A Probabilistic Proof of the Perron-Frobenius Theorem

Peter W. Glynn[1] and Paritosh Y. Desai[2]

[1]Department of Management Science and Engineering, Stanford University, Stanford CA
[2]Target Corporation, Sunnyvale, CA

Abstract

The Perron-Frobenius theorem plays an important role in many areas of management science and operations research. This paper provides a probabilistic perspective on the theorem, by discussing a proof that exploits a probabilistic representation of the Perron-Frobenius eigenvalue and eigenvectors in terms of the dynamics of a Markov chain. The proof provides conditions in both the finite-dimensional and infinite-dimensional settings under which the Perron-Frobenius eigenvalue and eigenvectors exist. Furthermore, the probabilistic representations that arise can be used to produce a Monte Carlo algorithm for computing the Perron-Frobenius eigenvalue and eigenvectors that will be explored elsewhere.

1 Introduction

The Perron-Frobenius theorem for non-negative matrices plays an important role in many areas of operations research, the management sciences, and applied mathematics in general. An important such setting is that of Markov population decision chains, in which Perron-Frobenius theory can be used to describe the rate at which the total reward aggregated across the entire population grows as a function of time; see, for example, [Rothblum and Veinott 1975], [Pliska 1976], [Rothblum 1984], [Rothblum and Whittle 1982], and [Rothblum and Veinott 1992].

Given the importance of the Perron-Frobenius theorem, it is convenient to have multiple perspectives on the result. In this paper, we explore the Perron-Frobenius theorem from a probabilistic viewpoint. In particular, we give a proof of the Perron-Frobenius theorem that relies on a representation of the Perron-Frobenius eigenvalue and its associated row and column eigenvectors in terms of expectations defined via the dynamics of a Markov chain. The representation takes advantage of regenerative structure that is often available in the Markov setting. Assuming a simple condition (see A1 in Section 3), we can quickly argue the existence of a positive eigenvalue and corresponding positive row and column eigenvectors using our probabilistic methods; see the first part of the proof of Theorem 1.

But this probabilistic representation carries additional advantages. In particular, our methods extend directly to the setting of infinite-dimensional non-negative matrices and even to continuous state space non-negative operators. Such generalizations are useful in a number of applied probability settings, most particularly when analyzing the quasi-stationary behavior of a Markov chain living on a countable or continuous state space. In a follow-on paper, we will show how our probabilistic representation also allows one to develop Monte Carlo algorithms for computing the Perron-Frobenius eigenvalue and eigenvectors efficiently and without state space discretization, even in the setting of continuous state space. While some elements of the ideas we present here exist within the probability literature (see,

for example, Section 2 of [Ney and Nummelin 1987]), our discussion is intended to provide a more unified and accessible framework within which the key probability aspects can be easily understood and identified.

This paper is organized as follows. In Section 2, we study a particular infinite-dimensional non-negative matrix in some depth, to help clarify some of the complications that arise in extending the Perron-Frobenius theorem to the infinite-dimensional setting. In particular, the example shows that in such an infinite-dimensional context, a continuum of positive eigenvalues, each accompanied by positive row and column eigenvectors, can exist. However, only one such eigenvalue describes the growth rate of successive powers of entries of the matrix, and it is this eigenvalue that we view as the natural generalization of the finite-dimensional Perron-Frobenius eigenvalue. Thus, the Perron-Frobenius eigenvalue and eigenvectors are not characterized via positivity considerations in this infinite-dimensional setting. In Section 3, we focus on finite and infinite-dimensional matrices, and provide our probabilistic representation in Theorem 1. We then show that it corresponds to the Perron-Frobenius eigenvalue in Theorem 2. The rest of the section is devoted to a discussion of sufficient conditions guaranteeing the validity of the hypotheses underlying Theorems 1 and 2. Finally, in Section 4, we quickly show how the ideas naturally extend to continuous state space non-negative operators.

2 An Example

We wish to provide a version of the Perron-Frobenius theorem that is valid for infinite-dimensional matrices, and that can be readily extended to continuous state space. Before stating and proving this result in the next section, we wish to provide an example that illustrates the additional complications that can arise in the infinite-dimensional setting.

Consider a birth-death Markov chain $X = (X_n : n \geq 0)$ having state space $\mathbf{N}_+ = \{0, 1, 2, \ldots\}$, with transition probabilities

$$
\begin{aligned}
P(x, x+1) &= p & , x \geq 0 \\
P(x, x-1) &= q \quad (\triangleq 1 - p) & , x \geq 1,
\end{aligned}
$$

with $P(0,0) = q$ (with $0 < p < 1$). Such a Markov chain arises, for example, as a slotted-time queue in which the number of arrivals in each slot of time is either 2 (with probability p) or 0 (with probability q), and that can serve one customer per slot of time whenever a customer is available to be served. If we wish to study the dynamics of X conditional on the system not having emptied by time n, we are led to the consideration of probabilities of the form $P_x(X_n = y, T > n)$, where $P_x(.) \triangleq P(\cdot | X_0 = x)$ and $T = \inf \{n \geq 0 : X_n = 0\}$. In other words, we are computing the quasi-stationary behavior of X, associated with non-entry into $\{0\}$ over the time interval $[0, n]$. Clearly,

$$
P_x(X_n = y, T > n) = B^n(x, y),
$$

where $B = (B(x, y) : x, y \in S)$ is the non-negative infinite-dimensional matrix in which $B(x, y) = P(x, y)$ for $x, y \in S = \{1, 2, \ldots\}$. The matrix B is *irreducible*, in the sense that for each $x, y \in S$, there exists $n = n(x, y)$ for which $B^n(x, y) > 0$ and is sub-stochastic (since the row sum for state 1 is $p < 1$).

The natural extension of the notion of a Perron-Frobenius eigenvalue and (column) eigenvector from the finite dimensional setting to the infinite-dimensional matrix B involves consideration of the eigensystem

$$
Bu = \lambda u \tag{1}
$$

for positive solutions u and λ. For this example, this corresponds to studying the linear system

$$pu(x+1) + qu(x-1) = \lambda u(x), x \geq 2, \tag{2}$$

with the boundary condition

$$pu(2) = \lambda u(1). \tag{3}$$

To solve the recursive equation (2), it is standard practice to consider solutions that are linear combinations of functions of the form z^x, thereby leading to consideration of a quadratic equation for z, namely,

$$pz^2 - \lambda z + q = 0.$$

The roots of this equation are given by

$$z_1 = \frac{\lambda - \sqrt{\lambda^2 - 4pq}}{2p},$$

and

$$z_2 = \frac{\lambda + \sqrt{\lambda^2 - 4pq}}{2p}, \tag{4}$$

provided $\lambda \geq 2\sqrt{pq}$. If $0 \leq \lambda < 2\sqrt{pq}$, then z_1 and z_2 are complex, and no linear combination of z_1^x and z_2^x can be non-negative for all $x \geq 1$, so that positive eigenvectors cannot exist for λ in this range. We therefore now restrict ourselves to consideration of the case in which $\lambda \geq 2\sqrt{pq}$.

