

Chapter 5: Invariant Measures

Invariant Measures (Stationary Distributions)

Invariant measures play a central role in understanding the long-time behavior of Markov chains. Even more, a standard motivation for studying Markov chains is to investigate or sample their invariant measures.

Definition 0.1 (Invariant measure). A positive measure m on the countable space S is **invariant** if

$$m_x = \sum_y m_y p_{y,x}, \quad x \in S \quad (1)$$

If $m \in \mathcal{P}(S)$ is a probability, we say it is an **invariant probability** or **stationary distribution** for the Markov chain.

For invariant probabilities, it is better to use the operator P , and its interpretation via the Markov process, to understand invariance. Indeed the following remark follows immediately from the definition of P (which however is defined on probabilities, not on any measure).

Remark. Let $m \in \mathcal{P}(S)$ be a probability on S . The following are equivalent

- m is invariant.
- $mP = m$.
- For each function $f \in \mathcal{B}(S)$, $\mathbb{E}_m[f(X_t)] = \mathbb{E}_m[f(X_0)]$ for all $t \geq 0$.
- For each $t \geq 0$, $\mathbf{Y} := \theta_t \mathbf{X}$ is a Markov chain with initial distribution m , where θ_t is the **time shift**.

This remark (which can be proved easily) clarifies the meaning of invariance: if \mathbf{X} starts at time 0 with distribution m , then X_t has distribution m for all $t \geq 0$.

Example 0.1. For $n \geq 2$, take $S = \mathbb{Z}_n = \{0, 1, \dots, n-1\}$, and consider the transition probabilities associated to *jumping clockwise with probability p and counterclockwise with probability $1-p$* , namely

$$p_{x,y} = \begin{cases} p & \text{if } y = x + 1 \pmod{n} \\ 1-p & \text{if } y = x - 1 \pmod{n} \\ 0 & \text{otherwise} \end{cases}$$

Then $m_x = 1/n$ is an invariant probability, since $1/n = p/n + (1-p)/n$.

Example 0.2. On $S = \mathbb{Z}$, consider the asymmetric simple random walk characterized by the transition probabilities

$$p_{x,y} = \begin{cases} p & \text{if } y = x + 1 \\ 1-p & \text{if } y = x - 1 \\ 0 & \text{otherwise} \end{cases}$$

The equation for the invariant measure becomes

$$m_x = pm_{x-1} + (1-p)m_{x+1}$$

So we have

- If $p = 0$ or $p = 1$, then $m_x = 1$ is the unique invariant measure up to multiplicative constants.
- For $p = 1/2$, all the solutions have the form $m_x = c_1 + c_2x$. However since we are looking for positive measures $m_x \geq 0$, necessarily we have $c_2 = 0$. Therefore in this case we have a unique invariant measure $m_x = 1$ up to multiplicative constants.
- For $p \in (0, 1)$, $p \neq 1/2$, all the measures $m_x = c_1 + c_2(p/1-p)^x$ are invariant for $c_1, c_2 \geq 0$.

Invariant Measures and Recurrence

Definition 0.2. The **return measure based at $x \in S$** is the measure (regarded as a measure over y)

$$\mu_y^x := \mathbb{E}_x \left[\sum_{t=0}^{\tau_x^+ - 1} \mathbf{1}_y(X_t) \right] \in [0, \infty] \quad (2)$$

namely μ_y^x counts the average time of visits to y before coming back to x (or over all times if one never comes back to x).

Theorem 0.1. Fix $x \in S$. Consider the system for the unknown variable $\mu \equiv (\mu_y)_{y \in S}$.

$$\begin{cases} \mu_x = 1 \\ \mu P \leq \mu \quad \text{meaning } \sum_z \mu_z P_{z,y} \leq \mu_y \end{cases} \quad (3)$$

Then:

- The return measure μ^x based at x is the minimal positive solution to Equation 3, in other words the minimal measure solving it.
- If x is recurrent, μ^x solves Equation 3 with equality, i.e. μ^x is invariant.
- If the Markov chain is recurrent and irreducible, then there exists a unique invariant measure μ up to a multiplicative constant, and moreover $\mu_y \in (0, \infty)$ for all y . In particular $\mu^{x'} = \mu^{x'} \mu^x$.

