Stochastic Operators, and an Infinite Dimensional Version of the Perron-Frobenius Theorem

Michael Taylor

1. Introduction

Let X be a compact Hausdorff space. The space $C(X)$ of continuous, real-valued functions on X is a Banach space, with norm

$$\|f\|_{\sup} = \sup_{x\in X} |f(x)|. \tag{1.1}$$

Let

$$A : C(X) \longrightarrow C(X) \tag{1.2}$$

be a bounded linear map. We say A is *positive* if

$$f \in C(X),\ f \ge 0 \Longrightarrow Af \ge 0. \tag{1.3}$$

We say A is *strictly positive* if

$$f \in C(X),\ f \ge 0,\ f \ne 0 \Longrightarrow Af(x) > 0, \quad \forall\, x \in X. \tag{1.4}$$

We say A is *primitive* if A is positive and some power A^m is strictly positive. We say A is *irreducible* if A is positive and

$$f \in C(X),\ f \ge 0,\ f \ne 0 \Longrightarrow \sup_k A^k f(x) > 0, \quad \forall\, x \in X. \tag{1.5}$$

The dual of $C(X)$ is

$$\mathcal{M}(X) = C(X)', \tag{1.6}$$

where $\mathcal{M}(X)$ denotes the space of finite, signed, regular Borel measures on X. The norm on $\mathcal{M}(X)$ is the total variation, which satisfies

$$\|\mu\|_{\mathrm{TV}} = \sup \{\langle f, \mu\rangle : f \in C(X),\ \|f\|_{\sup} \le 1\}, \tag{1.7}$$

where

$$\langle f, \mu\rangle = \int_X f\, d\mu. \tag{1.8}$$

The operator A in (1.2) has the adjoint

$$A^t : \mathcal{M}(X) \longrightarrow \mathcal{M}(X), \tag{1.9}$$

satisfying

$$\langle f, A^t \mu \rangle = \langle Af, \mu \rangle. \tag{1.10}$$

We have

$$\|A^t\|_1 = \|A\|_\infty, \tag{1.11}$$

where $\|A\|_\infty$ denotes the operator norm of A on $C(X)$ and $\|A^t\|_1$ that of A^t on $\mathcal{M}(X)$. Note that, if A is positive, then

$$A^t : \mathcal{M}_+(X) \longrightarrow \mathcal{M}_+(X), \tag{1.12}$$

where $\mathcal{M}_+(X)$ denotes the set of positive, finite, regular Borel measures on X.

A positive operator A on $C(X)$ is said to be a *stochastic* operator if, in addition,

$$A1 = 1. \tag{1.13}$$

For such operators, we have

$$A^t : \mathcal{P}(X) \longrightarrow \mathcal{P}(X), \tag{1.14}$$

where $\mathcal{P}(X)$ denotes the set of positive, regular Borel measures on X of total mass 1, i.e., probability measures on X.

The Perron-Frobenius theorem is a circle of results about the various sorts of operators defined above. The classical setting is the finite-dimensional case, i.e., where X is a finite point set. We establish such results here, in the infinite-dimensional setting. For a number of these results, we make the additional hypothesis that A in (1.2) is compact, which implies that A^t in (1.9) is compact.

In §2 we establish results in the Perron-Frobenius circle for stochastic operators. We first show that if A is a stochastic operator on $C(X)$, there exists $\mu \in \mathcal{P}(X)$ such that $A^t\mu = \mu$. (For this, compactness is not needed.) It follows that

$$A : V \longrightarrow V, \quad \text{where} \ \ V = \{f \in C(X) : \langle f, \mu \rangle = 0\}. \tag{1.15}$$

We then show that if A is stochastic and strictly positive, then, for $f \in C(X)$,

$$f \notin \operatorname{Span}(1) \Longrightarrow \|Af\|_{\sup} < \|f\|_{\sup}. \tag{1.16}$$

This is used to show in Propositions 2.3–2.4 that if A is a compact, stochastic operator on $C(X)$, and A is strictly positive, or more generally if A is primitive, then $A_V = A|_V$ has spectral radius $\rho(A_V) < 1$. Using this result, we establish the following in Proposition 2.5.

Proposition 1.1. *Let A be a compact, stochastic operator, and assume A is primitive. Then*

$$A^k \longrightarrow P, \quad \text{as } \ k \to \infty, \tag{1.17}$$

in operator norm, where P is the projection of $C(X)$ onto Span(1) *that annihilates V.*

It follows that

$$(A^t)^k \longrightarrow P^t \tag{1.18}$$

in operator norm, and P^t is the projection of $\mathcal{M}(X)$ onto Span(μ) that annihilates $W = \{\lambda \in \mathcal{M}(X) : \langle 1, \lambda \rangle = 0\}$. For P and P^t, we have the formulas

$$Pf = \langle f, \mu \rangle 1, \quad P^t\lambda = \langle 1, \lambda \rangle \mu, \tag{1.19}$$

given $f \in C(X)$, $\lambda \in \mathcal{M}(X)$. Making use of Proposition 1.1, we establish in Propositions 2.7–2.8 the following.