Suppose first that $\lambda > 2\sqrt{pq}$. If we take $u(x) = az_1^x + z_2^x = z_1^x(a + (z_2/z_1)^x)$, then equation (3) requires that

$$p(az_1^2 + z_2^2) = \lambda(az_1 + z_2),$$

in which case

$$a = -\frac{z_2}{z_1}\frac{\lambda - pz_2}{\lambda - pz_1}.$$

Since $0 < z_1 < z_2 < \frac{\lambda}{p}$, $a < 0$. However,

$$a + \left(\frac{z_2}{z_1}\right)^1 = \frac{z_2}{z_1}\left(1 - \frac{\lambda - pz_2}{\lambda - pz_1}\right)$$
$$= p\frac{z_2}{z_1}\frac{z_2 - z_1}{\lambda - pz_1} > 0,$$

so $a + \left(\frac{z_2}{z_1}\right)^x > 0$ for $x \geq 1$, and hence $u = (u(x) : x \geq 1)$ is a strictly positive column eigenvector solution to equation (1) for every $\lambda > 2\sqrt{pq}$. So, there is a continuum of positive λ's for which positive eigenvector solutions to equation (1) exist.

For $\lambda = 2\sqrt{pq}$, $z_1 = z_2 = \sqrt{\frac{q}{p}}$ and we are led to consideration of a solution to equation (2) of the form

$$u(x) = (q/p)^{\frac{x}{2}}(a + x)$$

for some choice of a. In this case, equation (3) requires that $a = 0$, so the solution $u(x) = x(q/p)^{\frac{x}{2}}$ for $x \geq 1$ is a positive column eigenvector solution of equation (1) when $\lambda = 2\sqrt{pq}$.

We can similarly consider the row eigenvector $\eta = (\eta(x) : x \in S)$ corresponding to eigenvalue λ, namely

$$\eta B = \lambda \eta. \tag{5}$$

3

Again, we are interested in positive solutions λ and η to equation (5). Here, the associated linear system is

$$p\eta(y-1) + q\eta(y+1) = \lambda\eta(y), y \geq 2, \tag{6}$$

subject to the boundary condition

$$q\eta(2) = \lambda\eta(1). \tag{7}$$

As in the column vector setting, η can be computed explicitly. The general solution takes the form

$$\eta(x) = \tilde{z}_1{}^x (b + (\frac{\tilde{z}_2}{\tilde{z}_1})^x)$$

for $x \geq 1$, where

$$\tilde{z}_1 = \frac{\lambda - \sqrt{\lambda^2 - 4pq}}{2q} \qquad (= \frac{1}{z_2}),$$

$$\tilde{z}_2 = \frac{\lambda + \sqrt{\lambda^2 - 4pq}}{2q} \qquad (= \frac{1}{z_1}),$$

To satisfy equation (7), we must then take

$$b = -\frac{\tilde{z}_2}{\tilde{z}_1}(\frac{\lambda - q\tilde{z}_2}{\lambda - q\tilde{z}_1}).$$

Because $0 < \tilde{z}_1 < \tilde{z}_2 < \frac{\lambda}{q}$, it follows that $(\eta(x) : x \geq 1)$ is a strictly positive row eigenvector solution to equation (5), for each $\lambda > 2\sqrt{pq}$. On the other hand, for $\lambda = 2\sqrt{pq}$, $\eta(x) = (\frac{p}{q})^{x/2}x$ for $x \geq 1$, so a positive row eigenvector also exists at this value of λ.

We conclude that in this example, there exist positive row and column eigenvector solutions to the eigensystem for B at all $\lambda \in [2\sqrt{pq}, \infty)$. So, the Perron-Frobenius eigenvalue and eigenvectors are not uniquely defined by positivity considerations in this infinite-dimensional irreducible setting.

We now explore this region of λ in more detail. Suppose u is a column eigenvector of B satisfying equation (1). Then, the matrix $\tilde{P} = (\tilde{P}(x, y) : x, y \in S)$ with entries given by

$$\tilde{P}(x, y) = \frac{B(x, y)u(y)}{\lambda u(x)}$$

for $x, y \in S$ is an irreducible stochastic matrix. Note that the Markov chain X associated with $\tilde{P}$ is again of birth-death form, with $\tilde{P}(1, 2) = 1$ and

$$\tilde{P}(x, x+1) = \frac{pu(x+1)}{\lambda u(x)},$$

$$\tilde{P}(x, x-1) = \frac{qu(x-1)}{\lambda u(x)},$$

for $x \geq 2$. For $\lambda > 2\sqrt{pq}$, $u(x) \sim z_2^x$ as $x \to \infty$, where we use the notation of $a(x) \sim b(x)$ as $x \to \infty$ to denote the fact that $a(x)/b(x) \to 1$ as $x \to \infty$.

Consequently,

$$\tilde{P}(x, x+1) \to \frac{p}{\lambda}z_2 = \frac{1}{2}(1 + \sqrt{1 - \frac{4pq}{\lambda^2}})$$

and

$$\tilde{P}(x, x-1) \to \frac{q}{\lambda}z_2^{-1} = \frac{1}{2}(1 - \sqrt{1 - \frac{4pq}{\lambda^2}})$$

4

as $x \to \infty$ Hence, $\tilde{P}$ is the one-step transition matrix of a transient Markov chain for $\lambda > 2\sqrt{pq}$. It follows that

$$\sum_{n=1}^{\infty} \tilde{P}^n(x, y) < \infty$$

for each $x, y \in S$; see, for example, p. 24 [Chung 1966]. This immediately implies that

$$\sum_{n}^{\infty} \lambda^{-n} B^n(x, y) < \infty \tag{8}$$

for each $x, y \in S$. On the other hand, for $\lambda = 2\sqrt{pq}$

$$\tilde{P}(x, x+1) = \frac{1}{2}(\frac{x+1}{x})$$

and

$$\tilde{P}(x, x-1) = \frac{1}{2}(\frac{x-1}{x})$$

for $x \geq 2$, while $\tilde{P}(1, 2) = 1$. Let $X = (X_n : n \geq 0)$ be the Markov chain having this transition matrix, and let $T_y = \inf\{n \geq 0 : X_n = yy\}$ be the hitting time of state y. If $v(x) = P_x(T_1 < \infty)$, then $v = (v(x) : x \geq 1)$ satisfies the linear system

$$v(x) = \tilde{P}(x, x+1)v(x+1) + \tilde{P}(x, x-1)v(x-1)$$

for $x \geq 2$, subject to $v(1) = 1$. The solution is

$$v(x) = \frac{\sum_{x}^{\infty} \frac{1}{y(y+1)}}{\sum_{1}^{\infty} \frac{1}{y(y+1)}}$$

for $x \geq 1$; see p. 31 [Hoel, Port, and Stone 1972]. Since $\sum_{1}^{\infty} y^{-2} < \infty$, evidently $v(2) < 1$, so that $\tilde{P}$ is the transition matrix of a transient Markov chain at $\lambda = 2\sqrt{pq}$. Consequently, equation (8) is also valid at $\lambda = 2\sqrt{pq}$.