Proof. For $x, y \in S$, we have that

$$\sum_{t=1}^{\tau_x^+} \mathbf{1}_y(X_t) = -\mathbf{1}_{x=y} + \mathbf{1}_{\tau_x^+ < \infty, x=y} + \sum_{t=0}^{\tau_x^+ - 1} \mathbf{1}_y(X_t) \leq \sum_{t=0}^{\tau_x^+ - 1} \mathbf{1}_y(X_t) \quad (4)$$

In other words, the inequality between the l.h.s. and the r.h.s is an a.s. equality unless $x = y$ and $\tau_x^+ = +\infty$ (in all the other cases, we either added and removed 0, or added and removed 1).

- Clearly $\mu_x^x = 1$ by the definition. On the other hand, by Equation 4

$$\begin{aligned} \mu_y^x &\geq \mathbb{E}_x \left[\sum_{t=1}^{\tau_x^+} \mathbf{1}_y(X_t) \right] = \mathbb{E}_x \left[\sum_{s=1}^{\infty} \sum_{t=1}^{\tau_x^+} \mathbf{1}_y(X_t) \mathbf{1}_{\tau_x^+ = s} \right] = \mathbb{E}_x \left[\sum_{t=1}^{\infty} \mathbf{1}_y(X_t) \mathbf{1}_{\tau_x^+ \geq t} \right] \\ &= \sum_{t=1}^{\infty} \sum_{z \in S} \mathbb{P}_x(\tau_x^+ > t-1, X_{t-1} = z, X_t = y) \end{aligned}$$

Since τ_x^+ is a stopping time, $\{\tau_x^+ > t - 1\}$ is determined by (X_1, \dots, X_{t-1}) (namely $\mathbf{1}_{\{\tau_x^+ > t-1\}}$ is a measurable function of (X_1, \dots, X_{t-1})). We can thus apply to Markov property to get

$$\mu_y^x \geq \sum_{t=1}^{\infty} \sum_{z \in S} p_{z,y} \mathbb{P}_x(\tau_x^+ > t - 1, X_{t-1} = z) = \sum_{z \in S} \sum_{t=0}^{\infty} p_{z,y} \mathbb{E}_x[\mathbf{1}_{\tau_x^+ > t} \mathbf{1}_z(X_t)] = \sum_{z \in S} \mu_z^x p_{z,y} = (P\mu^x)_y$$

This proves that μ^x solves Equation 3. Minimality follows easily iterating the inequality: for μ any solution to Equation 3 and $y \neq x$

$$\begin{aligned} \mu_y &\geq (\mu P)_y = p_{x,y} + \sum_{z_1 \neq x} \mu_{z_1} p_{z_1,y} \geq p_{x,y} + \sum_{z_1 \neq x} p_{x,z_1} + \sum_{z_1, z_2 \neq x} \mu_{z_2} p_{z_2,z_1} p_{z_1,y} \geq \dots \\ &\sum_{t=1}^T \mathbb{P}_x(X_t = y, \tau_x^+ \geq t) + \sum_{z_1, \dots, z_T \neq x} \mu_{z_T} p_{z_T, z_{T-1}} \dots p_{z_2, z_1} \end{aligned}$$

As we drop the last (non-negative) term as send $T \rightarrow \infty$, we get $\mu \geq \mu^x$.

