MATH4240 Tutorial 8 Notes

1 Periodicity

- The *period* d_x of a state x in a Markov chain is the gcd of all t that x can return to x after t steps,

  $$d_x = \gcd \left\{ t \geq 1 \;\middle|\; P^t(x,x) > 0 \right\}$$

  and $\{ t \geq 1 \mid P^t(x,x) > 0 \}$ contains all but finitely many multiples of d_x.
  (If the set is empty, the period is undefined or infinity, depending on your convention.)

- If $x \to y$ and $y \to x$, then $d_x = d_y$.
  In particular, in an irreducible closed set, every state has the same period

- An irreducible chain is called *aperiodic* / *acyclic* if the period is 1

Recall that a transition matrix is called *primitive* if there exists $N \geq 1$ such that every entry in P^n is positive for all $n \geq N$. The following lemma is easy to show:

Lemma 1.1. *The transition matrix of a finite irreducible chain is primitive if and only if the chain is aperiodic.*[1]

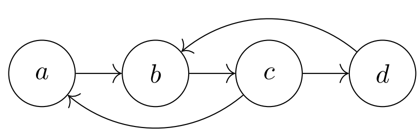
As noted before, a primitive matrix satisfies the conditions of the theorem from the lecture, so we have the following result:

Theorem 1.2. *If P is the transition matrix of a finite* ***irreducible aperiodic*** *chain, then $\lim_n P^n = \begin{pmatrix} \pi & \dots & \pi \end{pmatrix}^\mathsf{T}$ where π is the unique normalized left eigenvector of P associated to the eigenvalue 1, which is also the unique stationary distribution of the chain.*

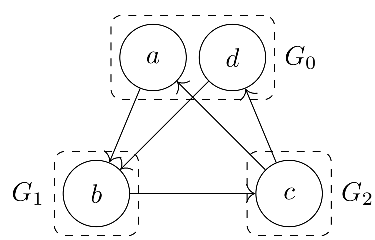
Together with the result on positive recurrence,

Corollary 1.3. *A finite* ***irreducible aperiodic*** *Markov chain is* ***positive recurrent****.*

Example 1. Consider the chain with transition diagram (ignoring the transition probabilities)



As noted in lecture, this chain has period $d = 3$. It may be easier to notice if we draw the diagram in this way:



Notice that by partitioning the state space into $G_0, \dots, G_{d-1} = G_2$, every state in G_i can only go to some states in G_{i+1} (and G_{d-1} to G_0), and there is no transition to states in other classes.

[1] An irreducible aperiodic Markov chain is sometimes also called *ergodic*.

In fact, it is easy to show the following result (you should try to prove it):

Theorem 1.4. *Consider the binary relation $\sim$ on the state space of an irreducible chain with transition matrix P where $x \sim y$ if $P^{nd}(x,y) \neq 0$ for some $n \geq 1$. Then it is an equivalence relation that satisfies the following properties:*

- *there are exactly d equivalence classes (**cyclic classes**)*
- *the cyclic classes can be labeled as $G_0, \ldots, G_{d-1}$ in a way that, in one step, a state in G_i can only reach another state in G_{i+1} (identifying $G_d = G_0$)*
- *the transition matrix can be rearranged into the following **cyclic block form**[2]*

$$P_{\text{cyc}} = \begin{array}{c} \\ G_0 \\ G_1 \\ G_2 \\ \vdots \\ G_{d-1} \end{array} \begin{array}{c} \begin{array}{ccccc} G_0 & G_1 & G_2 & \ldots & G_{d-1} \end{array} \\ \begin{bmatrix} 0 & P_1 & 0 & \ldots & 0 \\ 0 & 0 & P_2 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_0 & 0 & 0 & \ldots & 0 \end{bmatrix} \end{array}$$

  for some matrices $P_0, \ldots, P_{d-1}$.
- *P_{cyc}^d is a block diagonal matrix and each diagonal block is primitive*

By definition, each of P_i is a (row) stochastic matrix and so satisfies $P_i\vec{1} = \vec{1}$ (with appropriate sizes). This implies, on

$$v = \begin{pmatrix} \vec{1} & \omega\vec{1} & \omega^2\vec{1} & \ldots & \omega^{d-1}\vec{1} \end{pmatrix}^\mathsf{T}$$

with $\omega \in \mathbb{C}$ being a dth root of unity ($\omega^d = 1$), and $\vec{1}$ vectors having appropriate sizes,

$$\begin{aligned} Pv &= \begin{pmatrix} P_1\omega\vec{1} & P_2\omega^2\vec{1} & \ldots & P_{d-1}\omega^{d-1}\vec{1} & P_0\vec{1} \end{pmatrix}^\mathsf{T} \\ &= \begin{pmatrix} \omega\vec{1} & \omega^2\vec{1} & \ldots & \omega^d\vec{1} \end{pmatrix}^\mathsf{T} \\ &= \omega v \end{aligned}$$

so ω is an eigenvalue of P.