Proposition 1.2. *Let A be a compact, stochastic operator on $C(X)$, and assume A is irreducible. Then the measure $\mu \in \mathcal{P}(X)$ such that $A^t\mu = \mu$ is unique. Furthermore,* 1 *is an eigenvalue of A of algebraic multiplicity one, i.e., the generalized eigenspace $\mathcal{GE}(A, 1)$ is one-dimensional (equal to* Span(1)*).*

In §3 we turn to other classes of positive operators. We say a positive operator A on $C(X)$ is *crypto-stochastic* provided there exists

$$\psi \in C(X) \ \text{ such that } \ \psi(x) > 0, \ \forall\, x \in X, \quad \text{and} \quad A\psi = \psi. \tag{1.20}$$

Then, with $M_\psi f = \psi f$, $\widetilde{A} = M_\psi^{-1} A M_\psi$ is stochastic, and results of §2 apply. More generally, we say A is crypto-stochastic up to scaling if there exists $\lambda \in (0, \infty)$ such that $\lambda^{-1}A$ is crypto-stochastic. Clearly a necessary condition for A to have this property is that

$$A1(x) = \varphi(x) > 0, \quad \forall\, x \in X. \tag{1.21}$$

We show in Proposition 3.4 that if A is a positive, irreducible, compact operator on $C(X)$ satisfying three hypotheses, given as (H1)–(H3), then A is crypto-stochastic, up to scaling.

Turning away from crypto-stochastic operators, we establish the following in Proposition 3.5.

Proposition 1.3. *Let A be a positive, compact operator on $C(X)$. Assume that A satisfies (1.21). Then there exists $\lambda > 0$ and $\mu \in \mathcal{P}(X)$ such that $A^t\mu = \lambda\mu$.*

In §4 we take a look at positive infinite matrices that define bounded linear maps

$$A : \ell^\infty(\mathbb{N}) \longrightarrow \ell^\infty(\mathbb{N}). \tag{1.22}$$

Such maps are treated in [Sen]. We show that their analysis fits into the material developed in §§2–3, via the natural, positivity-preserving, isometric isomorphism

$$\ell^\infty(\mathbb{N}) \approx C(X_\mathbb{N}), \tag{1.23}$$

where $X_\mathbb{N}$ is the Stone-Cech compactification of $\mathbb{N}$, which can also be characterized as the maximal ideal space of $\ell^\infty(\mathbb{N})$, viewed as a commutative C^*-algebra.

2. Stochastic operators

Our first result in the circle of Perron-Frobenius theorems is the following. Actually, this result does not require A to be compact (nor does Proposition 2.2).

Proposition 2.1. *Assume A is a stochastic operator. Then there exists*

$$\mu \in \mathcal{P}(X) \quad \text{such that} \quad A^t\mu = \mu. \tag{2.1}$$

Proof. The set $\mathcal{P}(X)$ is a compact, convex subset of $\mathcal{M}(X)$, endowed with the weak* topology, and A^t is continuous on $\mathcal{M}(X)$ in this topology. Also,

$$A^t : \mathcal{P}(X) \longrightarrow \mathcal{P}(X). \tag{2.2}$$

The existence of a fixed point $\mu \in \mathcal{P}(X)$ is then a consequence of the Markov-Kakutani fixed point theorem (cf. [DS], p. 456).

Given μ as in (2.1), we set

$$V = \{f \in C(X) : \langle f, \mu\rangle = 0\}, \tag{2.3}$$

a closed linear subspace of $C(X)$, of codimension 1. We have a direct sum decomposition

$$C(X) = V \oplus \mathrm{Span}(1). \tag{2.4}$$

Also, whenever (2.1) holds,

$$A : V \longrightarrow V, \tag{2.5}$$

since

$$\langle Af, \mu\rangle = \langle f, A^t\mu\rangle = \langle f, \mu\rangle. \tag{2.6}$$

Proposition 2.2. *Let A be a stochastic operator. Assume in addition that A is strictly positive, so (1.4) holds. Then, for $f \in C(X)$,*

$$f \notin \mathrm{Span}(1) \Longrightarrow \|Af\|_{\sup} < \|f\|_{\sup}. \tag{2.7}$$

Proof. It suffices to treat the case $\|f\|_{\sup} = 1$, so $-1 \le f \le 1$. If $f(x) < 1$ for some x, there exists $\varphi \in C(X)$ such that $\varphi \ge 0$, $\varphi(x) > 0$, and $f + \varphi \le 1$. Hence $Af + A\varphi \le 1$. The hypothesis (1.4) implies $A\varphi(x) > 0$ for all $x \in X$, so $\sup Af(x) < 1$. Similarly, if $f(x) > -1$ for some $x \in X$, we have $\inf Af(x) > -1$. If $-1 \le f \le 1$ and $f \notin \mathrm{Span}(1)$, both of these conditions hold, and we have (2.7).

Before stating the next result, we note that if A is a stochastic operator on $C(X)$, then

$$\|A\|_\infty = \|A^t\|_1 = 1. \tag{2.8}$$

Also, having (2.5), let us denote the restriction of A to V by A_V.

