ESAIM: PS 23 (2019) 797–802
https://doi.org/10.1051/ps/2019007

ESAIM: Probability and Statistics
www.esaim-ps.org

GALTON–WATSON AND BRANCHING PROCESS REPRESENTATIONS OF THE NORMALIZED PERRON–FROBENIUS EIGENVECTOR

Raphaël Cerf[1,2] and Joseba Dalmau[3,*]

Abstract. Let A be a primitive matrix and let λ be its Perron–Frobenius eigenvalue. We give formulas expressing the associated normalized Perron–Frobenius eigenvector as a simple functional of a multitype Galton–Watson process whose mean matrix is A, as well as of a multitype branching process with mean matrix $e^{(A-I)t}$. These formulas are generalizations of the classical formula for the invariant probability measure of a Markov chain.

Mathematics Subject Classification. 60J80.

Received September 11, 2018. Accepted April 1, 2019.

1. Introduction

Let A be a primitive matrix of size N, *i.e.*, a non-negative matrix whose mth power is positive for some natural number m. The Perron–Frobenius theorem ([3], Thm. 1.1) states that there exist a positive real number λ and a vector u with positive coordinates such that $u^T A = \lambda u^T$. Moreover, the eigenvalue λ is simple, is larger in absolute value than any other eigenvalue of A, and any non-negative eigenvector of A is a multiple of u. The eigenvalue λ is the Perron–Frobenius eigenvalue of A and u is a Perron–Frobenius eigenvector of A. In the particular case of a stochastic matrix A, there is an intimate link between the Perron–Frobenius theorem and the theory of Markov chains. Indeed, if A is stochastic, then $\lambda = 1$ and the vector u represents the invariant probability measure of the Markov chain having A for transition matrix. For a general primitive matrix A, the Perron–Frobenius theorem carries along, but the probabilistic link to the theory of Markov chains falls apart (compare for instance Chaps. 5 and 6 in [3]). The purpose of this note is to provide a natural probabilistic interpretation of the Perron–Frobenius eigenvector, but this time in terms of branching processes. More precisely, we give probabilistic representations of the normalized Perron–Frobenius eigenvector $u/|u|_1$, as a functional of a multitype Galton–Watson process, as well as of a multitype branching process. Some of the formulas that we employ are mere reformulations of well-known formulas appearing for example in Chapter 6 of [3]. Our main contribution is to give them a probabilistic interpretation in terms of branching processes, which, by the way, renders the formulas more concise and elegant.

Keywords and phrases: Galton–Watson, branching process, Perron–Frobenius.

[1] Département de Mathématiques et Applications, Ecole Normale Supérieure, CNRS, PSL Research University, 75005 Paris, France.
[2] Laboratoire de Mathématiques d'Orsay, Université Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France.
[3] Centre de Mathématiques Appliquées, Ecole Polytechnique, CNRS, Université Paris-Saclay, 91405 Orsay, France.

[*] Corresponding author: joseba.dalmau@polytechnique.edu.

2. The Galton–Watson case

A multitype Galton–Watson process is a Markov chain

$$Z_n \,=\, \big(Z_n(1),\dots,Z_n(N)\big)\,,\qquad n\geq 0\,,$$

with state space $\mathbb{N}^N$. The number $Z_n(i)$ represents the number of individuals having type i in generation n. In order to build generation $n+1$ from generation n, each individual of type i present in generation n produces a random number of offspring, distributed according to a prescribed reproduction law, independently of the other individuals and the past of the process. The ensemble of all the offspring forms the generation $n+1$. The null vector is an absorbing state. For each $i\in\{\,1,\dots,m\,\}$, we denote by P_i and E_i the probabilities and expectations for the process started from a population consisting of a single individual of type i. From now onwards, we consider a multitype Galton–Watson model $(Z_n)_{n\geq 0}$ whose mean matrix is equal to A, *i.e.*, we suppose that