- b. If x is recurrent, then Equation 4 is an equality \mathbb{P}_x -a.s., and thus in the previous argument all inequalities are equalities.
- c. Assuming recurrence and irreducibility, let μ be an invariant measure. By irreducibility for each $x, y \in S$ there exists t such that $p_{x,y}^{(t)} > 0$, thus

$$\mu_y = \sum_z \mu_z p_{z,y}^{(t)} \geq p_{x,y}^{(t)} \mu_x \quad (5)$$

from which one deduces $\mu_y \in (0, \infty)$ for all $x, y \in S$, as long as there is one point with such property. Fix now $x \in S$, take $\mu := \mu^x$ and let now μ' be an invariant measure such that $\mu_x = 1$ (this fixes the multiplicative constant). By the minimality of μ , we have that $\lambda = \mu' - \mu \geq 0$, and $\lambda P^t = \lambda$, but $\lambda_x = 0$. Then reasoning as in Equation 5, we get $\lambda_y = 0$ for all $y \in S$. □

Definition 0.3. A point $x \in S$ is called **positive-recurrent** if $\mathbb{E}_x[\tau_x^+] < \infty$. In particular such an x is recurrent. A point $x \in S$ is called **null-recurrent** if it is recurrent and $\mathbb{E}_x[\tau_x^+] = \infty$.

Theorem 0.2. For an irreducible Markov Chain, the following are equivalent:

- a. Every state $x \in S$ is positive-recurrent.
- b. There exists a positive-recurrent state $x \in S$.
- c. There exists a (necessarily unique) invariant probability measure $m \in \mathcal{P}(S)$.

If any of these conditions holds, then $m_x = 1/\mathbb{E}_x[\tau_x^+]$. In particular

$$\sum_{x \in S} \frac{1}{\mathbb{E}_x[\tau_x^+]} = 1$$

Proof.

- **a.** \implies **b.**, is trivial.

• **b.** \implies **c.**, we know that μ^x is invariant and finite in each point, from point (b) in Theorem 0.1. Notice that

$$\sum_y \mu_y^x = \mathbb{E}_x \left[\sum_{t=0}^{\tau_x^+ - 1} \sum_{y \in S} \mathbf{1}_y(X_t) \right] = \mathbb{E}_x [\tau_x^+] < \infty \quad (6)$$

by assumption. So $m = \mu^x / \mathbb{E}_x [\tau_x^+]$ is the unique invariant probability measure.

• **c.** \implies **a.** Let m be an invariant probability, then $\mu := m/m_x$ satisfies Equation 3, and thus μ^x (as defined in Equation 2 and being the minimal solution to Equation 3), satisfies $\mu_y \geq \mu_x$ for all $y \in S$. Therefore by Equation 6

$$1 = \sum_y m_y \geq m_x \sum_y \mu_y^x = m_x \mathbb{E}_x [\tau_x^+]$$

which implies that x is positive-recurrent, as well as the explicit formula for m_x .

□

Example 0.3. Suppose that we shuffle a deck of N cards, repeating some random shuffling action over and over. Assume that:

- It is a complete mixing method, meaning that, in finitely many iterations of the shuffling, we can obtain any order of the cards (e.g. we do not just swipe the first and second card repeatedly).
- We do not see the cards when we mix them (that is, the probability of performing a given shuffling does not depend on the order of the cards).

Let us say that we start with a completely ordered deck. Then the average number of shuffling actions needed to come back to the original order is $N!$, no matter the allowed shuffling actions (method) or the exact probabilities of each of them. Indeed, the order of the cards X_t at each iteration is a Markov chain on the space S of permutations over N cards. The first hypothesis states that this Markov chain is irreducible, the second hypothesis states that $p_{\sigma \circ x, \sigma \circ y} = p_{x,y} = p_{e, x^{-1} \circ y}$ for each permutation σ . Thus the uniform measure on S is invariant, namely $m_x = 1/N!$ is the unique invariant probability. Therefore $\mathbb{E}_x [\tau_x^+] = N!$ for all $x \in S$.