Proposition 2.3. *Let A be a compact stochastic operator, and assume A is strictly positive. Then*

$$\alpha \in \operatorname{Spec} A_V \Longrightarrow |\alpha| < 1. \tag{2.9}$$

Hence the spectral radius of A_V is < 1, i.e.,

$$\rho(A_V) < 1. \tag{2.10}$$

Proof. Note that $A_V : V \to V$ is compact, so each nonzero $\alpha \in \operatorname{Spec}(A_V)$ must be an eigenvalue. The conclusion (2.9) then follows directly from (2.7). Also, compactness of A_V implies that $\operatorname{Spec} A_V$ is a countable subset of $\mathbb{C}$, whose only possible accumulation point is 0. Hence (2.10) follows from (2.9).

Remark. We recall the following useful formula for the spectral radius:

$$\rho(A_V) = \limsup_{k\to\infty} \|A_V^k\|^{1/k}. \tag{2.11}$$

The following result extends the scope of Proposition 2.3 a bit.

Proposition 2.4. *Let A be a compact stochastic operator, and assume A is primitive, i.e.,*

$$A^m \text{ is strictly positive for some } m \in \mathbb{N}. \tag{2.12}$$

Then the conclusions (2.9)–(2.10) hold.

Proof. We still have (2.5), and we can define A_V as before. Also $(A_V)^m = (A^m)_V$. Now if $\alpha \in \operatorname{Spec} A_V$, and $\alpha \neq 0$, compactness implies α is an eigenvalue of A_V, hence α^m is an eigenvalue of $(A_V)^m = (A^m)_V$. But Proposition 2.3 applies to A^m, so $|\alpha^m| < 1$. This gives (2.9), and (2.10) follows.

Remark. In case $X = \{1, 2\}$ so $C(X) = \mathbb{R}^2$, the following is an example of a stochastic matrix that is irreducible but not primitive:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.13}$$

Note that (2.9)–(2.10) fail for this matrix.

We can now prove the following key result.

Proposition 2.5. *Let A be a compact stochastic operator, and assume A is primitive. Then*

$$A^k \longrightarrow P, \quad \textit{as} \ \ k \to \infty, \tag{2.14}$$

in operator norm on $C(X)$, where P is the projection of $C(X)$ onto $\mathrm{Span}(1)$ *that annihilates V.*

Proof. We have

$$\begin{aligned} A^k &= A^k P + A^k(I-P) \\ &= P + A_V^k(I-P), \end{aligned} \tag{2.15}$$

so

$$\begin{aligned} \|A^k - P\|_\infty &= \|A_V^k(I-P)\|_\infty \\ &\le \|A_V^k\|_\infty \cdot \|I-P\|_\infty, \end{aligned} \tag{2.16}$$

and the fact that this converges to 0 (at an exponential rate) follows from (2.10)–(2.11).

Corollary 2.6. *In the setting of Proposition 2.5,*

$$(A^t)^k \longrightarrow P^t, \quad \textit{as} \ \ k \to \infty, \tag{2.17}$$

in operator norm on $\mathcal{M}(X)$. In this case, P^t is the projection of $\mathcal{M}(X)$ onto $\mathrm{Span}(\mu)$ *that annihilates*

$$W = \{\lambda \in \mathcal{M}(X) : \langle 1, \lambda \rangle = 0\}. \tag{2.18}$$

Proof. To get (2.17), just apply the transpose to (2.15)–(2.16):

$$(A^t)^k = P^t + (A_V^k(I-P))^t, \tag{2.19}$$

and note that

$$\|(A_V^k(I-P))^t\|_1 = \|A_V^k(I-P)\|_\infty. \tag{2.20}$$

Let us note that P is given by the formula

$$Pf = \langle f, \mu \rangle 1, \tag{2.21}$$

and then the identity

$$\langle f, P^t\lambda \rangle = \langle Pf, \lambda \rangle = \langle f, \mu \rangle \, \langle 1, \lambda \rangle \tag{2.22}$$

yields the formula

$$P^t\lambda = \langle 1, \lambda \rangle \mu, \quad \text{for} \ \ \lambda \in \mathcal{M}(X). \tag{2.23}$$

From (2.17) and (2.23), we deduce that, when A is a compact stochastic operator that is primitive, the measure μ in (2.1) is unique. The following extends the scope of this result.

Proposition 2.7. *Let A be a compact stochastic operator on $C(X)$, and assume A is irreducible. Then the measure μ in (2.1) is unique.*

Proof. Form

$$B = \sum_{k=1}^{\infty} 2^{-k} A^k = \frac{1}{2} A\Bigl(I - \frac{1}{2}A\Bigr)^{-1}, \tag{2.24}$$

which is a convergent series by (2.8), and defines a compact stochastic operator on $C(X)$. If A is irreducible, then B is strictly positive. Hence Proposition 2.3 and Corollary 2.6 apply to B. On the other hand, Proposition 2.1 applies to A, and clearly

$$\mu \in \mathcal{P}(X),\ A^t\mu = \mu \Longrightarrow B^t\mu = \mu. \tag{2.25}$$

By Corollary 2.6, applied to B, $(B^t)^k \to P^t$, given by (2.23). This establishes uniqueness of μ in (2.25).

