Lecture 26 · 16.7.2021 · 10:15–12:00 via Zoom

Invariant measures for irreducible chains (conclusion)

In the last lecture we have seen the concept of invariant measure and invariant probability, i.e.

$$\mu = \mu P.$$

An equivalent name is "stationary". Indeed the law of a Markov chain started from an invariant probability is stationary in time as we have seen.

We have also seen that an invariant measure ρ for an irreducible chain must satisfy $\rho(y) \in (0, +\infty)$ for all $y \in E$.

And moreover we have seen that for an irreducible and recurrent chain we can construct an invariant measure ν^x for any $x \in E$, via the formula

$$\nu^x(y) = \mathbb{E}_x\left[\sum_{n=1}^{S_x} \mathbb{1}_{X_n = y}\right] = \mathbb{E}_x\left[\sum_{n=0}^{S_x - 1} \mathbb{1}_{X_n = y}\right], \qquad y \in E.$$

where $S_x = \inf\{n \geqslant 1 : X_n = x\}$ is the return time to $x \in E$. Recall that $\nu^x(x) = 1$.

To address the question of uniqueness we prove the following.

Lemma 1. *Consider an irreducible discrete chain and let ρ, μ two invariant measures such that there exists $x_* \in E$ for which*

$$\frac{\rho(x)}{\mu(x)} \geqslant \frac{\rho(x_*)}{\mu(x_*)}, \qquad x \in E.$$

Then ρ and μ are proportional, i.e. $\rho(x) / \mu(x) = \text{constant}$.

Proof. Using the invariance of ρ and μ we have for any $n \geqslant 1$

$$\rho(x_*) = \rho P^n(x_*) = \sum_{y \in E} \rho(y) P^n(y, x_*) \geqslant \frac{\rho(x_*)}{\mu(x_*)} \sum_{y \in E} \mu(y) P^n(y, x_*)$$

$$= \frac{\rho(x_*)}{\mu(x_*)} \mu P^n(x_*) = \frac{\rho(x_*)}{\mu(x_*)} \mu(x_*) = \rho(x_*)$$

which means that the middle $\geqslant$ is actually $=$. Now this means that

$$\rho(y) P^n(y, x_*) = \frac{\rho(x_*)}{\mu(x_*)} \mu(y) P^n(y, x_*)$$

for any $y \in E, n \geqslant 1$. Since the chain is irreducible, for any $y \in E$ we can choose n such that $P^n(y, x_*) > 0$. Therefore we can divide by $P^n(y, x_*)$ and have the claim:

$$\frac{\rho(y)}{\mu(y)} = \frac{\rho(x_*)}{\mu(x_*)}, \qquad y \in E.$$

□

Note that in particular this implies that if the state space is finite then there all the invariant measures are proportional (indeed $\frac{\rho(y)}{\mu(y)}$ has always a minimum) and therefore there exist at most one invariant probability measure.

It is clear that an irreducible Markov chain in finite space must be recurrent which then implies that we have invariant measures ν^x as introduced above. Since the state space is finite this measure is certainly normalizable and we conclude that there exists a unique invariant probability given by

$$\pi(x) = \frac{\nu^x(x)}{\sum_{y\in E} \nu^x(y)} = \frac{1}{\mathbb{E}_x[S_x]}, \qquad x \in E.$$

We record this result in the following theorem.

Theorem. *An irreducible finite chain is necessarily recurrent and has a unique invariant probability given by*

$$\pi(x) = \frac{1}{\mathbb{E}_x[S_x]}, \qquad x \in E.$$

Proof. The argument has just been given above. □

If the chain is not finite but still recurrent we have to understand better the quantity

$$\rho(y) / \nu^x(y)$$

for some invariant measure ρ and some state $x \in E$. We need to prove that this quantity has a minimum.

Theorem. *If the chain is recurrent and irreducible then all the invariant measures are proportional and in particular we have $\nu^y = C_{x,y}\nu^x$ for all $x, y \in E$.*