Convergence to the invariant measure

If X_t is an irreducible, positive-recurrent Markov chain, it can be proven (as a consequence of the law of large numbers applied to the return times τ_x^+)

$$\lim_T \frac{1}{T} \sum_{t=0}^{T-1} f(X_t) = m(f) \quad (7)$$

a.s., for all function $f \in L^1(m)$. We will get refinements of this statement in later chapters. Here we want to focus on a stronger statement: does the law of X_t converge to m as $t \rightarrow \infty$? Notice that this is in principle a different question compared to Equation 7. Take $f = \mathbf{1}_x$, for some fixed $x \in S$. Then Equation 7 states that the number of times the chain passes in x in the first T steps, is roughly $m_x T$. This holds for any irreducible, positive-recurrent chain and is quite intuitive.

In Equation 7 with $f = \mathbf{1}_x$, taking expectation we get

$$\lim_T \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{P}(X_t = x) = m_x \quad (8)$$

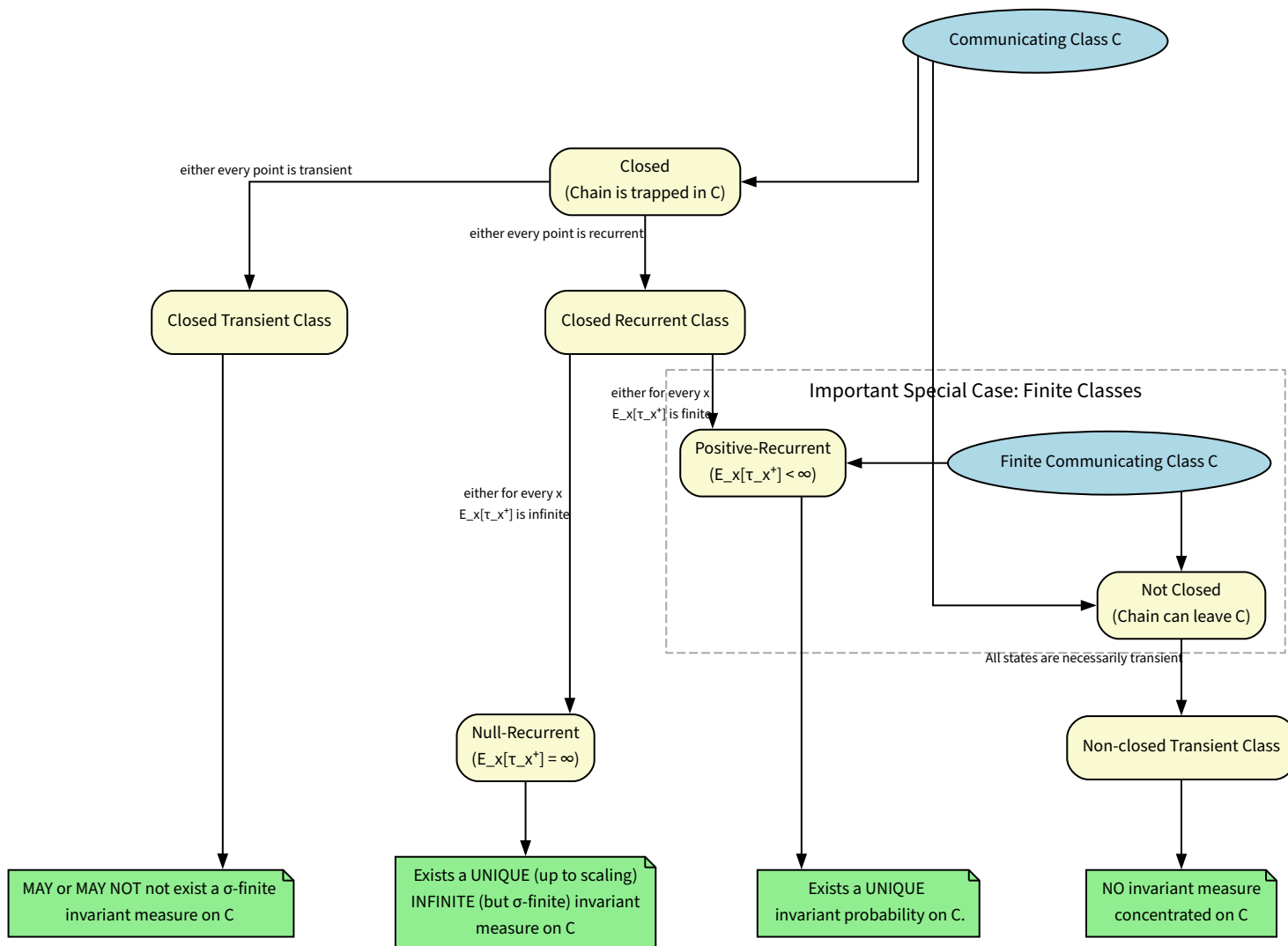



Figure 1

In other words, the Cesaro mean (in time) of $\mathbb{P}(X_t = x)$ converges to m_x . One may wonder whether $\mathbb{P}(X_t = x)$ converges to m_x , not just its Cesaro mean. In other words, we are asking if X_t converges to m in distribution. On the one hand, this result would be stronger than Equation 7, since it concerns the direct limit in t of $\mathbb{P}(X_t = x)$, or more in general of observables like $\mathbb{E}[f(X_t)]$. On the other hand, we are just making a statement for the distribution of X_t , not an a.s. statement, and on the other hand it is clear that for a generic chain, the chain X_t will continue to randomly jump and would not converge anywhere as $t \rightarrow \infty$, unless there are closed singletons (known as absorbing states). So we are asking whether we can trade a weaker stochastic notion of convergence (in distribution) for a stronger notion of limit in time (usual direct limit vs Cesaro mean).

