KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIWERSYTET GDAŃSKI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

# Markov operators and their stability - a model way of solving exercises

Tomasz Szarek

**PWP Interdysciplinary Doctoral Studies in Mathematical Modeling**
**UDA-POKL.04.01.01-00-026/13-00**
**Projekt jest współfinansowany przez Unię Europejską w ramach Europejskiego Funduszu Społecznego**

**Exercise 1:** Consider the system

$$x_{n+1} = T(x_n, \xi_n) \qquad \text{for } n = 0, 1, \dots, \tag{1}$$

where $T$ is a given transformation and the $\xi_n$ are independent random vectors satisfying:

- $T$ is defined on the space $X \times W$ with values in $X$, where $W$ is a closed subset of $\mathbb{R}$. For every fixed $y \in W$ the function $T(x, y)$ is continuous in $x$ and for every fixed $x \in X$ it is measurable in $y$.
- The random variables $\xi_0$, $\xi_1$, $\dots$, have values in $W$ and have te same distribution, that is, the measure

  $$\nu(B) = \text{prob}(\xi_n \in B) \qquad \text{for } B \in \mathcal{B}(W)$$

  is the same for all $n$.
- The initial random vector $x_0$ has values in $X$ and the vectors $x_0, \xi_0, \xi_1, \dots,$ are independent.

Find the transition operator (Markov operator) for the chain $(x_n)_{n\geq 1}$.

**Solution:** Observe that if $\xi_0, \xi_1, \dots$ are independent and $x_0$ is constant, then the vectors $x_0, \xi_0, \xi_1, \dots$, are also independent. the measures According to (1) the random vector $x_n$ is a function of $x_0$ and $\xi_0, \xi_1, \dots, \xi_{n-1}$. Hence $x_n$ and $\xi_n$ are independent. Using this we wil derive a reccurence formula for the measures

$$\mu_n(A) = \text{prob}(x_n \in A), \qquad A \in \mathcal{B}(X),$$

which statistically describe the behaviour of the dynamical system (1).

Thus, choose a bounded Borel measurable function $h : X \to \mathbb{R}$ and for some integer $n \geq 0$ consider the random variable $z_{n+1} = h(x_{n+1})$. Observe that

$$\mu_{n+1}(A) = \text{prob}(x_{n+1}^{-1}(B)).$$

Using this the mathematical expectation $E(z_{n+1})$ can be calculated as follows:

$$\begin{aligned} E(z_{n+1}) &= \int_\Omega h(x_{n+1}(\omega)) \,\text{prob}(d\omega) = \int_X h(x) \,\text{prob}(x_{n+1}^{-1}(dx)) \\ &= \int_X h(x)\mu_{n+1}(dx). \end{aligned}$$

For a Borel measurable function $f$ and a Borel measure $\mu$ we will write

$$\langle f, \mu \rangle = \int_X f(x)\mu(dx).$$

However, since $z_{n+1} = h(T(x_n, \xi_n))$ we have

$$
\begin{aligned}
E(z_{n+1}) &= \int_\Omega h(T(X_n(\omega), \xi_n(\omega)))\,\mathrm{prob}(d\omega) \\
&= \iint_{X\times W} h(T(x,y))\,\mathrm{prob}((x_n,\xi_n)^{-1}(dx,dy)).
\end{aligned} \tag{2}
$$

The independence of the random vectors $x_n$ and $\xi_n$ implies that

$$
\mathrm{prob}((x_n,\xi_n)^{-1}(A\times B)) = \mathrm{prob}(x_n^{-1}(A))\,\mathrm{prob}(\xi_n^{-1}(B)),
$$

which shows that the measure $\mathrm{prob}((x_n,\xi_n)^{-1}(C))$ is the product of measures

$$
\mu_n(A) = \mathrm{prob}(x_n^{-1}(A)) \quad \text{and} \quad \nu(B) = \mathrm{prob}(\xi_n^{-1}(B)).
$$

Thus, by the Fubini theorem we obtain

$$
E(z_{n+1}) = \int_X \left\{ \int_W h(T(x,y))\nu(dy) \right\} \mu_n(dx).
$$

Equating this expression with (2) we immediately obtain

$$
\langle h, \mu_{n+1} \rangle = \int_X \left\{ \int_W h(T(x,y))\nu(dy) \right\} \mu_n(dx).
$$

Using this we may calculate the value of $\mu_{n+1}(A)$ for an arbitrary measurable set $A \subset X$. Namely, setting $h = \mathbf{1_A}$ we obtain

$$
\mu_{n+1}(A) = \int_X \left\{ \int_W \mathbf{1_A}(\mathbf{T}(\mathbf{x},\mathbf{y}))\nu(\mathbf{dy}) \right\} \mu_n(dx).
$$

