EXAMPLE 1.5.7. (Queueing Chain)

Consider a railway ticket reservation counter in Chennai. Assume that people arrive at the counter in some random manner. Suppose one person is served in a unit of time. Administrators are usually interested in the number of people waiting in the queue at each time $n = 1, 2, \ldots$. We will now describe a simple Markov chain that models this problem. Let $\xi, \xi_1, \xi_2, \ldots$ be i.i.d random variables such that

$$P(\xi = k) = p_k, \quad k = 0, 1, 2, \ldots, \tag{1.5.16}$$

with $\sum_{k=0}^{\infty} p_k = 1$ (think of ξ_i as the number of people arriving in time unit i). Let X_0 be the number of people in the queue at time 0. Then the number of people in the queue at time $n \geq 1$ can be described by $X_n = \max\{X_{n-1} - 1, 0\} + \xi_n$. It can be verified that X_n is a Markov chain on state space $S = \{0, 1, 2, \ldots\}$ with some initial distribution μ and transition matrix $\mathbf{P}$ given by

$$\begin{pmatrix} p_0 & p_1 & p_2 & p_3 & p_4 & p_5 & \cdots \\ p_0 & p_1 & p_2 & p_3 & p_4 & p_5 & \cdots \\ 0 & p_0 & p_1 & p_2 & p_3 & p_4 & \cdots \\ 0 & 0 & p_0 & p_1 & p_2 & p_3 & \cdots \\ 0 & 0 & 0 & p_0 & p_1 & p_2 & \cdots \\ \vdots & & & \cdots & & \cdots & \cdots \end{pmatrix}.$$

It is not hard to see that the chain is irreducible if

$$0 < p_0 < 1 \text{ and if there exists } k > 1 \text{ such that } p_k > 0.$$

A natural question would be to decide whether the chain is recurrent or transient from an application point of view as well as a theoretical one.

The next result is a very useful technique for establishing recurrence for Markov chains and will help us answer the above question.

THEOREM 1.5.8. *(Lyapunov Function) Let X_n be an irreducible Markov chain on $S = \{0\} \cup \mathbb{N}$, with transition matrix $\mathbf{P}$. Then*

(a) the chain is transient if and only if there is a bounded non-constant function $f : S \to \mathbb{R}$ such that

$$\sum_{j=0}^{\infty} p_{ij} f(j) = f(i), \quad \text{for all } i \neq 0. \tag{1.5.17}$$

(b) the chain is recurrent if there is a function $f : S \to \mathbb{R}$ such that

$$\lim_{i\to\infty} f(i) = \infty \tag{1.5.18}$$

and

$$\sum_{j=0}^{\infty} p_{ij} f(j) \le f(i), \quad \textit{for all } i \neq 0. \tag{1.5.19}$$

We will defer the proof of the result and assuming it we will resolve the question for the example. Namely we will try to find Lyapunov functions f satisfying (1.5.17) or (1.5.18) and (1.5.19) for our example. We will do it for both the transient and recurrent cases (though one would be enough).

Transience: Define $f : S \to \mathbb{R}$ by $f(i) = a^i$ for some real number $0 < a < 1$. We need to show $\sum_{j=0}^{\infty} p_{ij} f(j) = f(i)$ for $i \neq 0$, so we to find $0 < a < 1$ such that

$$\sum_{j=i-1}^{\infty} p_{j-i+1} a^j = a^i.$$

In other words, dividing both sides of above by a^{i-1}, we need to find a such that

$$a = \sum_{j=i-1}^{\infty} p_{j-i+1} a^{j-i+1}, \; i \neq 0$$

or

$$a = \sum_{k=0}^{\infty} p_k a^k$$

. Let $g(a) = \sum_{k=0}^{\infty} p_k a^k$. Now $g(0) = p_0 > 0$, $g(1) = 1$ and $g'(1) = \sum_{k=0}^{\infty} k p_k$. The expression for $g'(1)$ we define as μ. Note that it is a potentially divergent series, but all terms of the series are positive, so $\mu > 0$ if it represnts a real number.