Proposition 2.8. *Let A be a compact stochastic operator on $C(X)$, and assume A is irreducible. Then* 1 *is an eigenvalue of A of algebraic multiplicity* 1*, i.e., the generalized eigenspace $\mathcal{GE}(A,1)$ is* 1*-dimensional.*

Proof. With B as in (2.24), we have $B - I = (A - I)(I - A/2)^{-1}$, and hence

$$f \in \mathcal{GE}(A,1) \Longleftrightarrow f \in \mathcal{GE}(B,1). \tag{2.26}$$

But Proposition 2.3 applies to B, and the conclusion (2.9) for B_V implies $\mathcal{GE}(B,1)$ has dimension 1.

3. Other classes of positive compact operators

We move on from compact stochastic operators to other classes of positive compact operators on $C(X)$. To begin, we say a positive operator A on $C(X)$ is *crypto-stochastic* if there exists

$$\psi \in C(X) \text{ such that } \psi(x) > 0,\ \forall\, x \in X, \text{ and } A\psi = \psi. \tag{3.1}$$

Then, with $M_\psi f = \psi f$, we have the positive operator

$$\widetilde{A} = M_\psi^{-1} A M_\psi, \quad \text{stochastic}, \tag{3.2}$$

and the results of §2 apply to $\widetilde{A}$. Note that if A is strictly positive, resp., primitive, or irreducible, so is $\widetilde{A}$. Note also that strict positivity of ψ is required in order that M_ψ^{-1} be a well defined, bounded operator on $C(X)$. In connection with this, we have the following.

Proposition 3.1. *Assume the positive operator A is irreducible. Then*

$$\psi \in C(X),\ \psi \geq 0,\ \psi \neq 0,\ A\psi = \psi \Longrightarrow \psi(x) > 0, \quad \forall\, x \in X. \tag{3.3}$$

Proof. Let

$$E = \sum_{k=1}^{\infty} \frac{1}{k!} A^k = e^A - I. \tag{3.4}$$

If A is irreducible, then E is strictly positive. Now

$$A\psi = \psi \Longrightarrow E\psi = (e-1)\psi. \tag{3.5}$$

But $\psi \geq 0,\ \psi \neq 0 \Rightarrow E\psi(x) > 0$ for all $x \in X$, so we have (3.4).

Clearly a necessary condition for a positive operator A on $C(X)$ to be crypto-stochasitc is that

$$A1(x) = \varphi(x) > 0, \quad \forall\, x \in X. \tag{3.6}$$

However, this condition is not sufficient. For example, in one picks a positive $\lambda \neq 1$ and a strictly positive compact stochastic operator A_0 on $C(X)$, the operator $A = \lambda A_0$ is positive and satisfies (3.6), but (3.1) cannot hold. This motivates the definition of a more general class of operators. We say a positive operator A on $C(X)$ is crypto-stochastic up to scaling if there exist

$$\psi \in C(X),\ \lambda \in (0, \infty) \text{ such that } \psi(x) > 0,\ \forall\, x \in X \text{ and } A\psi = \lambda\psi. \tag{3.7}$$

In such a case, the operator $A^{\#} = \lambda^{-1}A$ is crypto-stochastic.

These considerations lead to the problem of determining when a positive, compact operator on $C(X)$ is crypto-stochastic, up to scaling. In connection with this, we mention the following weaker problem.

Problem PF. Given a positive, compact operator on $C(X)$, find

$$\psi \in C(X),\ \lambda > 0 \ \text{ such that } \ \psi \ge 0,\ \psi \ne 0,\ \text{and} \ A\psi = \lambda\psi. \tag{3.8}$$

The weakening consists in not requiring ψ in (3.8) to be strictly positive. Part of the classical Perron-Frobenius theory is that this problem is always solvable when X is a finite point set, so, for some $n \in \mathbb{N}$, $C(X) \approx \mathbb{R}^n$. Here is that result. We phrase its formulation and proof in a way that lends itself to extension beyond the finite case.

Proposition 3.2. *Assume X has n points, $n \in \mathbb{N}$, and A is a positive operator on $C(X)$. Assume*

$$f \in C(X),\ f \ge 0,\ f \ne 0 \Longrightarrow Af \ne 0. \tag{3.9}$$

Then there exist $\lambda > 0$ and $\psi \in C(X)$ satisfying (3.8).

Proof. Let ν_0 be the probability measure on X that assigns the mass $1/n$ to each of its points. With the notation

$$C_+(X) = \{f \in C(X) : f \ge 0\}, \tag{3.10}$$

let

$$\Sigma = \{f \in C_+(X) : \langle f, \nu_0\rangle = 1\}. \tag{3.11}$$

Thus Σ is a compact, convex subset of $C(X)$. We define

$$\Phi : \Sigma \longrightarrow \Sigma \tag{3.12}$$

by

$$\Phi(f) = \frac{1}{\langle Af, \nu_0\rangle} Af. \tag{3.13}$$

The hypothesis (3.11) implies $\langle Af, \nu_0\rangle > 0$ for all $f \in \Sigma$, and by compactness we have a positive lower bound. Now the Brouwer fixed point theorem applies to (3.13). (A proof of his result can be found in Chapter 1 of [T].) Hence there exists $f \in \Sigma$ such that

$$Af = \langle Af, \nu_0\rangle f. \tag{3.14}$$

This proves Proposition 3.2.