Thus the operator $P : \mathcal{M}_{fin} \to \mathcal{M}_{fin}$ describing the evolution of measures is given by

$$
P\mu(A) = \int_X \left\{ \int_W \mathbf{1_A}(\mathbf{T}(\mathbf{x},\mathbf{y}))\nu(\mathbf{dy}) \right\} \mu(dx) \quad \text{for } \mu \in \mathcal{M}_{fin},\ A \in \mathcal{B}(X). \tag{3}
$$

Since $\nu$ is a probabilistic measure, it is obvious that $P$ is a Markov operator.

Setting

$$
Uh(x) = \int_W h(T(x,y))\nu(dy) \qquad \text{for } x \in X,
$$

we may rewrite (3) in the form

$$
\langle \mathbf{1}_A, P\mu \rangle = \langle U\mathbf{1}_A, \mu \rangle.
$$

Due to the linearity of the scalar product this implies

$$
\langle g_n, P\mu \rangle = \langle U g_n, \mu \rangle,
$$

where

$$g_n = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{\mathbf{A_i}}$$

is a simple function. Further, since every measurable function $h$ can be approximated by a sequence $\{g_n\}$ of simple functions, we obtain in the limit

$$\langle h, P\mu \rangle = \langle Uh, \mu \rangle, \tag{4}$$

if $\{g_n\}$ and $\{Ug_n\}$ satisfy the conditions of the Lebesgue dominated or Lebesgue monotone convergence theorem. In particular, (4) is valid if $h$ is Borel measurable and bounded or nonnegative. $\square$

**Exercise 2:** Show that the space $\mathcal{M}_{sig} := \mathcal{M} - \mathcal{M}$ with the Wasserstein distance is not complete.

**Solution:** Introduce in the space of Borel measures the total variation norm. Let $\nu \in \mathcal{M}_{sig}$ and let $(X_1, \ldots, X_n)$ be a measurable partition of $X$, that is

$$X = \bigcup_{i=1}^{n} X_i, \quad X_i \cap X_j = \emptyset \quad \text{for } i \neq j,\ X_i \in \mathcal{B}(X).$$

We set

$$\|\nu\|_0 = \sup \sum_{i=1}^{n} |\nu(X_i)|,$$

where the supremum is taken over all possible measurale partitions of $X$ (with arbitrary $n$). The value $\|\nu\|_0$ is the desired distance. Further observe that if $x \neq y$ then $\|\delta_x - \delta_y\|_0 = 2$. It is easy to check that $(\mathcal{M}_{sig}, \|\cdot\|_0)$ is a Banach space.

Let $\nu \in \mathcal{M}_{sig}$ and let $(X_1, \ldots, X_n)$ be a measurable partition of $X$. Fix $f \in C(X)$ such that $\|f\|_\infty \le 1$ and $\operatorname{Lip} f \le 1$. We have

$$|\langle f, \nu \rangle| = |\int_X f d\nu| \le \sum_{i=1}^{n} |\int_{X_i} f d\nu| \le \sum_{i=1}^{n} |\nu(X_i)|$$

and taking supremum over partitions $(X_1, \ldots, X_n)$ on the right hand side and supremum over functions $f$ satisfying $\|f\|_\infty \le 1$ and $\operatorname{Lip} f \le 1$ on the left hand side, we obtain

$$\|\nu\|_W \le \|\nu\|_0. \tag{5}$$

Assume now, contrary to our claim, that $(\mathcal{M}_{sig}, \|\cdot\|_W)$ is a Banach space similarly to $(\mathcal{M}_{sig}, \|\cdot\|_0)$. By (5) the linear operator $T = \text{id}$ is continuous and from the Banach open mapping theorem it follows that there exists $\gamma > 0$ such that

$$\|\nu\|_0 \le \gamma \|\nu\|_W.$$

However taking an arbitrary sequence $(x_n)_{n\geq 1}$ with $x_n \in X$ convergent to $x \in X$ and such that $x_n \neq x$ we see

$$2 = \|\delta_x - \delta_{x_n}\|_0 \leq \gamma \|\delta_x - \delta_{x_n}\|_W \leq \rho(x_n, x) \to 0 \quad \text{as } n \to \infty,$$

which contradicts the assumption that $(\mathcal{M}_{sig}, \|\cdot\|_W)$ is a Banach space. □

**Exercise 3:** Show that if $\|\mu_1 - \mu_2\|_W \leq \varepsilon^2$ for $\mu_1, \mu_2 \in \mathcal{M}_1$ and some $\varepsilon \in (0,1)$, then

$$\mu_1(\mathcal{N}(C,\varepsilon)) \geq \mu_2(C) - \varepsilon \qquad \text{for every Borel set } C \subset X.$$

(Here $\mathcal{N}(C,\varepsilon) = \{x \in X : \inf\{\rho(x,y) : y \in C\} < \varepsilon\}$.)

