gamma_i = \gcd(\mathbf{N}_{i,i})$. Obviously,

$$\gamma = \gcd(\{\gamma_i \mid i \in V\}). \tag{1}$$

Observe that each $\mathbf{N}_{i,i}$ is closed under addition (semi-group); let $\gamma_i = \gcd(\mathbf{N}_{i,i})$. We will use the following elementary lemma from number theory whose proof is based on Bézout theorem and left as exercise.

Lemma 5. *A set $\mathbf{N}$ of positive integers that is closed under addition contains all but a finite number of multiples of its greatest common divisor.*

Exercise 6. *Prove Lemma 5.*

Lemma 7. *For any $i \in [n]$, $\gamma_i = \gamma$.*

Proof. Let i, j be any pair of nodes in the associated graph, and let $a \in \mathbf{N}_{i,j}$ and $b \in \mathbf{N}_{j,i}$. The concatenation of a path from i to j with a path from j to i is a closed path starting at i. Hence $a + b \in \mathbf{N}_{i,i}$. From Lemma 5, we know that $\mathbf{N}_{j,j}$ contains all the multiples of γ_j greater than some integer. Consider any such multiple $k\gamma_j$ with k and γ_i relatively prime integers. By inserting one corresponding cycle at node j into the cycle at i with length $a + b$, we obtain a new cycle starting at i, i.e., $a + k\gamma_j + b \in \mathbf{N}_{i,i}$. It follows that γ_i divides both $a+b$ and $a+k\gamma_j+b$, and so γ_i divides γ_j. Similarly, we prove that γ_j divides γ_i, and so $\gamma_i = \gamma_j$. By (1), the common value of the γ_i's is actually equal to γ. □

We are now in position to give several criterions for an irreducible matrix to be primitive.

Theorem 8. *Let A be an irreducible matrix. The following assertions are equivalent:*

1. *The matrix A is primitive.*
2. *All the eigenvalues of A different from its spectral radius ρ_A satisfy $|\lambda| < \rho_A$.*

3. *The sequence of matrices* $\left(\frac{1}{\rho_A}A\right)^t$ *converges to a positive matrix.*

4. *There exists some* $i \in [n]$ *such that* $\gamma_i = 1$*.*

5. *The cyclicity of* A *is equal to 1.*

Proof. (1) $\Rightarrow$ (2) coincides with the last item of the Perron-Frobenius theorem.
To show that (2) $\Rightarrow$ (3), it suffices to observe that actually, the proof of Theorem 4 uses the assertion (2) only.
Suppose (3) and let $i \in [n]$. The sequence $\left(\frac{1}{(\rho_A)^t}A^t_{i\,i}\right)$ converges to a positive limit. Hence for t enough large, we have $A^t_{i\,i} > 0$. It follows that $\gamma_i = 1$, i.e., assertion (4) holds.
Lemma 7 shows that (4) and (5) are equivalent.
Now assume (4), i.e., $\gamma_i = 1$ for some $i \in [n]$. By Lemma 5, there exists some integer t_i such that $[t_i, +\infty[\subseteq \mathbf{N}_{i,i}$. Let $j, k \in [n]$; since A is irreducible there exist two positive integers u and v at most equal to n such that

$$A^u_{j\,i} > 0 \text{ and } A^v_{i\,k} > 0.$$

Hence for each $t \geqslant 2n + t_i$, we have

$$A^t_{j\,k} \geqslant A^u_{j\,i}A^{t-u-v}_{i\,i}A^v_{i\,k} > 0$$

which proves that A is primitive. □

5 The Leslie model of population growth

In 1945, Leslie introduced a model for the growth of a population and its projected age distribution that is closed to migration and where only one sex, usually the female, is considered. Thus the population is described by a vector whose size is the number of age groups and whose i-th component is the number of females in the i-th age group. For a thorough discussion of the Leslie model, see the book [1].

Let f_i be the expected number of daughters produced by a female in the i-th age group, and s_i the proportion of females in the i-th age group who survive to the next age group in one time unit.