Recalling Proposition 3.1, we see that if X is a finite point set, every positive, irreducible A on $C(X)$ is crypto-stochastic, up to scaling.

We return to cases where $C(X)$ is infinite dimensional, and investigate ways to extend the proof of Proposition 3.2 to cover positive, compact operators on $C(X)$, under some additional hypotheses. To start, we make the following three hypotheses:

(H1) There is a measure $\nu \in \mathcal{P}(X)$ such that $\nu(U) > 0$ for each nonempty open $U \subset X$. Equivalently,

$$f \in C(X),\ f \geq 0,\ f \neq 0 \Longrightarrow \langle f, \nu \rangle > 0. \tag{3.15}$$

(H2) The positive operator A satisfies

$$A : L^1(X, \nu) \longrightarrow C(X), \quad \text{compactly.} \tag{3.16}$$

(H3) With $C_+(X)$ as in (3.10), and

$$\Sigma = \{f \in C_+(X) : \langle f, \nu \rangle = 1\}, \tag{3.17}$$

there is a $\delta > 0$ such that

$$f \in \Sigma \Longrightarrow \|Af\|_{\sup} \geq \delta. \tag{3.18}$$

These hypotheses imply that $A(\Sigma)$ is a relatively compact, convex subset of $C(X)$. The following is a useful improvement of (3.18).

Lemma 3.3. *Under hypotheses (H1)–(H3), there exists* $\alpha > 0$ *such that*

$$f \in \Sigma \Longrightarrow \langle Af, \nu \rangle \geq \alpha. \tag{3.19}$$

Proof. If (3.19) fails, there exist $f_k \in \Sigma$ such that $\langle Af_k, \nu \rangle \leq 2^{-k}$. Since $A(\Sigma)$ is relatively compact in $C(X)$, we have a subsequence f_{k_j} such that $Af_{k_j} \to g \in C_+(X)$, uniformly. Consequently, $\langle g, \nu \rangle = 0$, which by (H1), implies $g = 0$. This contradicts the condition (3.18) in (H3).

Now define

$$\Phi : \Sigma \longrightarrow \Sigma, \quad \Phi(f) = \frac{1}{\langle Af, \nu \rangle} Af. \tag{3.20}$$

By (3.19), this is a well defined, continuous map, and the relative compactness of $A(\Sigma)$ in $C(X)$ yields

$$\Phi : \Sigma \longrightarrow \mathcal{K}, \tag{3.21}$$

where $\mathcal{K}$ is a compact, convex subset of $\Sigma \subset C(X)$. The Schauder fixed point theorem (a proof of which can be found in Chapter 13 of [T]) applies, to yield $\psi \in \mathcal{K} \subset \Sigma$ satisfying $\Phi(\psi) = \psi$, hence $A\psi = \langle A\psi, \nu \rangle \psi$. We have proved the first part of the following.

Proposition 3.4. *Let $A : C(X) \to C(X)$ be a positive operator. Assume hypotheses (H1)–(H3). Then there exist $\lambda > 0$ and $\psi \in C(X)$ such that (3.8) holds.*

If also A is irreducible, then $A^\# = \lambda^{-1} A$ satisfies

$$A^\# \psi = \psi, \quad \text{and} \quad \psi(x) > 0, \quad \forall\, x \in X, \tag{3.22}$$

and hence $A^\#$ is crypto-stochastic.

Proof. The first part was established above, and (3.22) follows from Proposition 3.1.

Suppose now that A_0 is a compact, positive operator on $C(X)$ and that (H1)–(H3) hold for $A = A_0^m$, for some $m \in \mathbb{N}$. If A_0 is irreducible, so is A, so, with λ as in Proposition 3.4, $A_1^m = A^\#$ is crypto-stochastic, where $A_1 = \lambda^{-1/m} A_0$, and we have (3.22). It follows that there exists $\mu \in \mathcal{P}(X)$ such that $(A^\#)^t \mu = \mu$, and, with V as in (2.3), ψ as in (3.22), we have

$$C(X) = V \oplus \operatorname{Span}(\psi), \quad A^\# : V \to V. \tag{3.23}$$

If $A^\#$ is primitive, so is $A^\#$. One deduces via Proposition 2.4 that $A^\#_V = A^\#|_V$ has spectral radius $\rho < 1$, and

$$\mathcal{GE}(A^\#, 1) = \operatorname{Span}(\psi). \tag{3.24}$$

Note also that

$$A^\#(A_1\psi) = A_1^{m+1}\psi = A_1(A^\#\psi) = A_1\psi, \tag{3.25}$$

hence $A_1\psi \in \operatorname{Span}(\psi)$. If $A_1\psi = \beta\psi$, then $A^\#\psi = \beta^m\psi = \psi$, so $\beta^m = 1$. Since A_1 is positive, this implies $\beta = 1$, so

$$A_1\psi = \psi. \tag{3.26}$$

Consequently A_1 itself is crypto-stochastic.

We temporarily leave results related to (H1)–(H3), and look directly for positive measures on X that are eigenvectors of A^t.