To explore the behavior of the sum appearing in equation (8) for $\lambda < 2\sqrt{pq}$, we note that

$$\begin{aligned} B^{2n+2}(x, x) &= P_x(X_{2n+2} = x, T > 2n+2) \\ &\geq P(x, x+1)P_{x+1}(T_x = 2n+1) \\ &= pP_1(T_0 = 2n+1), \end{aligned}$$

where X here is the Markov chain associated with the slotted time queuing model. This latter probability is a well known first passage time quantity and is given by

$$P_1(T_0 = 2n+1) = \frac{1}{2n+1}\binom{2n+1}{n+1}p^n q^{n+1};$$

see, for example, p. 352 of [Feller 1950]. Sterling's approximation (see p. 52 of [Feller 1950]) implies that

$$P_1(T_0 = 2n+1) \sim cn^{-\frac{3}{2}}(2\sqrt{pq})^{2n+1}$$

as $n \to \infty$, for some constant $c \in (0, \infty)$. Consequently,

$$\sum_{n=0}^{\infty} \lambda^{-n} B^n(x, x) = \infty \tag{9}$$

5

for $\lambda < 2\sqrt{pq}$.

This discussion makes clear that while positive eigenvalues and eigenvectors exist for all $\lambda \geq 2\sqrt{pq}$, $\lambda = 2\sqrt{pq}$ holds special significance; compare equation (8) to (9). This state of affairs holds for general infinite-dimensional irreducible sub-stochastic matrices $G = (G(x,y) : x, y \in S)$. In particular, for all such matrices, there exists a finite-valued non-negative R such that for all $x, y \in S$,

$$\sum_{n=0}^{\infty} \beta^n G^n(x,y) < \infty$$

for $\beta < R$, whereas

$$\sum_{n=0}^{\infty} \beta^n G^n(x,y) = \infty$$

for $\beta > R$ (see p. 200-201 of [Seneta 1981]); the quantity R is called the *convergence parameter* of G, and it is therefore natural to view R^{-1} as the Perron-Frobenius eigenvalue λ for such a matrix G. However, at this level of generality, there is no universal guarantee that corresponding positive column eigenvectors for G will exist, either at $\lambda = R^{-1}$ or $\lambda > R^{-1}$. The example we have studied in this section is intended to illustrate the fact that while no general guarantee exists, such eigenvectors can potentially exist at all $\lambda \geq R^{-1}$ in this setting.

Under certain additional assumptions on G, there exist positive vectors $c = (c(x) : x \in S)$ and $d = (d(x) : x \in S)$ for which

$$G^n(x,y) \sim \lambda^n c(x) d(y) \tag{10}$$

as $n \to \infty$, analogous to the behavior manifested in the setting of finite-dimensional irreducible aperiodic non-negative matrices. In the next section, we provide a simple probabilistic proof for equation (10) under probabilistically appropriate sufficient conditions.

3   A Probabilistic Proof of the Perron-Frobenius Theorem

Let $G = (G(x,y) : x, y \in S$ be an irreducible non-negative matrix in which S can be either finite or countably infinite. We assume that

$$s = \sup_{x \in S} \sum_y G(x,y) < \infty,$$

so that by passing to $B = (B(x,y) : x, y \in S)$ with

$$B(x,y) = G(x,y)/s,$$

we may assume, without loss of generality, that we are dealing with an irreducible sub-stochastic matrix.

Our probabilistic representation depends on augmenting the state space S. Specifically, for $\Delta \notin S$, let $S_\Delta = S \cup \{\Delta\}$, and put

$$P(x,y) = \begin{cases} B(x,y) & , x, y \in S \\ 1 - \sum_{y \in S} B(x,y) & , x \in S, y = \Delta \\ 1 & , x, y = \Delta \\ 0 & , x = \Delta, y \in S. \end{cases}$$

Then, $P = (P(x,y) : x, y \in S_\Delta)$ is a stochastic matrix on S_Δ. To construct a positive eigenvalue λ_* and corresponding positive eigenvectors for B, we let $X =$

$(X_n : n \geq 0)$ be the Markov chain on S_Δ having transition matrix P, and we let $T = \inf\{n \geq 0; X_n = \Delta\}$ be the hitting time of Δ. Fix $z \in S$ and let $\tau = \inf\{n \geq 1 : X_n = z\}$ be the first return time to z. Note that the strong Markov property (SMP) implies that X regenerates at visits to z, in the sense that $(X_{\tau+n} : n \geq 0)$ has the same distribution as X under P_z, and is independent of $(\tau, X_j : 0 \leq j < \tau)$.

For $x \in S_\Delta$, let $E_x(\cdot) \triangleq E(\cdot|X_0 = x)$. For $x \in S$, let $T_x = \inf\{n \geq 0 : X_n = x\}$ and $\tau_x = \inf\{n \geq 1 : X_n = x\}$. We are now ready to state our key assumption, under which a Perron-Frobenius eigenvalue and column/row eigenvectors will exist.

A1. *There exists $\theta \geq 0$ such that*

$$E_z e^{\theta\tau} I(T > \tau) = 1.$$

In A1, we use the interpretation

$$I(T > \tau) = \sum_{k=1}^{\infty} I(\tau = k, T > k)$$

so that, in particular, τ is finite-valued on $\{T > \tau\}$ (even when $T = \infty$).

Set

$$\lambda_* = e^{-\theta},$$
$$u_*(x) = E_x e^{\theta\tau} I(T > \tau),$$
$$\eta_*(y) = E_z \sum_{j=0}^{\tau-1} e^{\theta j} I(X_j = y, T > j),$$

for $x, y \in S$.

Theorem 1. *Under A1, $u_* = (u_*(x) : x \in S)$ and $\eta_* = (\eta_*(y) : y \in S)$ are positive finite-valued column and row vectors respectively, and*

$$Bu_* = \lambda_* u_*, \tag{11}$$
$$\eta_* B = \lambda_* \eta_*. \tag{12}$$

Proof. Note that by conditioning on X_1, we observe that for $x \in S$,

$$u_*(x) = e^\theta \sum_{y \in S, y \neq z} B(x,y) u_*(y) + e^\theta B(x,z). \tag{13}$$

In view of A1, $u_*(z) = 1$, so that equation (13) implies that $u_* = e^\theta Bu_*$. As for η_*,

suppose $y \neq z$. Since $X_0 \neq y$ under P_x,

$$\eta_*(y) = E_z \sum_{j=0}^{\tau-1} e^{\theta j} I(X_j = y, T > j)$$

$$= E_z \sum_{j=1}^{\tau-1} e^{\theta j} I(X_j = y, T > j)$$

$$= e^\theta E_z \sum_{k=0}^{\infty} e^{\theta k} I(X_{k+1} = y, T \wedge \tau > k+1)$$

$$= e^\theta \sum_{x \in S} E_z \sum_{k=0}^{\infty} e^{\theta k} I(X_k = x, T \wedge \tau > k) B(x, y)$$

$$= e^\theta \sum_{x \in S} E_z \sum_{k=0}^{\tau-1} e^{\theta k} I(X_k = x, T > k) B(x, y)$$

$$= e^\theta \sum_{x \in S} \eta_*(x) B(x, y)$$

$$= e^\theta (\eta_* B)(y),$$

where we use the notational convention $a \wedge b \triangleq \min(a, b)$. On the other hand,

$$e^\theta \sum_{x \in S} \eta_*(x) B(x, z) = e^\theta \sum_{x \in S} E_z \sum_{j=0}^{\infty} e^{\theta j} I(X_j = x, T \wedge \tau > j) B(x, z)$$

$$= E_z \sum_{k=1}^{\infty} e^{\theta k} I(X_k = z, T > k, \tau > k-1)$$

$$= E_z \sum_{k=1}^{\infty} e^{\theta k} I(\tau = k, T > k)$$

$$= E_z e^{\theta \tau} I(T > \tau) = 1$$

$$= E_z \sum_{j=0}^{\tau-1} e^{\theta j} I(X_j = z, T > j)$$

$$= \eta_*(z),$$

where we used A1 for the third last equality, and the fact that under P_z, $X_0 = z$ with $X_j \neq z$ for $1 \leq j < \tau$ for the second last equality. So, $\eta_* = e^\theta \eta_* B$.