If $\mu > 1$ then there must exist $0 < a_0 < 1$ such that $g(a_0) = a_0$. (*Reason:* The function $g(x) - x$ is positive at $x = 0$, vanishes at $x = 1$ and is strictly increasing at $x = 1$ and so negative just to the left of $x = 1$.) So, the function $f(i) = a_0^i$ satisfies (1.5.17), thereby establishing transience of the chain if $\mu > 1$.

Recurrence: Define $f : S \to \mathbb{R}$ by $f(i) = i$. Clearly f satisfies (1.5.18).

$$\begin{aligned}\sum_{j=0}^{\infty} p_{ij} f(j) &= \sum_{j=i-1}^{\infty} p_{j-i+1} j \\ &= \sum_{k=0}^{\infty} p_k (k+i-1) \\ &= \mu + i - 1\end{aligned}$$

So f will satisfy (1.5.19) if $\mu \leq 1$ and the chain will be recurrent *In conclusion, the chain will be transient if* $\mu > 1$ *and recurrent otherwise.*

■

Proof of Theorem 1.5.8 : (Not covered in class but is within the scope of the syllabus and is from [AS09].) Consider another Markov chain Y_n on S with transition matrix Q given by

$$q_{ij} = \begin{cases} 1 & \text{if } i = j = 0 \\ 0 & \text{if } i = 0, j \neq 0 \\ p_{ij} & \text{otherwise} \end{cases}$$

Let us write T_i^Y for the hitting time of i for the chain Y (just as T_i was the hitting time of i for the chain X). Fix $i \neq 0, j \neq 0$.

(a) Let us assume that the chain X_n is transient. Define $f : \mathbb{N} \to \mathbb{R}$ by

$$f(i) = \mathbb{P}(T_0^Y < \infty | Y_0 = i).$$

Using Proposition 1.4.5 we have

$$f(i) = \sum_{j=0}^{\infty} q_{ij} f(j). \tag{1.5.20}$$

As X_n is transient, this implies that there exists $k \neq 0$ such that $\mathbb{P}(T_0 < \infty \mid X_0 = k) < 1$. Then, by the definition of Q,

$$\mathbb{P}(T_0^Y < \infty \mid Y_0 = k) = \mathbb{P}(T_0 < \infty \mid X_0 = k) \tag{1.5.21}$$

Therefore, $f(k) < 1$ and $f(0) = 1$, we see that f is non-constant. The function f is clearly bounded.

For the converse, we proceed by first assuming the validity of equations (1.5.22) and (1.5.23) below (about the chain Y_n), whose justification we postpone to the end of the proof of the theorem.

$$\lim_{n\to\infty} q_{i0}^{[n]} = \mathbb{P}(T_0 < \infty \mid X_0 = i), \tag{1.5.22}$$

$$\lim_{n\to\infty} q_{ij}^{[n]} = 0 \text{ for all } j \neq 0. \tag{1.5.23}$$

Let us suppose if possible that there is a bounded non-constant solution f of (1.5.17) while the chain (X_n, P) is recurrent. In view of Proposition 1.5.4 and (1.5.22), we have

$$\lim_{n\to\infty} q_{i0}^{[n]} = 1 \tag{1.5.24}$$

For $i \neq 0$,

$$f(i) = \sum_{j=0}^{\infty} q_{ij} f(j).$$

For $i = 0$ the above equality trivially holds. So iterating n times and using (1.1.9), we will have

$$f(i) = \sum_{j=0}^{\infty} q_{ij}^{[n]} f(j), \text{ for all } i \in S, n \in \mathbb{N}. \tag{1.5.25}$$

Fix $i \in S$. We can find random variables Z_n on some probability space with $\mathbb{P}(Z_n = j) = q_{ij}^{[n]}$ for all $j \in S, n \in \mathbb{N}$. Let $Z \equiv 0$ be the trivial random variable. Then observe that for all $x \in \mathbb{R}$ we have

$$\begin{aligned} P(Z_n \leq x) &= \sum_{S \ni j \leq x} q_{ij}^{[n]} \\ &\to 1_{[0,\infty)}(x) \\ &= P(Z \leq x) \ ; \end{aligned}$$

where we have used (1.5.24), (1.5.23) and the finiteness of the set $\{S \ni j \leq x\}$ in the second step.