Proposition 3.5. *Let A be a positive, compact operator on $C(X)$. Assume that $\varphi = A1$ satisfies (3.6). Then there exist $\lambda > 0$ and $\mu \in \mathcal{P}(X)$ such that*

$$A^t \mu = \lambda \mu. \tag{3.27}$$

Proof. First note that there exists $\delta > 0$ such that

$$\langle 1, A^t\mu\rangle = \langle \varphi, \mu\rangle \geq \delta, \quad \forall\, \mu \in \mathcal{P}(X), \tag{3.28}$$

given (3.6). Hence we can define

$$\Psi : \mathcal{P}(X) \longrightarrow \mathcal{P}(X), \quad \Psi(\mu) = \frac{1}{\langle 1, A^t\mu\rangle} A^t\mu, \tag{3.29}$$

and Ψ is continuous. Since $A^t(\mathcal{P}(X))$ is a relatively compact, convex subset of $\mathcal{M}_+(X)$, we have

$$\Psi : \mathcal{P}(X) \longrightarrow \mathcal{K}, \tag{3.30}$$

where $\mathcal{K}$ is a compact, convex subset of $\mathcal{P}(X)$. It follows from the Schauder fixed point theorem that Ψ has a fixed point, say μ, in $\mathcal{K}$, and then

$$A^t\mu = \langle 1, A^t\mu\rangle\mu, \tag{3.31}$$

giving (3.27).

Having Proposition 3.5, we again make contact with (H1):

Proposition 3.6. *Let A be a opsitive operator on $C(X)$. If A is irreducible and $\mu \in \mathcal{P}(X)$ satisfies (3.27), with $\lambda > 0$, then*

$$f \in C(X),\ f \geq 0,\ f \neq 0 \Longrightarrow \langle f, \mu\rangle > 0. \tag{3.32}$$

Proof. For each $k \in \mathbb{N}$,

$$\lambda^k\langle f, \mu\rangle = \langle f, (A^t)^k\mu\rangle = \langle A^k f, \mu\rangle, \tag{3.33}$$

hence, for $E = e^A - I$ as in (3.4) and f as in (3.32),

$$(e-1)\langle f, \mu\rangle = \langle Ef, \mu\rangle > 0, \tag{3.34}$$

since irreducibility of A implies $Ef(x) > 0$ for all $x \in X$.

Remark. Compactness of A is not required for Proposition 3.6. This fact is particularly significant in light of Proposition 2.1.

It follows from Propositon 3.6 that, in the setting of Proposition 3.5, and with A irreducible, hypothesis (H1) holds with $\nu = \mu$. Furthermore, with

$$\Sigma = \{f \in C_+(X) : \langle f, \mu\rangle = 1\}, \tag{3.35}$$

we have

$$\lambda^{-1}A : \Sigma \longrightarrow \Sigma, \tag{3.36}$$

so also (H3) and (3.19) hold. Thus Proposition 3.4 implies the following.

Proposition 3.7. *Let A be a positive, compact, irreducible operator on $C(X)$, and assume $\varphi = A1$ satisfies (3.6). Take $\mu \in \mathcal{P}(X)$ such that (3.27) holds. Assume that (H2) holds with $\nu = \mu$, i.e.,*

$$A : L^1(X, \mu) \longrightarrow C(X), \quad \textit{compactly}. \tag{3.37}$$

Then $\lambda^{-1}A$ is crypto-stochastic.

4. Connections with infinite positive matrices

Here we look at infinite matrices $A = (a_{jk})$, defined for $j, k \in \mathbb{N}$, having a bound on the row sums:

$$\sum_{k=1}^{\infty} |a_{jk}| \le \alpha < \infty, \quad \forall\, j \in \mathbb{N}. \tag{4.1}$$

Then we have

$$A : \ell^\infty(\mathbb{N}) \longrightarrow \ell^\infty(\mathbb{N}), \tag{4.2}$$

acting as a bounded operator, by

$$(Af)_j = \sum_{k=1}^{\infty} a_{jk} f_k, \quad \|A\|_\infty \le \alpha. \tag{4.3}$$

Here, $\ell^\infty(\mathbb{N})$ denotes the space of bounded real sequences, i.e., the space of bounded functions $f : \mathbb{N} \to \mathbb{R}$, a Banach space with norm $\|f\|_\infty = \sup_k |f_k|$. We identify $f_k = f(k)$. We say A is positive if $a_{jk} \ge 0$ for each $j, k \in \mathbb{N}$. We say a positive matrix is stochastic if each row sum is 1, i.e., $\sum_k a_{jk} = 1$ for each j, or equivalently

$$A1 = 1, \tag{4.4}$$

where here 1 denotes the function on $\mathbb{N}$ that is identically 1.