$$\forall\, i,j\in\{\,1,\dots,N\,\}\qquad E_i\big(Z_1(j)\big)\,=\,A(i,j)\,.$$

There exist intimate links between the asymptotic behavior of the Galton–Watson process $(Z_n)_{n\geq 0}$, the Perron–Frobenius eigenvalue λ, and the associated normalized eigenvector u of A. For instance, the following classical result can be found in Chapter 2 of [2]. If the Perron–Frobenius eigenvalue λ of A is strictly larger than one, then the multitype Galton–Watson process has a positive probability of survival. Conditionally on the survival event, the vector of proportions of the different types converges almost surely to u when time goes to ∞, *i.e.*, conditionally on the survival event, with probability one,

$$\forall\, i\in\{\,1,\dots,N\,\}\qquad \lim_{n\to\infty}\frac{Z_n(i)}{Z_n(1)+\cdots+Z_n(N)}\,=\,u(i)\,.$$

In this note, we shall present a simpler formula linking the Galton–Watson process $(Z_n)_{n\geq 0}$ with λ and u, which works not only in the case $\lambda>1$, but also in the critical and subcritical cases. Let us fix $i\in\{\,1,\dots,N\,\}$. We shall stop the process $(Z_n)_{n\geq 0}$ on the type i by killing the descendants of individuals of type i in any generation $n\geq 2$. The resulting process is denoted by $(Z^i_n)_{n\geq 0}$. Thus, in the stopped process $(Z^i_n)_{n\geq 0}$, the individuals reproduce as in the Galton–Watson process $(Z_n)_{n\geq 0}$, however from generation 1 onwards, the individuals of type i do not have any descendants. Notice that the initial individual of type i produces offspring as in the Galton–Watson process $(Z_n)_{n\geq 0}$, only individuals of type i belonging to the subsequent generations are prevented from having offspring. Finally, for $u=(u(1),\dots,u(N))$ a vector in $\mathbb{R}^N$, we define

$$|u|_1\,=\,|u(1)|+\cdots+|u(N)|\,.$$

Theorem 2.1. *The normalized Perron–Frobenius eigenvector u of A is given by the formula*

$$\forall i\in\{\,1,\dots,N\,\}\qquad u(i)\,=\,\frac{1}{\displaystyle\sum_{n\geq 1}\lambda^{-n}E_i\big(\big|Z^i_n\big|_1\big)}\,.$$

Notice that $\big|Z^i_n\big|_1$ is simply the size of nth generation of the process $(Z^i_n)_{n\geq 0}$. In the case where $\lambda\geq 1$, the factor λ^{-n} is naturally interpreted as a killing probability. We introduce a random clock τ_λ, independent of the branching process $(Z_n)_{n\geq 0}$, and distributed according to the geometric law of parameter $1-1/\lambda$:

$$\forall n\geq 1\qquad P(\tau_\lambda\geq n)\,=\,\Big(\frac{1}{\lambda}\Big)^{n-1}\,.$$

The formula presented in the theorem can then be rewritten as

$$\forall i \in E \qquad u(i) \,=\, \frac{1}{E_i\left(\displaystyle\sum_{n=1}^{\tau_\lambda - 1} \big|Z_n^i\big|_1\right)}\,.$$

The nicest situation is when the Perron–Frobenius eigenvalue is equal to one. In this case, the formula becomes

$$\forall i \in \{\,1,\dots,N\,\} \qquad u(i) \,=\, \frac{1}{E_i\Big(\displaystyle\sum_{n\geq 1} \big|Z_n^i\big|_1\Big)}\,.$$