Thus $Z_n \xrightarrow{d} Z$. So we conclude from (1.5.25) that

$$f(i) = \lim_{n\to\infty} \sum_{j=0}^{\infty} q_{ij}^{[n]} f(j) = \lim_{n\to\infty} \mathbb{E} f(Z_n) = \mathbb{E} f(Z) = f(0).$$

So f is a constant function which is a contradiction. Hence (X_n, P) is transient.

(b) Suppose there is an $f : S \to \mathbb{R}$ satisfying (1.5.18) and (1.5.19). Let $\epsilon > 0$ be given. As $\lim_{i\to\infty} f(i) = \infty$ there exists an m such

$$f(i) > \frac{1}{\epsilon} \quad \text{for all } i \geq m. \tag{1.5.26}$$

Pick a constant C sufficiently large so that $C + f(i) > 0$ for all $i < m$. Notice that $f_1 = f + C$ is strictly positive and also satisfies (1.5.18) and (1.5.19). So we may simply assume that f is strictly positive. The inequality

$$\sum_{j=0}^{\infty} q_{ij} f(j) \leq f(i),$$

is valid for $i \neq 0$ by definition of Q and assumption on f and it is trivally valid for $i = 0$. Iterating this inequality n times and using (1.1.9)we obtain

$$\sum_{j=0}^{\infty} q_{ij}^{[n]} f(j) \leq f(i), \qquad \text{for all } i \in S, n \in \mathbb{N}. \tag{1.5.27}$$

Using (1.5.26) and (1.5.27) we have for all $i \in S, n \in \mathbb{N}$,

$$f(i) \geq \sum_{j=0}^{m-1} q_{ij}^{[n]} f(j) + \frac{1}{\epsilon}\Big(\sum_{j=m}^{\infty} q_{ij}^{[n]}\Big) = \sum_{j=0}^{m-1} q_{ij}^{[n]} f(j) + \frac{1}{\epsilon}\Big(1 - \sum_{j=0}^{m-1} q_{ij}^{[n]}\Big).$$

Now let $n \to \infty$ in the above. From (1.5.22) and (1.5.23), we have

$$f(i) \geq \mathbb{P}(T_0 < \infty \mid X_0 = i) f(0) + \frac{1}{\epsilon}(1 - \mathbb{P}(T_0 < \infty \mid X_0 = i)).$$

Rearranging the above we obtain

$$\mathbb{P}(T_0 < \infty \mid X_0 = i) \geq 1 - \epsilon[f(i) - \mathbb{P}(T_0 < \infty \mid X_0 = i) f(0)]$$

Since this inequality holds for all $\epsilon > 0$, we have $\mathbb{P}(T_0 < \infty \mid X_0 = i) \geq 1$ which means $\mathbb{P}(T_0 < \infty \mid X_0 = i) = 1$. Thus we may conclude that (X_n, P) is recurrent.

Finally, we justify the limits involving the transition matrix Q used above.