However, it is easy to see that $\mathbb{P}(X_t = x) \rightarrow m_x$ cannot be true, even on a finite state space, without further assumptions. Indeed, the chain on $Z_{2n} = \{0, 1, \dots, 2n\}$ that each steps jumps from i to $i + 1 \pmod{2n}$ with probability 1, features an obvious periodicity: $X_t - X_0$ is even if t is even, and odd if t is odd. So while the uniform measure is the unique invariant measure, $\mathbb{P}_0(X_t = x)$ will not converge (indeed, it will converge to the uniform measure on even/odd numbers, when taking the limit along even/odd times t). This periodic constraint phenomenon, however, is the only one that may hinder convergence, and justifies the following definition.

Definition 0.4 (aperiodic). A state x is **aperiodic** if there exists T such that $p_{x,x}^{(t)} > 0$ for all $t \geq T$.

Exercise 0.1. Let z be an aperiodic state of an *irreducible* Markov chain. Prove that for each $x, y \in S$ there exists a time T such that $p_{x,y}^{(t)} > 0$ for all $t \geq T$. In particular for an irreducible Markov chain, a state is aperiodic iff all the states are aperiodic. The chain is called aperiodic in this case.

 Solution

There exists s, s' such that $p_{x,z}^{(s)} > 0, p_{z,y}^{(s')} > 0$. Then

$$p_{x,y}^{(s+t+s')} \geq p_{x,z}^{(s)} p_{z,z}^{(t)} p_{z,y}^{(s')}$$

Thus if $p_{z,z}^{(t)} > 0$ for $t \geq T_z$, then $p_{x,y}^{(t)} > 0$ for all $t \geq T_z + s + s'$.

Exercise 0.2. Define the set $\Pi_x := \{t \geq 1 : p_{x,x}^{(t)} > 0\}$. The **period of x** is the greatest common divisor of Π_x . Prove that for an irreducible Markov chain, all points share the same period.

We are ready to prove that periodicity is indeed the only obstruction to the convergence of the law of X_t to the invariant probability. The proof, despite simple, is really influential and a favorite of many probabilists.

