Lecture 13 notes

klin@math.arizona.edu

February 27, 2020
Last revised: February 29, 2020

Today's agenda:

- A summary of key theorems regarding stationary distributions
- The Perron-Frobenius Theorem.
- A sketch of the Convergence Theorem for 2-state chains using the Perron-Frobenius Theorem.

Stationary distributions

First, there seems to be some confusion about basic definitions:

- A probability distribution on the state space S is *stationary* if for all $x \in S$,

$$\sum_{y \in S} \pi(y)p(y,x) = \pi(x). \tag{1}$$

- A Markov chain satisfied detailed balance with respect to a probability distribution π on the state space S if for all $x, y \in S$,

$$\pi(y)p(y,x) = \pi(x)p(x,y). \tag{2}$$

  Note that detailed balance is strictly stronger than stationarity: it implies stationarity but stationarity does not imply detailed balance.

- A Markov chain is said to *converge to equilibrium* if

$$p^n(x,y) \to \pi(y) \tag{3}$$

  as $n \to \infty$. Equivalently, a chain converges to equilibrium if for any initial distribution,

$$P(X_n = y) \to \pi(y). \tag{4}$$

Here is a summary of key theorems from the text.

Some of the theorems can be stated a bit more generally without much effort. For example, the assumptions of irreducibility and recurrence in Theorem 1.20 are not really necessary:

Theorem (1.20').

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = \begin{cases} \frac{1}{E_y T_y}, & T_y < \infty \\ 0, & T_y = \infty. \end{cases} \tag{5}$$

Proof. If $T_y = \infty$ then $N_n(y) = 0$, so the conclusion is trivial.

If $T_y < \infty$ (so the chain does reach y at some point), and y is recurrent, then the proof of Theorem 1.20 applies. (Check it!)

Finally, if y is transient, then[1]

$$P_y(N(y) = \infty) = \lim_{k\to\infty} P_y(N(y) \geqslant k) \tag{6a}$$

$$= \lim_{k\to\infty} \rho_{xy}\rho_{yy}^{k-1} \tag{6b}$$

$$= 0. \tag{6c}$$

So $N(y) < \infty$ with probability 1. Since $N_n(y) \leqslant N(y)$ for all n, the conclusion follows. □

Another theorem that generalizes slightly is Theorem 1.19. To avoid complications, here I just state a result for finite state spaces.

Theorem (1.19'). *If $|S| < \infty$ and S is a disjoint union $T \cup R$, where T is the set of all transient states and R is closed, irreducible, and aperiodic. Then $p^n(x, y) \to \pi(y)$ for all $x, y \in S$.*

Proof. We go through the different cases:

If $x, y \in R$, then the result is just Theorem 1.19.

If $x \in R$ and $y \in T$, then from the proof of the Decomposition Theorem we know that $p^n(x, y) = 0$.

If $x, y \in T$ and x does not communicate with y, then $p^n(x, y) = 0$ for all n.

If $x, y \in T$ and x does communicate with y, then there is an m such that $p^m(x, y) > 0$. And $p^{m+n}(x, y) \geqslant p^m(x, y) p^n(y, y)$.

$$\sum_{n=1}^{\infty} p^n(x, y) = \sum_{n=1}^{\infty} P_x(X_n = y) \tag{7a}$$

$$= E_x \sum_{n=1}^{\infty} \mathbb{1}_{(X_n = y)} \tag{7b}$$

$$= E_x N(y) \tag{7c}$$

$$= E_x\big(N(y) \big| T_y < \infty\big) \cdot \rho_{xy} + \underbrace{E_x\big(N(y) \big| T_y = \infty\big)}_{0} \cdot (1 - \rho_{xy}) \tag{7d}$$

$$= E_y N(y) \cdot \rho_{xy}. \tag{7e}$$

(The next-to-last line uses the Markov property.) Since y is transient, we know $E_y N(y) < \infty$, so $p^n(x, y) \to 0$.

[1] I said this incorrectly in lecture today. Duh!

Finally, suppose $x \in T$ and $y \in R$. We know from the Decomposition Theorem that there is some $z \in R$ such that $x \to z$ and $\rho_{zx} = 0$. Since R is recurrent (because it is closed, irreducible, and finite) we have $\rho_{zy} = 1$, and $x \to y$ as well. As in Eq. (7), we have

$$P_x(X_n = y) = \sum_{k=1}^{\infty} P_x(X_n = y | T_y = k) \cdot P_x(T_y = k) \tag{8}$$

For each k, we have

$$p^n(x,y) = P_x(X_n = y | T_y = k) \tag{9a}$$
$$= P_y(X_{n-k} = y) \tag{9b}$$

for $n > k$, by the Markov property. So[2]

$$\lim_{n\to\infty} p^n(x,y) = \lim_{n\to\infty} P_y(X_n = y) \tag{10a}$$
$$= \pi(y) \tag{10b}$$

by Theorem 1.19, and

$$\lim_{n\to\infty} P_x(X_n = y) = \pi(y) \sum_{k=1}^{\infty} P_x(T_y = k) \tag{11a}$$
$$= \pi(y) \cdot \rho_{xy}. \tag{11b}$$

The last thing we need is $\rho_{xy} = 1$. Heuristically, this is because $x \in T$, so eventually it reaches a recurrent state $z \in R$. Since R is closed and irreducible (and hence recurrent), $\rho_{zy} = 1$ for any z. So $\rho_{xy} = 1$.[3] □

Perron-Frobenius Theorem

Theorem. *Let P be a stochastic matrix, and suppose all states are recurrent. Then*

1) *1 is an eigenvalue of P, and it has a left eigenvector π with all nonnegative entries. For all eigenvalues λ of P, $|\lambda| \leqslant 1$.*

2) *If P is irreducible, then the eigenvalue 1 has multiplicity 1, i.e., there is a unique (left) eigenvector π such that $\pi \cdot P = \pi$, $\pi(x) > 0$ for all x, and $\sum_x \pi(x) = 1$. More generally, if the chain has k closed irreducible blocks, then the eigenvalue 1 has multiplicity k, and there are exactly k linearly independent stationary distributions.*

[2] This is not rigorous, because I do not justify exchanging the limit with the infinite sum. But it hopefully gives you an idea for why this is true – think about what the chain is doing!

[3] This can be made more airtight with a little effort.

3) *If P is irreducible with period d, then P has exactly d eigenvalues λ with $|\lambda| = 1$, corresponding to the roots of $\lambda^d = 1$, i.e., they are the dth roots of unity. All other eigenvalues of P have absolute value* < 1.

4) *If P is irreducible and aperiodic, then all eigenvalues $\lambda \neq 1$ have* $|\lambda| < 1$.

5) *If P is irreducible and satisfies detailed balance with respect to the (unique) stationary distribution π, then all eigenvalues are real.*

Notes:

1) See this snapshot of the Wikipedia page for more information.

2) Here are some examples and the associated Jupyter notebook.