We now turn to the finiteness and positivity of u_* and η_*. Set $\tau_y = \inf\{n \geq 1 : X_n = y\}$. Note that the irreducibility of B implies that $P_y(\tau < \tau_y \wedge T) > 0$ and $P_z(\tau_y < \tau \wedge T) > 0$ for $y \in S$. So,

$$u_*(y) = E_y e^{\theta \tau} I(T > \tau)$$

$$\geq E_y e^{\theta \tau} I(\tau_y \wedge T > \tau)$$

$$\geq P_y(\tau_y \wedge T > \tau) > 0.$$

Also, A1 and the SMP imply that

$$1 = E_z e^{\theta \tau} I(T > \tau)$$

$$\geq E_z e^{\theta \tau_y} I(\tau_y < \tau \wedge T) E_y e^{\theta \tau} I(T > \tau)$$

$$\geq P_z(\tau_y < \tau \wedge T) u_*(y),$$

so $u_*(y) < \infty$ for $y \in S$.

As for η_*, note that for $y \neq z$,

$$\sum_{j=0}^{\tau-1} e^{\theta j} I(X_j = y, T > j) > 0$$

on $\{\tau_y < \tau \wedge T\}$. So, $\eta_*(y) > 0$ since $P_z(\tau_y < \tau \wedge T) > 0$. Since we showed earlier that $\eta_*(z) = 1$, η_* is therefore positive. Also, the SMP implies that

$$\begin{aligned}\infty > u_*(y) =& E_y e^{\theta \tau} I(\tau < \tau_y \wedge T) \\ &+ E_y e^{\theta \tau_y} I(\tau_y < \tau \wedge T) u_*(y),\end{aligned}$$

and hence it follows that $E_y e^{\theta \tau_y} I(\tau_y < \tau \wedge T) < 1$. So, another application of the SMP yields

$$\eta_*(y) = E_z e^{\theta \tau_y} I(\tau_y < \tau \wedge T) E_y \sum_{j=0}^{\tau-1} e^{\theta j} I(X_j = y, T > j). \tag{14}$$

Note that $E_z e^{\theta \tau_y} I(\tau_y < T \wedge \tau) \leq E_z e^{\theta \tau} I(T > \tau) = 1$. If $\beta(y) \triangleq E_y \sum_{j=0}^{\tau-1} e^{\theta j} I(X_j = y, T > j)$, we see that

$$\beta(y) = 1 + E_y e^{\theta \tau_y} I(\tau_y < \tau \wedge T) \beta(y).$$

Since $E_z e^{\theta \tau_y} I(\tau_y < T \wedge \tau) < 1$, $\beta(y) < \infty$, and equation (14) therefore implies that $\eta_*(y) < \infty$, proving that η_* is finite valued. $\qquad\square$

With an additional assumption, much more can be said about λ_*, η_*, and u_*. In particular, we can show that λ_*^{-1} is indeed the convergence parameter of B, and u_* and η_* do indeed correspond to the Perron-Frobenius eigenvectors of the finite-dimensional theory.

A2. *With θ defined as in A1, assume that*

$$E_z \tau e^{\theta \tau} I(T > \tau) < \infty.$$

Recall that the period p of a non-negative irreducible matrix is always uniquely defined (as a *class property*) and finite-valued (even in the infinite-dimensional setting); see p. 20 [Nummelin 1984].

Theorem 2. *Assume A1 and A2. Then:*
  (i) *If u is a positive column vector satisfying $Bu = \lambda_* u$, then u is a positive multiple of u_*.*
  (ii) *If η is a positive row vector satisfying $\eta B = \lambda_* \eta$, then η is a positive multiple of η_*.*
  (iii) $\sum_{w \in S} \eta_*(w) u_*(w) < \infty.$
  (iv) *If p is the period of B, then*

$$\frac{1}{p} \sum_{j=0}^{p-1} \lambda_*^{-pn-j} B^{pn+j}(x, y) \to \frac{u_*(x) \eta_*(y)}{\sum_{w \in S} u_*(w) \eta_*(w)}$$

  as $n \to \infty$.

Proof. We start by observing that

$$\sum_{x \in S} \eta_*(x) u_*(x) = \sum_{x \in S} E_z \sum_{j=0}^{\tau-1} e^{\theta j} I(X_j = x, T > j) u_*(x)$$

$$= \sum_{x \in S} E_z \sum_{j=0}^{\infty} e^{\theta j} I(T \wedge \tau > j) I(X_j = x) E_x e^{\theta \tau} I(T > \tau)$$

$$= \sum_{x \in S} E_z \sum_{j=0}^{\infty} e^{\theta \tau} I(T \wedge \tau > j, X_j = x)$$

$$= E_z \sum_{j=0}^{\tau-1} e^{\theta \tau} I(T \wedge \tau > j)$$

$$= E_z \tau e^{\theta \tau} I(T > \tau) < \infty$$

by A2.

Set

$$P_*(x, y) = \frac{B(x, y) u_*(y)}{\lambda_* u_*(x)}$$

for $x, y \in S$. Observe that P_* is an irreducible stochastic matrix that inherits the same periodicity as does B. If $\pi_*(x) = u_*(x)\eta_*(x) / \sum_{w \in S} u_*(w)\eta_*(w)$, then $\pi_* = (\pi_*(x) : x \in S)$ is a probability on S, and

$$\sum_{x \in S} \pi_*(x) P_*(x, y) = \sum_{x \in S} \frac{\eta_*(x) B(x, y)}{\lambda_*} u_*(y) \cdot \frac{1}{\sum_{w \in S} u_*(w)\eta_*(w)}$$

$$= \frac{\eta_*(y) u_*(y)}{\sum_{w \in S} u_*(w)\eta_*(w)}$$

$$= \pi_*(y),$$

so that $\pi_* = (\pi_*(x) : x \in S)$ is a stationary distribution of P_*. Consequently, the Markov chain $X = (X_n : n \geq 0)$ having transition matrix P_* is positive recurrent.