Justification of (1.5.22): First, we have

$$\begin{aligned}\lim_{n\to\infty} q_{i0}^{[n]} &= \lim_{n\to\infty}\sum_{k=1}^{n}\mathbb{P}(T_0^Y = k \mid Y_0 = i)q_{00}^{n-k}\\ &= \lim_{n\to\infty}\sum_{k=1}^{n}\mathbb{P}(T_0^Y = k \mid Y_0 = i)\\ &= \lim_{n\to\infty}\sum_{k=1}^{n}\mathbb{P}(T_0 = k \mid Y_0 = i) \qquad (1.5.28)\\ &= \mathbb{P}(T_0 < \infty \mid X_0 = i)\end{aligned}$$

Justification of (1.5.23): Since P is irreducible there is an s such that $p_{j0}^{[s]} > 0$ and so by definition of Q, $q_{j0}^{[s]} > 0$ as well. Together with the fact that $q_{0j} = 0$ this implies

$$P_j(T_j^Y < \infty) \le 1 - q_{j0}^s < 1.$$

So,

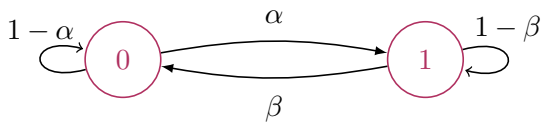
$$\begin{aligned}\sum_{n=1}^{\infty} q_{ij}^{[n]} &= \sum_{n=1}^{\infty}\mathbb{P}(Y_n = j \mid Y_0 = i)\\ &= \sum_{m=0}^{\infty}\mathbb{P}(N^Y(j) \ge m \mid Y_0 = i)\\ &= \sum_{m=0}^{\infty}\mathbb{P}(T_j^Y < \infty \mid Y_0 = i)(\mathbb{P}(T_j^Y < \infty \mid Y_0 = j))^{m-1}\\ &= \mathbb{P}(T_j^Y < \infty \mid Y_0 = i)\sum_{m=0}^{\infty}(\mathbb{P}(T_j^Y < \infty \mid Y_0 = j)^{m-1}.\end{aligned}$$

The geometric infinite series converges since $\mathbb{P}(T_j^Y < \infty \mid Y_0 = j) < 1$. Therefore $\sum_{n=1}^{\infty} q_{ij}^{[n]}$ converges which implies (1.5.23).

■

1.6 EXISTENCE AND CONVERGENCE TO STATIONARITY

Recall from Example 1.1.4 were we considered a Markov chain X_n on $S = \{0, 1\}$ with $P(X_0 = 0) = \mu_0 \in (0, 1), P(X_0 = 1) = \mu_1 \in (0, 1)$, and whose dynamics is given by directed graph with S as the vertex set.



with $0 < \alpha, \beta < 1$. We saw that

$$\sup_{i \in \{0,1\}} \mid \mathbb{P}(X_n = i) - \pi(\{i\}) \mid \leq C(1-\alpha-\beta)^n,$$

with the stationary distribution π given by $\pi(\{0\}) = \frac{\beta}{\alpha+\beta}$ and $\pi(\{1\}) = \frac{\alpha}{\alpha+\beta}$. While not every Markov chain may have a stationary distribution, it is nevertheless true that every irreducible Markov chain on a finite state space S has a unique stationary distribution. Our treatment is borrowed from [FW, Lemma 3.1 Page 158] and [LPW, Lemma 1.16, Theorem 4.9].

THEOREM 1.6.1. *Let X_n be an irreducible Markov chain on a finite state space S. Then X_n has a unique stationary distribution.*

Proof : We shall first show existence of a stationary distribution. For $i \in S$, a directed spanning tree on S with root at i is by definition a graph g which satisfies the following conditions:

(a) g has verices indexed by S,

(b) the edges of g are directed in such a way that for any $j \in S$, there is a unique directed path from j to i,

(c) $i \in S$ does not have outgoing edge and $j \in S \setminus \{i\}$ has one outgoing edge.