Exercise 9. *The point of the exercise is to study the growth of a population in the Leslie model.*

1. *Show that the transition after one time unit is given by an irreducible matrix* L *(called the Leslie matrix).*

2. *Show that* L *has a unique positive eigenvector up to some positive scalar.*

3. *Under what condition* L *is primitive? If so, show that asymptotically the total population grows (or declines) at some rate* r *and that the relative size of each age group to the total population converges to some limit that is independent of the initial population.*

4. *Consider the population of Atlantic salmon who die immediately after spawning, and assume that there are three age groups. The corresponding Leslie matrix is equal to*

$$\begin{pmatrix} 0 & 0 & f \\ s_1 & 0 & 0 \\ 0 & s_2 & 0 \end{pmatrix}.$$

   What happens asymptotically?

6 Stochastic matrices and ergodic matrices

A non-negative matrix A is *stochastic* if the entries in each row sum to 1:

$$\forall i \in [n], \ \sum_{k=1}^{n} A_{i\,k} = 1.$$

In other words, each row of A is a probability vector.

We easily check that the set of stochastic matrices is compact, and that the (finite or infinite) product of stochastic matrices is a stochastic matrix. By definition the column vector $\mathbb{1}$ all of whose entries equal 1 is an eigenvector with eigenvalue 1. Moreover, 1 is its spectral radius: let v an eigenvector with the eigenvalue λ, and let $i \in [n]$ such that $|v_i|= \max_{k=1}^{n}|v_k|$. Thus

$$|\lambda v_i|= |\sum_{k=1}^{n} A_{i\,k} v_k|\leqslant \sum_{k=1}^{n} A_{i\,k}|v_k|\leqslant \sum_{k=1}^{n} A_{i\,k} = |v_i|.$$

So $|\lambda|\leqslant 1$.

A stochastic matrix that is primitive is said to be *ergodic*. From the Perron-Frobenius theorem, we know that if A is an irreducible stochastic matrix, then 1 is an eigenvalue with algebraic (and geometric) multiplicity 1 with a positive eigenvector. Moreover if A is ergodic, then all the other eigenvalues of A satisfy $|\lambda|< 1$.

Exercise 10.
1. *Find an example of a stochastic matrix such that 1 is not a simple eigenvalue (algebraic multiplicity greater than one).*
2. *Find an example of a stochastic matrix with an eigenvalue different from 1 whose absolute value is equal to 1.*

The positive vector y defined in Section 3 for any irreducible matrix A and any positive eigenvector x corresponding to the eigenvalue ρ_A is called the *Perron vector* of A when x is the $\mathbb{1}$ vector; it is denoted by π_A or simply π when no confusion can arise. Recall that it is defined by

$$\pi > 0, \quad \sum_{i=1}^{n} \pi_i = 1, \quad A^T \pi = \pi.$$

As an immediate corollary of Theorem 4, we derive the following convergence result for powers of an ergodic matrix.

Corollary 11. *Let A be a stochastic matrix. If A is ergodic, then the sequence of matrices $\left(A^t\right)_{t\in\mathbb{N}}$ converges to a (stochastic) matrix with range 1. More precisely*

$$\lim_{t\to+\infty} A^t = \mathbb{1}\pi^T$$

where π is the Perron vector of A.

Exercise 12. *Let A be the stochastic matrix. Find a necessary and sufficient condition on A for the Perron vector of A to be collinear with the vector $\mathbb{1}$.*

In the case of Exercise 12, A is said to be a *doubly stochastic* matrix.

Exercise 13. *Let G be a bidirectional graph with n vertices, and let A be the stochastic matrix whose associated graph is G and whose non-zero entries in each row are all equal. Compute the Perron vector of A.*

Exercise 14. *Let G be the directed graph whose set of nodes is $[n]$ and formed by the union of a directed cycle C_n consisting of the edges $(i, i+1)$ (where i is taken modulo n) and $n-1$ edges $(i, 1)$ for $i \in [n-1]$. Let A be the stochastic matrix whose associated graph is G and whose non-zero entries in each row are all equal. Show that the Perron vector π of A is given by*

$$\pi_i = \begin{cases} 1/\left(2 - 2^{-n+1}\right) & \text{if } i = 1 \\ \\ 2^{-i+1}\pi_1 & \text{if } i \in [n] \end{cases}$$

Compare with the case of a bidirectional graph (previous exercise).

References

[1] Hal Caswell. *Matrix Population Models, Construction, Analysis and Interpretation.* Sinauer Associates, 2nd Edition, Sunderland, MA, USA, 2000.