Suppose that $u = (u(x) : x \in S)$ is a positive solution of $Bu = \lambda_* u$. Then, if $w(x) \triangleq u(x)/u_*(x)$ for $x \in S$,

$$\sum_{y \in S} P_*(x, y) w(y) = w(x) \tag{15}$$

for $x \in S$. It follows from equation (15) that $(w(X_n) : n \geq 0)$ is a positive martingale adapted to the history of X. The supermartingale convergence theorem (see, for example, p. 335 [Chung 1974]) then establishes that there exists a finite-valued random variable (r.v.) M_∞ such that

$$w(X_n) \to M_\infty \qquad P_x \quad a.s.$$

as $n \to \infty$. Since X visits each state $x \in S$ infinitely often (due to recurrence), the a.s. (almost sure) convergence implies that $w(\cdot)$ must be constant. In other words, u must be a positive multiple of u_*.

Suppose that $\eta = (\eta(y) : y \in S)$ is a positive solution of $\eta B = \lambda_* \eta$. Put $\pi(x) = \eta(x) u_*(x)$, and note that $\pi = (\pi(x) : x \in S)$ satisfies

$$\sum_{x \in S} \pi(x) P_*(x, y) = \pi(y)$$

for $y \in S$, so that π is a positive solution of the equilibrium equation for P_*. This implies that π must be a multiple of π_*; see Theorem 9.2 [Nummelin 1984]. This, in turn, proves that η must be a positive multiple of η_*.

Finally, the positive recurrence of X implies that for each $x, y \in S$,

$$\frac{1}{p}\sum_{j=0}^{p-1} P_*^{np+j}(x,y) \to \pi_*(y)$$

as $n \to \infty$; (see p. 98 of [Gantmacher 1959] for the proof holds when $|S| < \infty$; an analogous proof when $|S| = \infty$). But this is exactly equivalent to part *iv)* of the theorem. $\qquad\square$

An easy consequence of part *iv)* is that λ_* is the convergence parameter of B. Thus, λ_* describes the decay rate of B^n, and u_* and η_* arise naturally in describing the large n behavior of B^n; see part *iv* of the theorem . Thus, λ_*, u_*, and η_* are the natural Perron-Frobenius eigenvalue, column eigenvector, and row eigenvector associated with B.

We now turn to a simple sufficient condition guaranteeing A1 and A2. Let

$$\theta_1 = \sup\{\gamma : E_z e^{\gamma T} < \infty\},$$
$$\theta_2 = \sup\{\gamma : E_z e^{\gamma(T \wedge \tau)} < \infty\}.$$

Theorem 3. *If $\theta_2 > \theta_1$, then A1 and A2 hold.*

Proof. Note that for $\gamma \in \mathbf{R}$, the SMP implies that

$$E_z e^{\gamma T} = E_z e^{\gamma \tau} I(T > \tau) E_z e^{\gamma T} + E_z e^{\gamma T} I(T < \tau). \qquad (16)$$

Choose $\theta_0 \in (\theta_1, \theta_2)$. By definition of θ_1 and θ_2, it follows that $E_z e^{\theta_0 T} = \infty$ but $E_z e^{\theta_0(T \wedge \tau)} < \infty$. In view of equation (16), we are then led to the conclusion that $E_z e^{\theta_0 \tau} I(T > \tau) \geq 1$. Since $h(\gamma) \triangleq E_z e^{\gamma \tau} I(T > \tau) \leq E_z e^{\gamma(T \wedge \tau)} < \infty$, the function $h(\cdot)$ is finite-valued, strictly increasing, and infinitely differentiable on $(-\infty, \theta_2]$. Hence, there exists a unique $\theta < \theta_0$ such that $E_z e^{\theta \tau} I(T > \tau) = 1$, proving A1.

Furthermore, by an argument analogous to that used in studying moment generating functions (see, for example, p. 119 of [Billingsley 1979]),

$$h'(\gamma) = E_z \tau e^{\gamma \tau} I(T > \tau),$$

for $\gamma < \theta_2$, so that $E_z \tau e^{\theta \tau} I(T > \tau) = h'(\theta) < \infty$, proving A2. $\qquad\square$

Given the importance of A1 and A2, it is of some interest to determine whether A1 and A2 are *solidarity properties*, in the sense that if they hold at one $z \in S$, then they hold together at all $z \in S$. Let $P_*(\cdot)$ be the probability in the path-space of X under which X has transition matrix $(P_*(x,y) : x, y \in S)$, and let $E_*(\cdot)$ be the corresponding expectation operator. Then, for any function $f : S^{n+1} \to \mathbf{R}_+$,

$$E_*[f(X_0, \ldots, X_n)|X_0 = x]$$

$$= \sum_{x_1, \ldots, x_n} f(x, x_1, \ldots, x_n) P_*(x, x_1) P_*(x_1, x_2) \cdots P_*(x_{n-1}, x_n)$$

$$= \sum_{x_1, \ldots, x_n} f(x, x_1, \ldots, x_n) B(x, x_1) B(x_1, x_2) \cdots B(x_{n-1}, x_n) \frac{u_*(x_n)}{u_*(x)} \lambda_*^{-n}$$

$$= E_x e^{\theta n} f(X_0, \ldots, X_n) I(T > n) \frac{u_*(X_n)}{u_*(X_0)}.$$

Hence, since $X_n = y$ on $\{\tau_y = n\}$,

$$P_*(\tau_y = n | X_0 = y) = E_y e^{\theta n} I(\tau_y = n, T > n). \qquad (17)$$

Since X is recurrent under P_*,

$$1 = P_*(\tau_y < \infty | X_0 = y) = E_y e^{\theta \tau_y} I(\tau_y < T),$$

proving that A1 holds with $y \in S$ replacing z. As for A2, equation (17) yields

$$\begin{aligned}
E_y \tau_y e^{\theta \tau_y} I(T > \tau_y) &= \sum_{n=1}^{\infty} n E_y e^{\theta n} I(\tau_y = n, T > n) \\
&= \sum_{n=1}^{\infty} n P_*(\tau_y = n | X_0 = y) \\
&= E_*[\tau_y | X_0 = y].
\end{aligned}$$

But since X is positive recurrent under P_* (and positive recurrence is a solidarity property),

$$E_*[\tau_y | X_0 = y] < \infty$$

for all $y \in S$, so that A2 also holds for all $y \in S$. We have therefore established the following property.

Proposition 1. *If A1 and A2 hold at one $z \in S$, then A1 and A2 hold at each $y \in S$.*

So, A1 and A2 are indeed solidarity properties, so we are free to choose any $z \in S$ to verify these conditions.

Our final goal in this section is to provide an easily verifiable condition on B guaranteeing that $\theta_2 > \theta_1$ (as in the statement of Theorem 3. Specifically, consider the hypothesis:

A3. *There exists $v \in S$ and c_1, $c_2 \in \mathbf{R}_+$ such that*

$$0 < c_1 = \inf_{x \in S} \frac{B(x, y)}{B(v, y)} \le \sup_{x \in S} \frac{B(x, y)}{B(v, y)} = c_2 < \infty.$$

Assumption A3 requires that the rows of B be similar, in the sense that they must all be within a uniform constant multiple of one another. This assumption is automatic when B is a finite-dimensional strictly positive matrix, but is a (very) strong condition when B is infinite-dimensional.