To relate the study of such matrices to material in §§1–3, we use the natural isometric isomorphism

$$\ell^\infty(\mathbb{N}) \approx C(X_{\mathbb{N}}), \tag{4.5}$$

where $X_{\mathbb{N}}$ denotes the Stone-Cech compactification of $\mathbb{N}$. This is a compact Hausdorff space. There is a natural inclusion

$$\mathbb{N} \subset X_{\mathbb{N}}, \tag{4.6}$$

as an open, dense subset. Since $\ell^\infty(\mathbb{N})$ is not separable, $X_{\mathbb{N}}$ is not metrizable. Using the isomorphism (4.5), we identify A in (4.2) with

$$A : C(X_{\mathbb{N}}) \longrightarrow C(X_{\mathbb{N}}). \tag{4.7}$$

Positivity in (4.2) turns into positivity in (4.7), since the isomorphism (4.5) is also positivity preserving. Also, if (4.4) holds on $\ell^\infty(\mathbb{N})$, it holds on $C(X_{\mathbb{N}})$.

As in (1.6), we have the duality

$$\ell^\infty(\mathbb{N})' \approx C(X_{\mathbb{N}})' = \mathcal{M}(X_{\mathbb{N}}), \tag{4.8}$$

where $\mathcal{M}(X_{\mathbb{N}})$ is the space of finite, signed, regular Borel measures on $X_{\mathbb{N}}$. There are a natural injection and a natural projection

$$J : \ell^1(\mathbb{N}) \longrightarrow \mathcal{M}(X_{\mathbb{N}}), \quad \Pi : \mathcal{M}(X_{\mathbb{N}}) \longrightarrow \ell^1(\mathbb{N}), \tag{4.9}$$

induced by (4.6).

As in (1.9), the map (4.7) has a transpose

$$A^t : \mathcal{M}(X_{\mathbb{N}}) \longrightarrow \mathcal{M}(X_{\mathbb{N}}). \tag{4.10}$$

We also have a map

$$A^\tau : \ell^1(\mathbb{N}) \longrightarrow \ell^1(\mathbb{N}), \tag{4.11}$$

given by

$$A^\tau = \Pi A^t J. \tag{4.12}$$

Results of §§2–3 yield conditions under which A^t has an invariant measure $\mu \in \mathcal{P}(X_{\mathbb{N}})$. Such μ is also an invariant element of A^τ if and only if $\operatorname{supp} \mu \subset \mathbb{N}$. We will see examples below for which $\ell^1(\mathbb{N})$ does not have a positive element that is invariant under A^τ.

While A^t has a richer structure that A^τ, the operator A^τ does not lose information about A. In fact, since $\ell^\infty(\mathbb{N}) = \ell^1(\mathbb{N})'$, a bounded operator on $\ell^1(\mathbb{N})$ has a transpose:

$$B : \ell^1(\mathbb{N}) \to \ell^1(\mathbb{N}) \Longrightarrow B^t : \ell^\infty(\mathbb{N}) \to \ell^\infty(\mathbb{N}), \tag{4.13}$$

and we see that, for A in (4.2),

$$(A^\tau)^t = A. \tag{4.14}$$

Note also that $\ell^1(\mathbb{N})$ is weak* dense in $\mathcal{M}(X_{\mathbb{N}})$ (though not norm dense), and A^t is the unique extension of A^τ to a linear operator on $\mathcal{M}(X_{\mathbb{N}})$ that is continuous in the weak* topology of $\mathcal{M}(X_{\mathbb{N}})$.

At this point it is useful to note that, in place of the set of positive integers $\mathbb{N}$, we could use any countably infinite set S, and extend results derived above from the setting of (4.2) to

$$A : \ell^\infty(S) \to \ell^\infty(S), \quad \ell^\infty(S) \approx C(X_S), \quad C(X_S)' = \mathcal{M}(X_S), \tag{4.15}$$

where X_S is the Stone-Cech compactification of X_S. Particularly useful examples include

$$\mathbb{Z}, \quad \mathbb{Z}^n, \quad S\ell(n,\mathbb{Z}). \tag{4.16}$$

An example involving $S = \mathbb{Z}$ is

$$A : \ell^\infty(\mathbb{Z}) \longrightarrow \ell^\infty(\mathbb{Z}), \quad Af(k) = f(k+1). \tag{4.17}$$

An element $\mu \in \mathcal{P}(X_{\mathbb{Z}})$ invariant under A^t defines a linear functional on $\ell^\infty(\mathbb{Z})$ known as an *invariant mean.* The existence of such invariant means is a special case of Proposition 2.1. It is clear that such μ satisfies $\mu(\mathbb{Z}) = 0$, and there does not exist an A^τ-invariant element of $\ell^1(\mathbb{Z})$. Such results hold for a number of related operators, such as

$$A : \ell^\infty(\mathbb{Z}) \longrightarrow \ell^\infty(\mathbb{Z}), \quad Af(k) = \frac{1}{2} f(k+1) + \frac{1}{2} f(k-1). \tag{4.18}$$

The Markov process associated to this operator is "the standard random walk" on $\mathbb{Z}$.