The denominator is naturally interpreted as the expected number of descendants from an individual of type i, if the descendants of type i are forbidden to reproduce. Let us remark also that, by multiplying the matrix A by a constant factor, we can adjust the value of the Perron–Frobenius eigenvalue without altering the Perron–Frobenius eigenvector. More precisely, suppose that the mean matrix of the Galton–Watson process is given by

$$\forall\, i,j \in \{\,1,\dots,N\,\} \qquad E_i\big(Z_1(j)\big) \,=\, c\,A(i,j)\,,$$

where c is a positive constant. If we take $c = 1/\lambda$, then we obtain indeed a critical branching process and the Perron–Frobenius eigenvalue is 1. In practice, the exact value of the Perron–Frobenius eigenvalue might be unknown, so we can simply choose a value c large enough so that the Perron–Frobenius eigenvalue becomes larger than one, and we can introduce the random killing clock as above.

In the particular case where the matrix A is a stochastic matrix, and each individual produces exactly one child, the Perron–Frobenius eigenvalue λ is equal to 1 and the process $(Z_n)_{n\geq 0}$ is simply a Markov chain with transition matrix A. The stopped process $(Z_n^i)_{n\geq 0}$ is the Markov chain stopped at the time τ_i of the first return to i. So, in this situation, the population Z_n^i has size 1 until time τ_i and 0 afterwards, therefore

$$\sum_{n\geq 1} \lambda^{-n} E_i\big(\big|Z_n^i\big|_1\big) \,=\, E_i\left(\sum_{n\geq 1}^{\tau_i} 1\right) \,=\, E_i(\tau_i)$$

and we recover the classical formula for the invariant probability measure of a Markov chain.

Let us come to the proof of the theorem. The theorem is in fact a consequence of the following proposition.

Proposition 2.2. *Let $i \in \{\,1,\dots,N\,\}$. The vector v defined by*

$$\forall j \in \{\,1,\dots,N\,\} \qquad v(j) \,=\, \sum_{n\geq 1} \lambda^{-n} E_i\big(Z_n^i(j)\big)$$

is the Perron–Frobenius eigenvector of A satisfying $v(i) = 1$.

Indeed, the formula appearing in the theorem is obtained by normalizing the above vector. We now proceed to the proof of the proposition.

Proof. We fix $i \in \{\,1,\dots,N\,\}$ and we define a vector v *via* the formula stated in the proposition. Let us examine first $v(i)$. By definition of the stopped process $(Z_n^i)_{n\geq 0}$, we have

$$v(i) \,=\, \sum_{n\geq 1} \sum_{i_1,\dots,i_{n-1}\neq i} \lambda^{-n} A(i,i_1)\cdots A(i_{n-1},i)\,.$$

The lemma of [1] yields that the above sum is equal to 1, thus $v(i) = 1$. Let next j belong to $\{ 1, \dots, N \}$. We have

$$(vA)(j) \,=\, \sum_{1 \le k \le N} v(k)A(k,j) \,=\, A(i,j) + \sum_{\substack{1 \le k \le N \\ k \ne i}} v(k)A(k,j).$$

We compute next

$$\begin{aligned}
\sum_{\substack{1 \le k \le N \\ k \ne i}} v(k)A(k,j) \,&=\, \sum_{\substack{1 \le k \le N \\ k \ne i}} \sum_{n \ge 1} \lambda^{-n} E_i\big(Z_n^i(k)\big)A(k,j) \\
&= \sum_{n \ge 1} \lambda^{-n} E_i\Big(\sum_{\substack{1 \le k \le N \\ k \ne i}} Z_n^i(k)A(k,j)\Big) \\
&= \sum_{n \ge 1} \lambda^{-n} E_i\Big(E\Big(Z_{n+1}^i(j) \,\Big|\, Z_n^i \Big)\Big) \\
&= \sum_{n \ge 1} \lambda^{-n} E_i\big(Z_{n+1}^i(j)\big) \\
&= \lambda v(j) - E_i\big(Z_1^i(j)\big).
\end{aligned}$$