Theorem 0.3. For an **irreducible aperiodic** Markov chain \mathbf{X} with **invariant probability** m (in particular \mathbf{X} is positive-recurrent) it holds

$$\lim_{t \rightarrow \infty} \sup_{x \in S} |\mathbb{P}(X_t = x) - m_x| = 0$$

Proof. Let \mathbf{X}' be a Markov chain on the same state space, independent of \mathbf{X} , and starting with initial distribution m . Notice that the pair $\boxtimes := (\mathbf{X}, \mathbf{X}')$ is a Markov chain on the state space $\Sigma := S \times S$ with transition probabilities

$$q_{(x,x'),(y,y')} := p_{x,y} p_{x',y'}$$

From this definition, it is intuitive and easy to check that

$$q_{(x,x'),(y,y')}^{(t)} := p_{x,y}^{(t)} p_{x',y'}^{(t)} \tag{9}$$

The Markov chain \boxtimes

- admits the measure $\mu_{(x,x')} := m_x m_{x'}$ as **invariant probability**.
- is **irreducible** (for this we need the aperiodicity of the original chain). Indeed from Equation 9 and Exercise 0.1, for all $(x, x'), (y, y') \in \Sigma$, $q_{(x,x'),(y,y')}^{(t)} > 0$ for t large enough (so \mathfrak{N} is actually aperiodic, although we just need irreducibility here).

Therefore \mathfrak{N} is positive-recurrent from Theorem 0.2. Let $\Delta := \{(x, x), x \in S\} \subset \Sigma$ be the diagonal in Σ . Then

$$\tau := \inf\{t \geq 0 : X_t = X'_t\} = \tau_\Delta(\mathfrak{N})$$

is the hitting time of Δ for \mathfrak{N} , and thus a stopping time. In particular

$$\mathbb{P}(X_t = x, \tau \leq t) = \mathbb{P}(X'_t = x, \tau \leq t)$$

as indeed both coincide with $\mathbb{P}(Y_t = x, \tau \leq t)$ where $Y_t = X_t \mathbf{1}_{t \leq \tau} + X'_t \mathbf{1}_{t > \tau}$. Therefore

$$|\mathbb{P}(X_t = x) - m_x| = |\mathbb{P}(X_t = x) - \mathbb{P}(X'_t = x)| = |\mathbb{P}(X_t = x, \tau > t) - \mathbb{P}(X'_t = x, \tau > t)| \leq \mathbb{P}(\tau > t)$$

However, since \mathfrak{N} is irreducible and positive-recurrent $\mathbb{P}(\tau < \infty) = 1$, and thus the r.h.s. vanishes as we take $t \rightarrow \infty$. \square

The operator P and invariant measures

Proposition 0.1. *Let m be an invariant measure. For $q \in [1, \infty]$, the operator P is a contraction in $L^q(m)$. Namely*

$$\begin{aligned} \int |Pf|^q dm &\equiv \sum_x |Pf(x)|^q m_x \leq \sum_x |f(x)|^q m_x \equiv \int |f|^q dm, \quad q \in [1, \infty) \\ \|Pf\|_{L^\infty(m)} &\equiv \sup_{x: m_x > 0} |Pf(x)| \leq \sup_{x: m_x > 0} |f(x)| \equiv \|f\|_{L^\infty(m)} \end{aligned}$$

Proof. Take $q \in (1, \infty)$. By **Jensen's inequality** applied to the function $u \mapsto |u|^q$, we have that

$$|Pf(x)|^q = |\mathbb{E}_x[f(X_1)]|^q \leq \mathbb{E}_x[|f(X_1)|^q] = Pg(x)$$

where $g(\cdot) = |f(\cdot)|^q$. If we sum (integrate over m) we get, since $mP = m$, $m(|Pf|^q) \leq m(Pg) = m(g) = m(|f|^q)$. Or more explicitly

$$\sum_x |Pf(x)|^q m_x \leq \sum_x m_x \sum_y p_{x,y} g(y) = \sum_{x,y} m_x p_{x,y} g(y) = \sum_y m_y g(y) = \sum_y m_y |f(y)|^q$$

where in the last equality we used the invariance $\sum_x m_x p_{x,y} = m_y$. The case with $q = \infty$ is an immediate consequence of the fact that $|Pf(x)| \leq c$ if $|f| \leq c$. \square