Put $\psi(y) = B(v, y) / \sum_w B(v, w)$ for $y \in S$ and set $\psi = (\psi(y) : y \in S)$. To exploit A3, we need to use a regenerative structure for X based on the distribution of ψ, rather than regenerations based on the consecutive hitting times of z. In particular, note that by definition of c_1, $B(x, y) \ge \delta \psi(y)$ for $x, y \in S$ with $\delta = c_1 \sum_w B(v, w)$ so that

$$B(x, y) = \delta \psi(y) + \tilde{B}(x, y) \tag{18}$$

for $x, y \in S$, where $\tilde{B}(x, y) \triangleq B(x, y) - \delta \psi(y) \ge 0$ (so that $\tilde{B} = (\tilde{B}(x, y) : x, y \in S)$ is a sub-stochastic matrix). By summing both sides of equation (18) over $y \in S$, we conclude that $\delta \le 1$. Hence, we may view a transition of X at time n as follows:

- With probability δ, distribute X_{n+1} according to ψ;

- With probability $\tilde{B}(X_n, y)$, set $X_{n+1} = y$;

- With probability $1 - \sum_{y \in S} B(X_n, y)$, set $X_{n+1} = \Delta$.

At times τ at which X distributes itself according to ψ, X regenerates and has distribution ψ, independent of $X_0, \ldots, X_{\tau-1}$, and τ. Let $E_\psi(\cdot) \triangleq \sum_x \psi(x) E_x(\cdot)$ and $P_\psi(\cdot) \triangleq \sum_x \psi(x) P_x(\cdot)$. Put

$$\theta_1 = \sup\{\theta : E_\psi e^{\theta T} < \infty\},$$
$$\theta_2 = \sup\{\theta : E_\psi e^{\theta(T \wedge \tau)} < \infty\}.$$

Theorems 1 through 3 generalize naturally to the randomized regenerations τ just introduced.

Theorem 4. *If $\theta_2 > \theta_1$, then there exists a solution θ to the equation*

$$E_\psi e^{\theta\tau} I(T > \tau) = 1. \tag{19}$$

Put $\lambda_* = e^{-\theta}$, and set

$$u_*(x) = E_x e^{\theta\tau} I(T > \tau),$$

$$\eta_*(y) = E_\psi \sum_{j=1}^{\tau-1} e^{\theta j} I(X_j = y, T > \tau).$$

Then:
(i) $u_* = (u_*(x) : x \in S)$ is a positive finite-valued solution of $Bu = \lambda_* u$. Any other positive finite-valued solution must be a positive multiple of u_*.
(ii) $\eta_* = (\eta_*(y) : y \in S)$ is a positive finite-valued solution of $\eta B = \lambda_* \eta$. Any other positive finite-valued solution must be a positive multiple of η_*.
(iii) $\sum_{w \in S} \eta_*(w) u_*(w) < \infty$.
(iv) If p is the period of B, then

$$\frac{1}{p} \sum_{j=0}^{p-1} \lambda_*^{-pn-j} B^{pn+j}(x, y) \to \frac{u_*(x) \eta_*(y)}{\sum_{w \in S} \eta_*(w) u_*(w)}$$

as $n \to \infty$.

Proof. The proofs of Theorems 2 and 3 extend to this setting without change. It follows that there exists a root θ to equation (19). Note that equation (19) implies that

$$u_*(x) = e^\theta P_x(\tau = 1) + e^\theta \sum_{y \in S} \tilde{B}(x, y) u_*(y)$$

$$= e^\theta \delta + e^\theta \sum_{y \in S} \tilde{B}(x, y) u_*(y)$$

$$= e^\theta \delta \sum_{y \in S} \psi(y) u_*(y) + e^\theta \sum_{y \in S} \tilde{B}(x, y) u_*(y)$$

$$= e^\theta \sum_{y \in S} [\delta \psi(y) + \tilde{B}(x, y)] u_*(y)$$

$$= e^\theta \sum_{y \in S} B(x, y) u_*(y) = \lambda_*^{-1} (B u_*)(x)$$

for $x \in S$, verifying the fact that u_* is a column eigenvector satisfying $Bu = \lambda_* u$.

As for η_*,

$$e^\theta \sum_{x \in S} \eta_*(x) B(x,y) = E_\psi \sum_{x \in S} \sum_{j=0}^\infty I(X_j = x, T \wedge \tau > j) e^{\theta(j+1)} B(x,y)$$

$$= E_\psi \sum_{x \in S} \sum_{j=1}^\infty e^{\theta j} I(X_{j-1} = x, T \wedge \tau > j-1)(\tilde{B}(x,y) + \delta \psi(y))$$

$$= E_\psi \sum_{j=1}^{\tau-1} e^{\theta j} I(X_j = y, T > j) + E_\psi e^{\theta \tau} I(T > \tau) \psi(y)$$

$$= \psi(y) + E_\psi \sum_{j=1}^{\tau-1} e^{\theta j} I(X_j = y, T > j)$$

$$= E_\psi \sum_{j=0}^\infty e^{\theta j} I(X_j = y, T > j)$$

$$= \eta_*(y),$$

proving that η_* is a solution of $\eta B = \lambda_* \eta$.

As for the positivity of u_*, this follows from the fact that $u_*(y) \geq e^\theta P_y(\tau = 1) = \delta e^\theta > 0$. For η_*, note that $\eta_*(y) \geq E_\psi I(T_y < \tau \wedge T) e^{\theta T_y} > 0$ if $P_\psi(T_y < T \wedge \tau) > 0$. Suppose, however, that $P_\psi(T_y < T \wedge \tau) = 0$. Then, the regenerative structure of τ guarantees that $P_\psi(T_y < T \wedge \tau_n) = 0$, where τ_n is the time of the nth randomized regeneration. Sending $n \to \infty$, we conclude that $P_\psi(T_y < T) = 0$, contradicting the irreducibility of B.

To obtain the finiteness of u_*, observe that

$$1 = E_\psi e^{\theta \tau} I(T > \tau)$$
$$\geq P_\psi(T_y < T \wedge \tau) u_*(y),$$

and use the just established positivity of $P_\psi(T_y < T \wedge \tau)$. As for the finiteness of η_*, note that Theorems 2 and 3 establish that $E_\psi \tau e^{\theta \tau} I(T > \tau) < \infty$. But $\tau e^{\theta \tau} I(T > \tau) \geq \sum_{j=0}^{\tau-1} e^{\theta j} I(X_1 = y, T > j)$, yielding the result. The uniqueness of u_* and η_* follow as in Theorem 2. $\square$

We are now ready to prove our next result.