In many cases, the operator A might have special structure that allows one to replace $\ell^\infty(S)$ by a smaller subspace, such as

$$\begin{aligned} \ell^\infty_\#(\mathbb{N}) &= \{f \in \ell^\infty(\mathbb{N}) : \lim_{k\to\infty} f(k) \text{ exists}\}, \quad \text{or} \\ \ell^\infty_\#(\mathbb{Z}) &= \{f \in \ell^\infty(\mathbb{Z}) : \lim_{k\to+\infty} f(k) \text{ and } \lim_{k\to-\infty} f(k) \text{ exist}\}, \end{aligned} \tag{4.19}$$

having natural isomorphisms

$$\begin{aligned} \ell^\infty_\#(\mathbb{N}) &\approx C(\widehat{\mathbb{N}}), \quad \widehat{\mathbb{N}} = \mathbb{N} \cup \{\infty\}, \\ \ell^\infty_\#(\mathbb{Z}) &\approx C(\widehat{\mathbb{Z}}), \quad \widehat{\mathbb{Z}} = \mathbb{Z} \cup \{+\infty, -\infty\}. \end{aligned} \tag{4.20}$$

For example, (4.18) has the variant

$$A : \ell^\infty_\#(\mathbb{Z}) \longrightarrow \ell^\infty_\#(\mathbb{Z}), \quad Af(k) = \frac{1}{2} f(k+1) + \frac{1}{2} f(k-1). \tag{4.21}$$

Using the same letter A in (4.18) and (4.21) is perhaps an abuse of notation, a practice we continue by writing

$$A^t : \mathcal{M}(\widehat{\mathbb{Z}}) \longrightarrow \mathcal{M}(\widehat{\mathbb{Z}}). \tag{4.22}$$

In the setting of (4.21), we can write the set of elements of $\mathcal{P}(\widehat{\mathbb{Z}})$ that are invariant under A^t as

$$\{a\delta_{+\infty} + (1-a)\delta_{-\infty} : 0 \le a \le 1\}. \tag{4.23}$$

While Proposition 2.1 applies to the operators in (4.17), (4.18), and (4.21), most of the rest of §2 does not. For one thing, the operators just mentioned are not compact. Also, the operators in (4.18) and (4.21) have the *weak* irreducibility property that

$$f \in \ell^\infty(\mathbb{Z}),\ f \ge 0,\ f \neq 0 \Longrightarrow \sup_m A^m f(k) > 0, \quad \forall\, k \in \mathbb{Z}. \tag{4.24}$$

However, whenever $f(k) \to 0$ as $|k| \to \infty$, so does $A^m f(k)$ for each m, so $\sup_m A^m f$ vanishes on $X_{\mathbb{Z}} \setminus \mathbb{Z}$ in case (4.18), and on $\widehat{\mathbb{Z}} \setminus \mathbb{Z}$ in case (4.21), for each such f, violating the condition (1.5) for irreducibility. This also explains why Proposition 3.6 does not apply to A in (4.18) and (4.21). For a positive operator $A : \ell^\infty(S) \to \ell^\infty(S)$, irreducibility in the sense of (1.5), for $A : C(X_S) \to C(X_S)$ (which, for emphasis, we will call *strong* irreducibility) is equivalent to the property

$$f \in \ell^\infty(S),\ f \ge 0,\ f \neq 0 \Longrightarrow \sup_m A^m f(k) \ge \delta(f) > 0, \quad \forall\, k \in S. \tag{4.25}$$

We next look at a family of operators on $\ell^\infty(\mathbb{N})$, examined in [Sen], given by the infinite matrices

$$A = \begin{pmatrix} q_1 & p_1 & & & \cdots \\ q_2 & 0 & p_2 & & \\ q_3 & 0 & 0 & p_3 & \\ \vdots & & & & \ddots \end{pmatrix}, \tag{4.26}$$

with

$$0 < q_j < 1, \quad q_j + p_j = 1. \tag{4.27}$$

Note that

$$\{q_j\} \text{ bounded away from } 0 \Longrightarrow A \text{ is strongly irreducible,} \tag{4.28}$$

as defined above, since then (4.25) holds. One also readily verifies that

$$p_j \to 0 \ \text{ as } \ j \to \infty \Longrightarrow A \ \text{ is compact on } \ \ell^\infty(\mathbb{N}). \tag{4.29}$$

One can further extend the scope of the approach to positive infinite matrices described above. For example, we can consider bounded linear operators

$$A : L^\infty(Y, \sigma) \longrightarrow L^\infty(Y, \sigma), \tag{4.30}$$

where (Y, σ) is a measure space. The positivity condition becomes

$$f \in L^\infty(Y, \sigma), \ f \geq 0 \Longrightarrow Af \geq 0. \tag{4.31}$$

This fits into the framework of §§2–3 as follows. The space $L^\infty(Y, \sigma)$ is a commutative C^*-algebra, and we have a natural, positivity-preserving, isometric isomorphism

$$L^\infty(Y, \sigma) \approx C(X), \tag{4.32}$$

where X is the maximal ideal space of the C^*-algebra $L^\infty(Y, \sigma)$. In (4.2), we have $Y = \mathbb{N}$ and $\sigma =$ counting measure.

References

[DS] N. Dunford and J. Schwartz, Linear Operators, Vol. 1, Wiley, New York, 1958.

[Nel] E. Nelson, The adjoint Markoff process, Duke Math. J. 25 (1958), 671–690.

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

[T] M. Taylor, Partial Differential Equations, Vols. 1–3, Springer-Verlag, New York, 1996 (2nd ed., 2011).