The adjoint operator

In particular it is useful to consider P acting on $L^2(m)$. Since P is bounded in $L^2(m)$, the operator $P^\dagger: L^2(m) \rightarrow L^2(m)$ is defined by

$$m(fPg) := m(gP^\dagger f) \quad \text{for all } f, g \in L^2(m) \quad (10)$$

It is immediate to check that P^\dagger has entries such that

$$m_x p_{x,y}^\dagger = m_y p_{y,x} \quad (11)$$

Remark. If m is invariant, then $p_{x,y}^\dagger$ as in Equation 11 are transition probabilities, namely $\sum_y p_{x,y}^\dagger = 1$.

Conversely, if for some measure m , $p_{x,y}^\dagger$ as in Equation 11 are transition probabilities, then m is invariant.

Reversibility

Exercise 0.3. Let X_0, X_1, \dots be a Markov chain on a countable state space S , with transition probabilities $(p_{x,y})_{x,y}$, and invariant probability m . Assume that X_0 (and thus all of the X_t) has distribution m . Let $T \geq 1$, and consider the backward-time chain Y_0, Y_1, \dots, Y_T defined as $Y_t = X_{T-t}$. Check that $(Y_t)_{t \leq T}$ is a Markov chain, and find its transition probabilities.

Definition 0.5. A measure m is called **reversible** for the transition probabilities $(p_{x,y})$ if for all $x, y \in S$

$$m_x p_{x,y} = m_y p_{y,x} \quad (12)$$

or equivalently if $P^\dagger = P$.

If a Markov chain admits a unique invariant measure, the chain is called **reversible** if the invariant measure is reversible.

Notice that reversibility implies invariance (take the sum over y in Equation 12).

Markov triples

Definition 0.6. A **Markov triple** is a triple (S, P, m) where

- S is a countable set.
- P is a Markov operator, meaning $Pf(x) = \mathbb{E}_x[f(X_t)]$ for some Markov chain \mathbf{X} , and for all bounded measurable $f: S \rightarrow \mathbb{R}$.
- m is a σ -finite invariant measure on S , namely $m(Pf) = m(f)$, for $f \in L^1(S)$.

Definition 0.7. A measurable map $\phi: S \rightarrow S'$ is a **morphism of Markov triples**

$\phi: (S, P, m) \rightarrow (S', P', m')$ if for all bounded measurable $f: S' \rightarrow \mathbb{R}$

$$P(f \circ \phi) = (P'f) \circ \phi \quad \text{and} \quad m' = m \circ \phi^{-1}.$$

In this case, $(Y_t := \phi(X_t))_t$ is a Markov chain with Markov operator P' and invariant measure m' , whenever \mathbf{X} is a Markov chain with transition operator P and invariant measure m .

Hereafter, given a Markov triple (S, P, m) we denote $\langle f, g \rangle_m := \sum_x f(x)g(x)m_x$ the scalar product in $L^2(m)$.

Definition 0.8. Let (S, P, m) be a Markov triple. If $f, g \in L^2(m)$ define the **Dirichlet form**

$$\mathcal{E}(f, g) = \langle f, (I - P)g \rangle_m$$

Proposition 0.2. Denote P^\dagger the *adjoint* of P w.r.t. the measure m and $P^s := (P + P^\dagger)/2$, and similarly $p_{x,y}^s := (p_{x,y} + p_{x,y}^\dagger)/2$, $\mathcal{E}^\dagger(f, g) := \langle f, (I - P^\dagger)g \rangle_m$, $\mathcal{E}^s(f, g) := \langle f, (I - P^s)g \rangle_m$. Then it holds

- $\mathcal{E}^\dagger(f, g) = \mathcal{E}(g, f)$. In particular m is reversible iff \mathcal{E} is symmetric.
- $\mathcal{E}(f, f) = \mathcal{E}^s(f, f) = \frac{1}{2} \sum_{x,y \in S} m_x p_{x,y}^s (f(x) - f(y))^2$.

$$c. \mathcal{E}^s(f, g) = \frac{1}{2} \sum_{x, y \in S} m_x p_{x, y}^s (f(x) - f(y))(g(x) - g(y)).$$

$$\mathcal{E}(f, f) = \langle f, (I - P)g \rangle_m$$

Proof.