Theorem 5. *If A3 holds, then the conclusions of* Theorem 4 *are in force.*

Proof. It remains only to show that $\theta_2 > \theta_1$. Note that

$$P_x(T \wedge \tau > n) = \sum_{x_1, \ldots, x_n} \tilde{B}(x,x_1) \tilde{B}(x_1,x_2) \cdots \tilde{B}(x_{n-1},x_n). \tag{20}$$

But A3 implies that $B(x,y) \leq (c_2/c_1) \delta \psi(y)$, so $-\delta \psi(y) \leq (c_1/c_2) B(x,y)$. Consequently,

$$\tilde{B}(x,y) \leq (1 - \frac{c_1}{c_2}) B(x,y)$$

for $x, y \in S$. Substituting this inequality into equation (20) yields

$$P_x(T \wedge \tau > n) \leq (1 - \frac{c_1}{c_2})^n \sum_{x_1, \ldots, x_n} B(x,x_1) \cdots B(x_{n-1},x_n)$$

$$= (1 - \frac{c_1}{c_2})^n P_x(T > n).$$

Since $\sum_{k=0}^{j-1} e^{\theta k}(e^{\theta} - 1) = e^{\theta j} - 1$, evidently

$$E_x e^{\theta(T \wedge \tau)} = 1 + (e^{\theta} - 1)E_x \sum_{k=0}^{(T \wedge \tau) - 1} e^{\theta k}$$

$$= 1 + (e^{\theta} - 1)\sum_{k=0}^{\infty} e^{\theta k} P_x(T \wedge \tau > k)$$

$$\leq 1 + (e^{\theta} - 1)\sum_{k=0}^{\infty} e^{\theta k}(1 - \frac{c_1}{c_2})^k P_x(T > k)$$

$$= 1 + (e^{\theta} - 1)E_x \sum_{k=0}^{T-1} (e^{\theta}(1 - \frac{c_1}{c_2}))^k$$

$$= 1 + \frac{(e^{\theta} - 1)}{(e^{\hat{\theta}} - 1)}(E_{\psi} e^{\hat{\theta}T} - 1),$$

where $\hat{\theta} = \theta + \log(1 - \frac{c_1}{c_2})$. We conclude that $\theta_2 \geq \theta_1 - \log(1 - (c_1/c_2)) > \theta_1$, proving the result. $\qquad\square$

In fact, under A3, the period of B is obviously one (since $B(y, y) > 0$ for each $y \in S$ for which $\psi(y) > 0$), so we conclude that A3 ensures that

$$B^n(x, y) \sim \lambda_*^n \frac{u_*(x)\eta_*(y)}{\sum_{w \in S} u_*(w)\eta_*(w)}$$

as $n \to \infty$. In other words, A3 guarantees the validity of equation (10).

Actually, A3 can be further generalized.

A3′. *There exists $m \geq 1, v \in S$, and $c_1, c_2 \in \mathbf{R}_+$ such that:*

$$0 < c_1 = \inf_{x \in S} \frac{B^m(x, y)}{B^m(v, y)} \leq \sup_{x \in S} \frac{B^m(x, y)}{B^m(v, y)} = c_2 < \infty.$$

Note that when B is finite-dimensional and has period one, the irreducibility that we are requiring throughout this section guarantees that B^m is strictly positive for some $m \geq 1$, and A3′ is then immediate.

Assuming A3′, we observe that if we choose $y \in S$ so that $B^m(v, y) > 0$, then $B^{2m}(y, y) \geq B^m(y, y)B^m(y, y) \geq c_1^2[B(v, y)]^2$ and

$$B^{2m+1}(y, y) \geq B^m(y, y) \sum_{x \in S} B(y, x)B^m(x, y)$$

$$\geq c_1^2 B^m(v, y) \sum_x B(y, x)B^m(v, y) > 0,$$

so the greatest common divisor of $\{k \geq 1 : B^k(y, y) > 0\}$ equals 1. Hence, B has period one, which implies that B^m is irreducible. Consequently, Theorem 5 can be applied to B^m, so that the conclusions of Theorem 4 apply also to B^m.

Note also that if u_*, η_*, and λ_* are as in Theorem 4 (when applied to B^m),

$$B^m u_* = \lambda_* u_*$$

and

$$\eta_* B^m = \lambda_* \eta_*.$$

But this implies that

$$B^m(Bu_*) = \lambda_* B u_* \tag{21}$$

15

and

$$(\eta_* B)B^m = \lambda_* \eta_* B, \tag{22}$$

so that $\eta_* B$ and Bu_* are positive row and column eigenvectors associated with B^m. The uniqueness of u_* and η_* (up to positive multiples) then implies that $Bu_* = a_1 u_*$ and $\eta_* B = a_2 \eta_*$ for some $a_1, a_2 > 0$. Consequently, $B^m u_* = a_1^m u_*$ and $\eta_* B^m = a_2^m \eta_*$. The positivity of u_* and η_* then ensure that $a_1^m = \lambda_* = a_2^m$, due to equations (21) and (22). So, λ_*, u_*, and η_* satisfy

$$Bu_* = \lambda_*^{1/m} u_*,$$

and

$$\eta_* B = \lambda_*^{1/m} \eta_*,$$

establishing the following result.

Theorem 6. *If A3′ holds, then we can define λ_*, u_*, and η_* by applying Theorem 4 to B^m. Furthermore,*

$$B^n(x,y) \sim \lambda_*^{n/m} \frac{u_*(x)\eta_*(y)}{\sum_{w \in S} u_*(w)\eta_*(w)}$$

as $n \to \infty$, and u_* and η_* are the unique (up to a positive multiple) column and row eigenvectors of B associated with eigenvalue $\lambda_*^{1/m}$.

4 Extension to Continuous State Space

In this section, we briefly indicate how the theory of Section 3 easily extends to continuous state space. In particular, suppose that S is a metric space equipped with the measurable subsets induced by its metric topology, and suppose that $B = (B(x,dy) : x,y \in S)$ is a non-negative sub-stochastic kernel, so that for each $x \in S$, $B(x,\cdot)$ is a non-negative measure on S for which $B(x,S) \le 1$. We further assume that there exists a probability $\nu(\cdot)$ on S and $m \ge 1$ so that for $x, y \in S$,

$$B^m(x,dy) = b_m(x,y)\nu(dy)$$

for some positive density $(b_m(x,y) : x,y \in S)$. Finally, we assume that:

A4. *There exists $v \in S$ and $c_1, c_2 \in \mathbf{R}_+$ such that*

$$0 < c_1 = \inf_{x \in S} \frac{b_m(x,y)}{b_m(v,y)} \le \sup_{x \in S} \frac{b_m(x,y)}{b_m(v,y)} = c_2 < \infty.$$

For (measurable) $C \subseteq S$, set

$$\psi(C) = \frac{\int_C b_m(v,y)\nu(dy)}{\int_S b_m(v,y)\nu(dy)}.$$

As in Section 3, we can now write

$$B^m(x,dy) = \delta\psi(dy) + \tilde{B}(x,dy) \tag{23}$$

for $x, y \in S$, and let τ be the first $k \ge 1$ at which X_{km} has the distribution ψ. Under A4, the proof of Theorem 5 holds without change, and we therefore obtain the existence of $\theta \ge 0$ such that

$$E_\psi e^{\theta\tau} I(T > \tau m) = 1$$

16

with

$$E_\psi \tau e^{\theta\tau} I(T > \tau m) < \infty.$$

Furthermore, if we set $\lambda_* = e^{-\theta}$ and

$$u_*(x) = E_x e^{\theta\tau} I(T > \tau m),$$

$$\eta_*(C) = E_\psi \sum_{j=0}^{\tau-1} e^{\theta j} I(X_{jm} \in C, T > jm)$$

for $x \in S$ and (measurable) $C \subseteq S$, the following theorem holds.