**Solution:** Fix $\varepsilon \in (0,1)$. Let $\mu_1, \mu_2 \in \mathcal{M}_1$ be given. Fix a Borel set $C \subset X$. Define $f(x) = \max(\varepsilon - \rho(C,x), 0)$. It is easy to chceck that $\|f\|_\infty \leq 1$ and $\operatorname{Lip} f \leq 1$. Since $f(x) = 0$ for $x \notin \mathcal{N}(C,\varepsilon)$ and $f(x) = \varepsilon$ for $x \in C$, we have

$$\begin{aligned}\varepsilon\mu_2(C) - \varepsilon\mu_1(\mathcal{N}(C,\varepsilon)) &\leq | < f, \mu_1 > - < f, \mu_2 > | \\ &\leq \|\mu_1 - \mu_2\|_W \leq \varepsilon^2.\end{aligned}$$

Therefore,

$$\mu_1(\mathcal{N}(C,\varepsilon)) \geq \mu_2(C) - \varepsilon$$

and the assertion follows. □

**Exercise 4:** Let $P$ be a nonexpensive Markov operator. Assume that for every $\varepsilon > 0$ there is a Borel set $A$ with $\operatorname{diam} A \leq \varepsilon$, a real number $\alpha > 0$ and an integer $\bar{n}$ such that

$$P^{\bar{n}}\mu(A) \geq \alpha \qquad \text{for } \mu \in \mathcal{M}_1. \tag{6}$$

Prove that $P$ is asymptotically stable.

**Solution:** Fix $\mu_1, \mu_2 \in \mathcal{M}_1$ and an $\varepsilon > 0$. Choose $A \subset X$ and $\alpha$, $0 < \alpha < 1$ according to (6).

We define four sequences of probability measures $(\mu_i^k)$, $(\nu_i^k)$ $\;, k = 0, 1, 2, \ldots$, $i = 1, 2$. If $k = 0$ we define $\mu_i^0 = \nu_i^0 = \mu_i$. If $k \geq 1$ is fixed and $\mu_i^{k-1}$, $\nu_i^{k-1}$ are given, we define

$$\nu_i^k(B) = \frac{P^{\bar{n}}\mu_i^{k-1}(B \cap A)}{P^{\bar{n}}\mu_i^{k-1}(A)} \quad \mu_i^k(B) = \frac{1}{1-\alpha}\{P^{\bar{n}}\mu_i^{k-1}(B) - \alpha\nu_i^k(B)\}. \tag{7}$$

Since $P^{\bar{n}}\mu_i^{k-1}(A) \geq \alpha$, we have

$$P^{\bar{n}}\mu_i^{k-1}(B) \geq P^{\bar{n}}\mu_i^{k-1}(B \cap A) = P^{\bar{n}}\mu_i^{k-1}(A)\nu_i^k(B) \;\; \geq \alpha\nu_i^k(B).$$

Observe that $\nu_i^k(X \setminus A) = 0$ and consequently

$$\|\nu_1^k - \nu_2^k\|_W = \sup_f |\int_X f d\nu_1^k - \int_X f d\nu_2^k| \;\; = \sup_f |\int_A f d\nu_1^k - \int_A f d\nu_2^k| \leq \operatorname{diam} A \leq \varepsilon. \tag{8}$$

Using equations (7) we have

$$P^{k\cdot\bar n}\mu_i = \alpha P^{(k-1)\cdot\bar n}\nu_i^1 + \alpha(1-\alpha)P^{(k-2)\cdot\bar n}\nu_i^2 \\ +\ldots+\alpha(1-\alpha)^{k-1}\nu_i^k + (1-\alpha)^k\mu_i^k \qquad \text{for } k>1.$$

Since $P$ is nonexpansive, this implies

$$\|P^{k\cdot\bar n}\mu_1 - P^{k\cdot\bar n}\mu_2\|_W \le \alpha\|\nu_1^1-\nu_2^1\|_W + \alpha(1-\alpha)\|\nu_1^2-\nu_2^2\|_W \\ +\ldots+\alpha(1-\alpha)^{k-1}\|\nu_1^k-\nu_2^k\|_W + (1-\alpha)^k\|\mu_1^k-\mu_2^k\|_W.$$