As the diagonal blocks of P_{cyc}^d are primitive, from the result mentioned in lecture (also Tutorial 6), we can see that, if $\lambda \in \mathbb{C}$ is an eigenvalue of P with modulus 1, λ^d must be an eigenvalue with modulus 1 to each of the diagonal block of P_{cyc}^d, and so $\lambda^d = 1$.

This gives us the following result[3]:

Theorem 1.5 (Nonnegative Perron–Frobenius theorem, partial)**.** *If P is the transition matrix of a finite irreducible chain, then P has period $d \geq 1$ if and only if P has exactly d eigenvalues with unit modulus, and in such case they are exactly the dth roots of unity: $e^{2\pi i \frac{1}{d}}, \ldots, e^{2\pi i \frac{d}{d}} = 1$.*

So now you have another way to compute the period of an irreducible chain[4]: instead of finding all loops or drawing the diagram in a clever way, just find *all* eigenvalues of the transition matrix and count those with unit modulus. However, note that sometimes it *may* be easier to just count the paths than to find all eigenvalues (especially when you have many states).

---

[2] Like canonical form, it seems that there isn't a common notation for this.

[3] For more, see this note by M. Boyle.

[4] There is also an algorithmic one, just like how irreducible closed sets can be found algorithmically. See this note by Jarvis and Shier for a BFS approach (following Denardo's result) and a corresponding Python implementation.

2 With Mean Recurrence Time

Recall from the last tutorial that the following holds on an irreducible chain:

$$\frac{1}{n}N_n(y) \xrightarrow{n\to\infty} 1/m_y \quad \text{with probability 1}$$
$$\frac{1}{n}E_x\left(N_n(y)\right) \xrightarrow{n\to\infty} 1/m_y$$

Suppose the chain has period d, and $x \in S$ be a state. As the label is not important, we may assume that $x \in G_0$. Then for each $y \in G_r$, $P^n(x,y) \neq 0$ only if $n \equiv r \pmod{d}$. If we group the terms together, for $n = kd$,

$$\frac{1}{n}N_n(y) = \frac{1}{kd}\sum_{t=1}^{kd}\mathbf{1}_y(X_t) = \frac{1}{kd}\sum_{m=0}^{k-1}\mathbf{1}_y(X_{md+r}) \xrightarrow{n\to\infty} 1/m_y \quad \text{with probability 1}$$

and so

$$\frac{1}{k}\sum_{m=0}^{k-1}\mathbf{1}_y(X_{md+r}) \xrightarrow{k\to\infty} d/m_y \quad \text{with probability 1}$$
$$\frac{1}{k}\sum_{m=0}^{k-1}P^{md+r}(x,y) \xrightarrow{k\to\infty} d/m_y$$

In particular,

$$\frac{1}{k}\sum_{m=0}^{k-1}P^{md}(x,x) \xrightarrow{k\to\infty} d/m_x$$

Of course, these are obvious once you have the strong convergence $\lim_m P^{md+r}$ (see lecture notes).

Suppose for the moment the chain is positive recurrent, and so has unique stationary distribution $\pi(x) = 1/m_x$, and the transition matrix is already in cyclic block form. Let $Q_0, \ldots, Q_{d-1}$ be the diagonal blocks of P^d, so $P^d = \operatorname{diag}(Q_0, \ldots, Q_{d-1})$. As noted, each Q_i is primitive and so the corresponding d-step chain has its own (unique) stationary distribution π_i on G_i. In particular, as $x \in G_0$,

$$\frac{1}{k}\sum_{m=0}^{k-1}P^{md}(x,x) = \frac{1}{k}\sum_{m=0}^{k-1}Q_0^m(x,x) \xrightarrow{k\to\infty} 1/m_{x,Q_0} = \pi_0(x)$$

Comparing with the previous limit, we have

$$\pi_0(x) = d/m_x = d\pi(x)$$

that is,

$$\pi(x) = \pi_0(x)/d$$

As x is arbitrary,

Theorem 2.1. *The stationary distribution of the original chain with transition matrix P is formed by putting together the stationary distributions on the cyclic classes of the d-step chain (with transition matrix P^d) and normalizing by the period:*

$$\pi = (\pi_0, \pi_1, \ldots, \pi_{d-1})/d$$

and, with P_i from the cyclic block form,

$$\pi_i P_{i+1} = \pi_{i+1}$$

Of course, this can be proven by inspecting the transition matrices (instead of going through the limits). This also means that, to find the stationary distribution of a periodic chain, you just need to find the stationary distribution of the d-step chain on *one* cyclic class, which could be easier to solve than the whole stationary equation (with the additional cost of finding cyclic classes, d-step matrix, and computing the matrix multiplications).