Theorem 7. *If A4 holds, then for $x \in S$ and (measurable) $C \subseteq S$ for which $\eta_*(C) > 0$,*

$$B^n(x, C) \sim \lambda_*^{n/m} \frac{u_*(x)\eta_*(C)}{\int_S u_*(w)\eta_*(dw)}$$

as $n \to \infty$.

Proof. Under A4, the proof of Theorem 4 shows that

$$B^m u_* = \lambda_* u_* \tag{24}$$

and

$$\eta_* B^m = \lambda_* \eta_*. \tag{25}$$

The proof of Theorem 4 also shows that both $u_*(\cdot)$ and $\eta_*(\cdot)$ are finite-valued. We further note that $u_*(x) \geq e^\theta P_x(T > m, \tau = 1) = \lambda^{-1}\delta$, so $u_*(\cdot)$ is bounded below by a positive constant. Also, A4 and equation (24) imply that

$$\begin{aligned}
\lambda_* u_*(x) &= (B^m u_*)(x) \\
&= \int_S b_m(x, y) u_*(y) \nu(dy) \\
&\leq c_2 \int_S b_m(v, y) u_*(y) \nu(dy) \\
&= c_2 (B^m u_*)(v) \\
&= c_2 \lambda_* u_*(v)
\end{aligned}$$

so $u_*(x) \leq c_2 u_*(v)$.

Set

$$P_*(x, dy) = \frac{B^m(x, dy) u_*(y)}{\lambda_* u_*(x)}$$

for $x, y \in S$, and note that $P_* = (P_*(x, dy) : x, y \in S)$ is (as in Theorem 2) a stochastic kernel. Furthermore, equation (23) and the upper bound on u_* imply that

$$P_*(x, dy) \geq \frac{\delta \psi(dy) u_*(y)}{\lambda_* u_*(x)} \geq \frac{\delta}{\lambda_*} \frac{\psi(dy) u_*(y)}{c_2 u_*(v)} \tag{26}$$

uniformly in $x \in S$. It follows that X is a uniformly ergodic Markov chain under P_*; see p. 394 of [Meyn and Tweedie 2009] for the definition and basic properties. As in Section 3, π_* is a stationary distribution of P_*, where

$$\pi_*(dx) = \frac{u_*(x)\eta_*(dx)}{\int_S u_*(w)\eta_*(dw)}$$

for $x \in S$. The uniform ergodicity of P_* implies that π_* is the unique sigma-finite nonnegative solution of $\pi P_* = \pi$ (up to positive multiples; see p. 234 of

[Meyn and Tweedie 2009]) and this implies, as in Section 3, that η_* is the unique (up to positive multiples) solution of $\eta B^m = \lambda_* \eta$. Furthermore, if u is a nonnegative solution of $B^m u = \lambda_* u$, then $P_* w = w$, where $w(x) = u(x)/u_*(x)$ for $s \in S$. Again, since $(w(X_n) : n \geq 0)$ is a non-negative supermartingale, $w(X_n) \to M_\infty$ a.s. as $n \to \infty$, where M_∞ is finite-values a.s. Since X hits every set with positive ν-measure infinitely often a.s., this implies that there can not exist $d \in \mathbf{R}$ and $\epsilon > 0$ such that $\{x : w(x) < d - \epsilon\}$ and $\{x : w(x) > d + \epsilon\}$ both have positive ν-measure. Consequently, w nust be a constant ν-a.e. In view of A4, this implies that $P_* w$ must be constant, so that w is constant. This proves that u_* is the unique positive solution of $B^m = \lambda_* u$ (up to positive multiples). Consequently, as in the argument leading to Theorem 6, we find that

$$Bu_* = \lambda_*^{1/m} u_*$$

and

$$\eta_* B = \lambda_*^{1/m} \eta_*.$$

We can now define the stochastic kernel,

$$\tilde{P}_*(x, dy) = \frac{B(x, dy)u_*(y)}{\lambda_*^{1/m} u_*(x)}$$

for $x, y \in S$ and note that $\tilde{P}_*^m = P_*$. Thus, equation (26) ensures that

$$\tilde{P}_*^m(x, dy) \geq \psi(dy)u_*(y)$$

uniformly in $x \in S$, so that if X evolves on S under $\tilde{P}_*$, X is uniformly ergodic. Since π_* is also stationary for $\tilde{P}_*$, this implies that

$$\tilde{P}_*^n(x, C) \to \pi_*(c)$$

as $n \to \infty$, which proves the theorem. $\qquad\square$

Theorem 7 is, of course, the natural continuous state space analog to the Perron-Frobenius theorem that is available in the finite-dimensional setting.

Acknowledgements

References

[Billingsley 1979]   P. Billingsley. *Probability and Measure*, Wiley, 1979.

[Chung 1966]   K. L. Chung, *Markov Chains with Stationary Transition Probabilities*, Springer-Verlag, Berlin, 1966.

[Chung 1974]   K. L. Chung, *A Course in Probability Theory*, 2nd ed., Academic Press, New York, 1974.

[Feller 1950]   W. Feller, *An Introduction to Probability Theory and its Applications, Vol I*, John Wiley and Sons, New York, 1950.

[Gantmacher 1959]   F. R. Gantmacher, *The Theory of Matrices, Vol II* , Chelsea Publishing Company, New York, 1959.

[Hoel, Port, and Stone 1972] P. G. Hoel, S. C. Port, C. J. Stone, *Introduction to Stochastic Processes*, Houghton Mifflin Company, Boston, 1972.

[Meyn and Tweedie 2009] S. Meyn, R. Tweedie, *Markov Chains and Stochastic Stability*, Cambridge University Press, 2009.

[Nummelin 1984]   E. Nummelin, *General Irreducible Markov Chains and Non-Negative Operators (Cambridge Tracts in Mathematics)*, Cambridge University Press, 1984.

[Ney and Nummelin 1987] P. Ney, E. Nummelin, Markov Additive Processes I. Eigenvalue Properties and Limit Theorems. *Annals of Probability*, V 15, No 2, 1987, 561-592

[Pliska 1976]   S. R. Pliska, Optimization of multitype branching processes, *Management Science*, 23 (1976), 117-124.

[Rothblum 1984]   Rothblum, U. G. (1984). Multiplicative Markov Decision Chains. *Mathematics of Operations Research* 9, 1, 6-24.

[Rothblum and Veinott 1975] Rothblum, U. and Veinott, A. F., Jr. (1975). Cumulative Average Optimality for Normalized Markov Decision Chains. Draft version. Unpublished

[Rothblum and Veinott 1992] U. G. Rothblum, A. F. Veinott, Jr., Markov Branching Decision Chains: Immigration-Induced Optimality. Technical Report No. 45, Department of Operations Research, Stanford University, 1992.

[Rothblum and Whittle 1982] U. G. Rothblum and P. Whittle, Growth Optimality for Branching Markov Decision Chains. *Mathematics of Operations Research* Vol. 7, No. 4, Nov., 1982, 582-601.

[Seneta 1981]   E. Seneta, *Non-negative Matrices and Markov Chains*, Springer-Verlag, 1981.