Remember that the initial individual of type i reproduces as in the Galton–Watson process $(Z_n)_{n \ge 0}$, therefore $E_i\big(Z_1^i(j)\big) = A(i,j)$ and putting together the previous computations, we obtain

$$\forall j \in \{ 1, \dots, N \} \qquad (vA)(j) \,=\, \lambda v(j)\,.$$

The matrix A is primitive, there exists some m such that A^m is strictly positive. From the previous equation taken for $j = i$, we obtain

$$(vA^m)(i) \,=\, \lambda^m v(i)\,.$$

Since $v(i) = 1$ and all the elements of the matrix A^m are strictly positive, all of the $v(j)$ must be finite. Moreover,

$$\lambda^m v(j) \,=\, (vA^m)(j) \,\ge\, v(i)A^m(i,j) \,>\, 0\,,$$

hence the coefficient $v(j)$ is positive. The vector v is strictly positive, therefore v is a left Perron–Frobenius eigenvector of A, as wanted. $\square$

3. The branching process case

A multitype branching process is a continuous-time Markov process

$$Z_t \,=\, \big(Z_t(1), \dots, Z_t(N)\big), \quad t \ge 0,$$

with state space $\mathbb{N}^N$. The number $Z_t(i)$ represents the number of individuals carrying the type i at time t. Individuals reproduce independently of each other, at a rate dependent on their type. When an individual reproduces, it gives birth to a random number of offspring, distributed according to a prescribed reproduction law, independently of the other individuals and the past of the process. The null vector is an absorbing state. For each $i \in \{ 1, \dots, m \}$, we denote by P_i and E_i the probabilities and expectations for the process started from a population consisting of a single individual of type i. We consider a multitype branching process whose mean matrix has generator $A - I$, in other words, we suppose that

$$\forall\, i,j \in \{\,1,\dots,N\,\} \quad \forall t \geq 0 \qquad E_i\big(Z_t(j)\big) \;=\; \big(e^{(A-I)t}\big)(i,j),$$

where the exponential appearing in the formula is the matrix exponential. This mean matrix corresponds to the process where individuals reproduce at rate 1, and the reproduction laws are the same as in the discrete Galton–Watson process. There exist well-known links between the asymptotic behavior of the branching process $(Z_t)_{t\geq 0}$, the Perron–Frobenius eigenvalue λ, and the associated eigenvector u of A, analogous to those of the Galton–Watson case. We shall next present a simple formula in the spirit of Theorem 2.1. Let us fix $i \in \{\,1,\dots,N\,\}$. We will stop the process $(Z_t)_{t\geq 0}$ on the type i by killing the descendants of type i at any time $t \geq 0$. The resulting process is denoted by $(Z^i_t)_{t\geq 0}$. We denote by $E_{(i)}$ the expectation for the process $(Z^i_t)_{t\geq 0}$ starting from a random population, drawn according to the original reproduction law of an individual of type i, so that $E_{(i)}(Z^i_0(j)) = A(i,j)$. All individuals of type i die at rate 1 without producing offspring.

Theorem 3.1. *The normalized Perron–Frobenius eigenvector u of A is given by the formula*

$$\forall\, i \in \{\,1,\dots,N\,\} \qquad u(i) \;=\; \frac{1}{\displaystyle\int_0^\infty e^{-(\lambda-1)t} E_{(i)}\big(\big|Z^i_t\big|_1\big)\,\mathrm{d}t}\,.$$

As for the Galton–Watson case, this result is a direct consequence of the following proposition.

Proposition 3.2. *Let $i \in \{\,1,\dots,N\,\}$. The vector v defined by*

$$\forall j \in \{\,1,\dots,N\,\} \qquad v(j) \;=\; \int_0^\infty e^{-(\lambda-1)t} E_{(i)}\big(Z^i_t(j)\big)\,\mathrm{d}t$$

is the Perron–Frobenius eigenvector of A satisfying $v(i) = 1$.