Let $G(i)$ be the set of all directed spanning trees on S with root at i. For every $i \in S$, define

$$\pi_0(i) = \sum_{g \in G(i)} \prod_{(j \to k) \in g} p_{jk}, \tag{1.6.1}$$

where $j \to k$ denotes any directed edge in the tree g. As X_n is irreducible each $\pi_0(i) > 0$. We shall first verify that

$$\pi_0(i) = \sum_{j \in S} \pi_0(j) p_{ji}, \forall i \in S,$$

which is equivalent to showing that

$$\pi_0(i) \sum_{j \neq i} p_{ij} = \sum_{j \neq i} \pi_0(j) p_{ji} \tag{1.6.2}$$

or that

$$(\sum_{g\in G(i)}\prod_{(j\to k)\in g}p_{jk})\sum_{j\neq i}p_{ij}=\sum_{j\neq i}(\sum_{g\in G(j)}\prod_{(k\to j)\in g}p_{kj})p_{ji} \tag{1.6.3}$$

The equality follows from the following observation, which can be easily verified. Both sides of (1.6.3) are equal to $\sum_{g\in L}\prod_{(j\to k)\in g}p_{jk}$, where L is the set of all graphs satisfying:

(a) every point $k\in S$ has only one directed edge pointing way from k to l, $(l\neq k, l\in S)$

(b) the graph has exactly one closed cycle and $i\in S$ belongs to that cycle.

So the normalised version of π_0 - defined by $\pi(i)=\left(\sum_{j\in S}\pi_0(j)\right)^{-1}\pi_0(i)$ - will be a stationary distribution for X_n.

Now we shall prove uniqueness. As S is finite, let $\pi=[\pi_1(1)-\pi_2(1),\pi_1(2)-\pi_2(2),\ldots]$ be thought of as a row vector. As π is stationary we have

$$\pi\mathbf{P}=\pi \qquad \text{equivalently} \qquad \pi(\mathbf{P}-I)=0 \qquad \text{equivalently} \qquad (\mathbf{P}^T-I)\pi^T=0$$

We will show that the null space of $P-I$ has dimension 1 which will imply that the null space of $\mathbf{P}^T-I$ is also one dimension which implies that there is only one probability vector π that satisfies $(\mathbf{P}^T-I)\pi^T=0$.

Let $h:S\to\mathbb{R}$ be a solution to $\mathbf{P}h=h$, with h being thought of as a column vector. Set $M=\max\{h(i):i\in S\}$ and $h(i_0)=M$ for some $i_0\in S$. Suppose $p_{i_oj}>0$ and $h(j)<M$ then

$$h(i_o)=h(j)p_{ji_0}+\sum_{k\in S,k\neq j}h(k)p_{kj_0}<M(p_{i_oj}+\sum_{k\in S,k\neq j}p_{ki_0})=M.$$

This is a contradiction. Therefore $h(j)=M$. By irreducibility it follows that $h(k)=M$ for any $k\in M$. Hence the

$$\text{null space of } P-I=\left\{h:h=\alpha\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix},\alpha\in\mathbb{R}\right\}$$

and has dimension 1. ■

We will now establish a rate of convergence to stationary distribution for an irreducible aperiodic Markov chain on a finite state space.

THEOREM 1.6.2. *Let X_n be a Markov chain on S with transition matrix $\mathbf{P}$ and initial distribution μ. Let X_n be irreducible and aperiodic Markov chain with $|S|<\infty$. Let π be*

its unique stationary distribution. Then there exist $C > 0$ *and* $0 < \theta < 1$ *such that for all* $n \geq 1$,

$$\sup_{i,j \in S} \mid p_{ij}^n - \pi(\{j\}) \mid \leq \mid S \mid \cdot C \cdot \theta^n \tag{1.6.4}$$