From above, condition (8) and the obvious inequality $\|\mu_1^k-\mu_2^k\|_W \le 2$ it follows that

$$\|P^{k\cdot\bar n}(\mu_1-\mu_2)\|_W \le \varepsilon + 2(1-\alpha)^k.$$

Using the nonexpansivness of $P$ we obtain

$$\|P^n(\mu_1-\mu_2)\|_W \le \varepsilon + 2(1-\alpha)^k \qquad \text{for } n \ge k\cdot\bar n.$$

Since $\varepsilon>0$ was arbitrary and $k$ does not depend on $\mu_1$, $\mu_2$ we have

$$\|P^n\mu_1 - P^n\mu_2\|_W \le \varepsilon$$

for $n\ge n_0$ and every measures $\mu_1,\mu_2\in\mathcal{M}_1$.

So, we are given

$$\|P^n\mu - P^m\mu\|_W \le \varepsilon$$

for $n,m\ge n_0$ and every $\mu\in\mathcal{M}_1$. Really, if $n>m$ we have

$$P^n\mu = P^m(P^{n-m}\mu)$$

and because $m\ge n_0$

$$\|P^m(\mu - P^{n-m}\mu)\|_W \le \varepsilon.$$

Since $\mathcal{M}_1$ is a complete metric space, the sequence $(P^n\mu)_{n\ge1}$ converges to some $\mu_\star\in\mathcal{M}_1$. Obviously $P\mu_\star=\mu_\star$ and

$$\lim_{n\to\infty}\|P^n\mu-\mu_\star\|_W = \lim_{n\to\infty}\|P^n(\mu-\mu_\star)\|_W = 0$$

for every $\mu\in\mathcal{M}_1$. □

**Exercise 5.** Let $P$ be a Feller operator and $U$ its dual. Let $V: X\to[0,\infty)$ be a continuous function, bounded on bounded sets, such that

$$\lim_{\rho(x,x_0)\to\infty} V(x) = \infty$$

for some $x_0\in X$ (*Lyapunov function*). Assume, additionally, that there exist $a<1$ and $b>0$ such that

$$UV(x) \le aV(x)+b \qquad \text{for } x\in X. \tag{9}$$

Show that there exists a bounded Borel set $B \subset X$ such that

$$\liminf_{n\to\infty} P^n\mu(B) \geq \frac{1}{2}$$

for any $\mu \in \mathcal{M}_1$.

**Solution:** From (9), we obtain by an induction argument

$$U^n V(x) \leq a^n V(x) + \frac{b}{1-a} \qquad \text{for } n \in \mathbb{N}. \tag{10}$$

Fix $\varepsilon > 0$. Let $A \in \mathcal{B}(X)$ be bounded and $\mu \in \mathcal{M}_1^A$. Set

$$B = \{x : V(x) \leq q\},$$

where $q > 2b(\varepsilon(1-a))^{-1}$. From (10) and the Chebyshev inequality we obtain

$$\begin{aligned}
P^n\mu(B) &\geq 1 - \frac{1}{q}\int_X V(x)P^n\mu(\mathrm{d}x) \\
&= 1 - \frac{1}{q}\int_X U^n V(x)\mu(\mathrm{d}x) \\
&\geq 1 - \frac{1}{q}\left(a^n \int_X V(x)\mu(\mathrm{d}x) + \frac{b}{1-a}\right) \\
&\geq 1 - \frac{\varepsilon}{2} - \frac{a^n}{q}\int_X V(x)\mu(dx) \\
&\geq 1 - \frac{\varepsilon}{2} - \frac{a^n}{q}\sup_{x\in A} V(x).
\end{aligned}$$

Consequently, there exists an integer $n_0$ such that

$$P^n\mu(B) \geq 1 - \varepsilon \qquad \text{for} \quad n \geq n_0, \quad \mu \in \mathcal{M}_1^A. \tag{11}$$

Let $\varepsilon > 0$ be such that $(1-\varepsilon)^2 > 1/2$. Fix $\mu \in \mathcal{M}_1$. From Ulam's theorem it follows that there exists a campact, thus bounded, set $A \subset X$ such that $\mu(K) \geq 1 - \varepsilon$. Define the measure $\mu_A(\cdot) = \mu(\cdot \cap A)/\mu(A)$. Obviously $\mu_A \in \mathcal{M}_1^A$. Since $P^n\mu \geq (-\varepsilon)P^n\mu_A$, from (11) it follows that

$$P^n\mu(B) \geq (1-\varepsilon)P^n\mu_A(B) \geq (1-\varepsilon)(1-\varepsilon) \geq 1/2$$

for all $n$ sufficiently large. □