Proof. As in the discrete case, we do the proof by verifying that the vector v is indeed an eigenvector of A. Let $k \in \{\,1,\dots,N\,\}$, and let us start by computing the integral involved in the definition of $v(k)$. Differentiating the expectation with respect to t yields

$$\frac{\mathrm{d}}{\mathrm{d}t} E_{(i)}\big(Z^i_t(k)\big) \;=\; \sum_{\substack{1\leq j\leq N\\ j\neq i}} E_{(i)}\big(Z^i_t(j)\big)A(j,k) - E_{(i)}\big(Z^i_t(k)\big).$$

Thus, integrating by parts, for any $T > 0$,

$$\begin{aligned}\int_0^T e^{-(\lambda-1)t} E_{(i)}\big(Z^i_t(k)\big)\,\mathrm{d}t \;=& -\frac{1}{\lambda-1} e^{-(\lambda-1)T} E_{(i)}\big(Z^i_T(k)\big) + \frac{1}{\lambda-1} E_{(i)}\big(Z^i_0(k)\big)\\ &+ \frac{1}{\lambda-1}\int_0^T e^{-(\lambda-1)t}\Bigg(\sum_{\substack{1\leq j\leq N\\ j\neq i}} E_{(i)}\big(Z^i_t(j)\big)A(j,k) - E_{(i)}\big(Z^i_t(k)\big)\Bigg)\,\mathrm{d}t.\end{aligned}$$

Let B be the matrix obtained from A by filling with zeros its ith row. The first expectation on the right-hand side can be rewritten as

$$E_{(i)}\big(Z^i_t(k)\big) \;=\; \sum_{1\leq j\leq N} A(i,j)\big(e^{(B-I)t}\big)(j,k).$$

Yet it follows from part (e) of Theorem 1.1 of [3] that the spectral radius of B is strictly less than λ. Therefore, when t goes to infinity, the matrix exponential $e^{(B-I)t}$ behaves as $e^{\mu' t}$, for some μ' strictly smaller than $\lambda - 1$. Sending T to infinity in the above integrals we obtain the following identity:

$$\lambda v(k) \;=\; A(i,k) + \sum_{\substack{1\le j\le N\\ j\neq i}} v(j)A(j,k).$$

Thus, the proof will be achieved if we manage to show that $v(i) = 1$. Yet, the previous formula holds for $k = i$ too, and we may use it iteratively over $v(j)$ in order to get, for any $n \geq 1$,

$$v(i) \;=\; \sum_{h=1}^{n} \frac{1}{\lambda^h} \sum_{i_1,\dots,i_h\neq i} A(i,i_1)\cdots A(i_h,i) + \frac{1}{\lambda^n} \sum_{i_1,\dots,i_{n+1}\neq i} v(i_1)A(i_1,i_2)\cdots A(i_{n+1},i).$$

Again, calling B the matrix obtained from A by filling with zeros its ith row, the last term can be written as

$$\frac{1}{\lambda^n} \sum_{\substack{1\le i_1\le N\\ i_1\neq i}} v(i_1)B^n(i_1,i),$$

which converges to 0 when n goes to ∞. Thus,

$$v(i) \;=\; \sum_{n\ge 1} \sum_{i_1,\dots,i_{n-1}\neq i} \lambda^{-n} A(i,i_1)\cdots A(i_{n-1},i).$$

This last quantity is equal to 1, as shown in the lemma of [1]. □

References

[1] R. Cerf and J. Dalmau, A Markov chain representation of the normalized Perron-Frobenius eigenvector. *Electron. Commun. Probab.* **22** (2017) 52.

[2] T.E. Harris, The Theory of Branching Processes. Springer-Verlag, Berlin, Prentice-Hall, Inc., Englewood Cliffs, NJ (1963).

[3] E. Seneta, Nonnegative Matrices and Markov Chains, 2nd edn. *Springer Series in Statistics*. Springer-Verlag, New York (1981).