Proof : Now without loss of generality let $S = \{1, 2, \ldots, N\}$ for some $N \geq 1$. We shall first show that there exists $r \geq 1$ such

$$p_{ij}^r > 0 \qquad \text{for all } i, j \in S. \tag{1.6.5}$$

Since the chain is aperiodic, for each $i \in S$ there is $n_i \geq 1$ such that $p_{ii}^n > 0$ for all $n \geq n_i$. For each $j \in S$ and $j \neq i$ by irreducibility there is a n_j such that $p_{ij}^{n_j} > 0$. Let $r = \max\{n_i : i \in S\} + \max\{n_j : j \in S\}$. Then for any $i, j \in S$,

$$p_{ij}^r \geq p_{ii}^{r-n_j} p_{ij}^{n_j} > 0,$$

which establishes (1.6.5). Let $0 < \delta < 1$ be such that

$$p_{ij}^r \geq \delta \pi(\{i\}) > 0,$$

for all $i, j \in S$. Now without loss of generality let $S = \{1, 2, \ldots, N\}$ for some $N \geq 1$. Define the matrix $\Pi_{N \times N}$ by

$$\Pi = \begin{pmatrix} \pi(\{1\}) & \pi(\{2\}) & \ldots & \pi(\{N\}) \\ \pi(\{1\}) & \pi(\{2\}) & \ldots & \pi(\{N\}) \\ \pi(\{1\}) & \pi(\{2\}) & \ldots & \pi(\{N\}) \\ \vdots & \vdots & \vdots & \vdots \\ \pi(\{1\}) & \pi(\{2\}) & \ldots & \pi(\{N\}) \end{pmatrix}.$$

Let $\theta := 1 - \delta$ and define the matrix $Q_{N \times N}$ by

$$Q := \frac{1}{\theta}(P^r + (1 - \theta)\Pi).$$

By construction $Q = (q_{ij})_{1 \leq i,j \leq N}$ is such that $0 \leq q_{ij} \leq 1$ and $\sum_{j=1}^N q_{ij} = 1$ for all $i \in S$. Rearranging the above, using the fact that $Q^k \Pi = \Pi$ and$\Pi P^r = \Pi$ along with induction it is easy to see that for all $k \geq 1$,

$$P^{rk} = \theta^k Q^k + (1 - \theta^k)\Pi.$$

Let $n \geq 1$. Then there is $k \geq 1$ and $0 \leq l \leq r-1$ such that $n = rk + l$. We multiply the above by P^l to get

$$\begin{aligned} P^{rk+l} &= \theta^k Q^k P^l + (1-\theta^k)\Pi P^l \\ &= \theta^k Q^k P^l + (1-\theta^k)\Pi, \end{aligned}$$

as π is as stationary distribution. Therefore, by rearranging we have

$$P^{rk+l} = \theta^k (Q^k P^l - \Pi).$$

If we consider each in the matrix equality above and take absolute values we have that for each $i \in S$ we have

$$p_{ij}^{rk+j} - \pi(\{j\}) = \theta^k((Q^k P^l)_{ij} - \pi(\{j\})), \qquad \text{for all } j \in S.$$

Consequently, we have

$$\sup_{i,j \in S} \mid p_{ij}^n - \pi(\{j\}) \mid = \theta^k \mid (Q^k P^l)_{ij} - \pi(\{j\}) \mid \leq C\theta^n,$$

for some $C > 0$. ∎

Now observe that, for any $j \in S$ and $n \geq$,

$$\begin{aligned} \mid \mathbb{P}(X_n = j) - \pi(\{j\}) \mid &= \mid \sum_{i \in S} p_{ij}^n \mathbb{P}(X_0 = i) - \pi(\{j\}) \mid \\ &= \sum_{i \in S} \mid p_{ij}^n - \pi(\{j\}) \mid \mathbb{P}(X_0 = i) \\ &\leq \mid S \mid \sup_{i,j \in S} \mid p_{ij}^n - \pi(\{j\}) \mid \end{aligned}$$

Therefore using Theorem 1.6.2 we have for all $n \geq 1$ there is a $C > 0$ and $0 < \theta < 1$ such that

$$\sup_{j \in S} \mid \mathbb{P}(X_n = j) - \pi(\{j\}) \mid \leq \mid S \mid \cdot C \cdot \theta^n$$