Proof. Let ρ be an invariant measure, then we have for all $x, z \in E$ with $z \neq x$

$$\rho(z) = (\rho P)(z) = \rho(x)P(x,z) + \sum_{y\neq x} \rho(y)P(y,z).$$

We now iterate this relation to get

$$\rho(z) = \rho(x)P(x,z) + \sum_{y\neq x} \rho(x)P(x,y)P(y,z) + \sum_{y\neq x}\sum_{y'\neq x} \rho(y')P(y',y)P(y,z)$$

$$= \cdots$$

$$\geqslant \rho(x)P(x,z) + \sum_{y\neq x} \rho(x)P(x,y)P(y,z) + \rho(x)\sum_{y\neq x}\sum_{y'\neq x} P(x,y')P(y',y)P(y,z) + \cdots$$

$$= \rho(x)\left[\sum_{k\geqslant 1} \mathbb{P}_x(X_1 \neq x, X_2 \neq x, \ldots, X_k = z)\right] = \rho(x)\left\{\sum_{k\geqslant 0} \mathbb{E}_x[\mathbb{1}_{S_x>k}\mathbb{1}_{X_k=z}]\right\} = \rho(x)\nu^x(z)$$

since $z \neq x$. Therefore we have found that

$$\frac{\rho(z)}{\nu^x(z)} \geqslant \frac{\rho(x)}{\nu^x(x)}, \qquad z \in E,$$

which by Lemma 1 it implies that ρ must be proportional to ν^x and therefore all the invariant measures are proportional. In particular we have $\nu^y = C_{x,y}\nu^x$. □

Now it remains the question whether we can normalize such unique invariant measure to get an invariant probability. Note that we have

$$\nu^x(E) = \sum_{y\in E} \mathbb{E}_x\left[\sum_{k=0}^{S_x-1} \mathbb{1}_{X_k=y}\right] = \mathbb{E}_x\left[\sum_{k=0}^{S_x-1} 1\right] = \mathbb{E}_x[S_x]$$

So if we have $\mathbb{E}_x[S_x] < \infty$ for some $x \in E$ then

$$\nu^y(E) = C_{x,y}\nu^x(E) < \infty$$

so we have $\mathbb{E}_y[S_y] < \infty$ for all $y \in E$ and the chain is positive recurrent, moreover if we define the invariant probability

$$\pi^x(z) := \frac{\nu^x(z)}{\nu^x(E)}$$

then we have for any $x, y \in E$

$$\pi^x(z) = \frac{\nu^x(z)}{\nu^x(E)} = \frac{\nu^y(z)}{\nu^y(E)} = \pi^y(z)$$

and there exist only one invariant probability $\pi = \pi^x$ for all $x \in E$ and such that

$$\pi(x) = \pi^x(x) = \frac{\nu^x(x)}{\nu^x(E)} = \frac{1}{\nu^x(E)} = \frac{1}{\mathbb{E}_x(S_x)}.$$

We can summarize these considerations in the following theorem:

Theorem. *If the chain is recurrent and irreducible (but not necessarily finite) then all the invariant probabilites are proportional to ν^x for some $x \in E$. The chain is positive recurrent if and only if the invariant probability is normalizable and in that case there exists only one invariant probability π such that*

$$\pi(x) = \frac{1}{\mathbb{E}_x(S_x)}, \qquad x \in E.$$

I want to conclude giving an application of these results.

Example. Let (E, G) be a finite connected graph where E is the set of vertices and $G \subseteq \{\{x, y\}: x, y \in E\}$ the set of edges. We write $x \sim y$ if $\{x, y\} \in G$. And let $d(x) = \#\{y \in E: x \sim y\}$ the degree of the vertex $x \in E$. We can define a Markov chain on E by letting

$$P(x, y) = \frac{\mathbb{1}_{x\sim y}}{d(x)}.$$

This chain is irreducible and recurrent since it is finite and therefore there exists a unique invariant measure π. To compute it note that the measure $\mu(x) = d(x)$ is invariant:

$$(\mu P)(x) = \sum_{y\in E} \mu(y)P(y, x) = \sum_{y\in E} \mu(y)\frac{\mathbb{1}_{y\sim x}}{d(y)} = \sum_{y\in E} \mathbb{1}_{y\sim x} = \sum_{y\in E} \mathbb{1}_{x\sim y} = d(x) = \mu(x).$$

Therefore π must be proportional to μ, and since

$$\mu(E) = \sum_{y \in E} \mu(y) = \sum_{y \in E} d(y) = \sum_{y \in E} \sum_{x \in E} \mathbb{1}_{y \sim x} = 2|G|$$

we have

$$\pi(x) = \frac{\mu(x)}{\sum_{y \in E} \mu(y)} = \frac{\mu(x)}{2|G|} = \frac{d(x)}{2|G|}.$$

As a consequence we also have that a formula for the average return time:

$$\mathbb{E}_x[S_x] = \frac{1}{\pi(x)} = \frac{2|G|}{d(x)}.$$

Remark 2. If the graph in the example is not finite then the chain is not positive recurrent. It can be either transient or null recurrent (i.e. recurrent but not positive recurrent). Examples are $\mathbb{Z}^d$ which give a transient chain for $d \geqslant 3$ and a recurrent chain for $d = 1, 2$ but not positive recurrent.

The lectures end here