- a. It follows immediately from the definition of P^\dagger .
- b. Notice that from a.

$$\begin{aligned} \mathcal{E}(f, f) &= \mathcal{E}^\dagger(f, f) = \frac{1}{2} (\mathcal{E}(f, f) + \mathcal{E}^\dagger(f, f)) = \mathcal{E}^s(f, f) \\ &= \sum_{x, y} (m_x p_{x, y}^s (f(x) - f(y))f(x)) = \frac{1}{2} \sum_{x, y} (m_x p_{x, y}^s (f(x) - f(y))f(x) + m_y p_{y, x}^s (f(y) - f(x))f(y)) \end{aligned}$$

and since $m_x p_{x, y}^s = m_y p_{y, x}^s$, the terms are easily rearranges in a square as in the statement.

- c. It follows by polarization from b., since \mathcal{E}^s is a symmetric bilinear form.

□

Definition 0.9. An invariant measure m is **ergodic** if any of the following equivalent conditions holds

- If $Pf = f$ m -a.s., then f is constant m -a.s.
- If $\mathcal{E}(f, f) = 0$, then f is constant.

Another look at linear equations

Let (S, P, m) be a Markov triple, and let $c: S \times S \rightarrow \mathbb{R}$ and $g: S \rightarrow \mathbb{R}$ be given, and let $A \subset S$. We assume that $c \geq 0$ to be sure that all the sums are well-defined. Define

$$u(x) := \mathbb{E}_x \left[g(X_{\tau_A}) \mathbf{1}_{\tau_A < \infty} + \sum_{t=1}^{\tau_A} c(X_{t-1}, X_t) \right] \in [0, \infty]$$

Then, conditioning on the first step X_1 , we get that u satisfies (it is indeed the minimal solution if $c, g \geq 0$)

$$\begin{cases} (I - P)u = f & \text{on } A^c \\ u = f & \text{on } A \end{cases} \quad (13)$$

where

$$\begin{cases} f(x) = \sum_y p_{x, y} c(x, y) & \text{for } x \in A^c \\ f(x) = g(x) & \text{for } x \in A \end{cases}$$

Remark. Assume that (S, P, m) is ergodic and A non-empty. Equation Equation 13 admits at most one solution in $L^2(m)$. Indeed, let u, v be two solutions, and set $w = u - v$. Then w satisfies Equation 13 with $f \equiv 0$. In particular $w(x)((I - P)w)(x) = 0$ everywhere on S (if $x \in A$ then $w(x) = 0$, otherwise $((I - P)w)(x) = 0$). Thus $\mathcal{E}(w, (I - P)w) = 0$, and w is constant by ergodicity. But $w = 0$ on A , thus $w \equiv 0$.

i Abstraction

Let S be a Polish space, and recall that in this case we have that the transition probabilities are measurable maps $p: S \rightarrow \mathcal{P}(S)$, meaning that $p(x, A) = \mathbb{P}(X_1 \in A | X_0 = x)$ represents the probability of jumping to A when at point $x \in S$. In this case, we have that P has the following three properties

- $P\mathbf{1} = \mathbf{1}$.
- $Pf \geq 0$ if $f \geq 0$.
- $Pf_n \rightarrow 0$ if f_n is a monotone sequence converging pointwise to 0.

It is not hard to check that these properties characterize the Markov operators. In other words, if a linear operator acting on bounded measurable functions satisfies them, then $Pf(x) = \mathbb{E}_x[f(X_1)]$ for some Markov chain with transition probabilities $p(x, A) := (P\mathbf{1}_A)(x)$.