Classification of Markov chains

We say that the states i and j *intercommunicate* if there are integers n, m such that $p_{ij}(n) > 0$ and $p_{ji}(m) > 0$. In other words it is possible to go from each state to the other after a finite number of steps.

Theorem 1 *Suppose i, j intercommunicate, then they are either both transient or both persistent.*

Proof: Since i, j intercommunicate there are integers n, m such that

$$h = p_{ij}(n)p_{ji}(m) > 0 \tag{10}$$

Hence for any r,

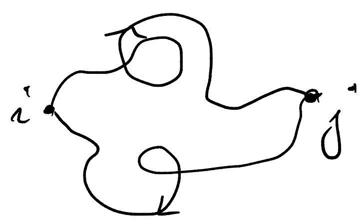
$$p_{ii}(n+m+r) \geq p_{ij}(n)p_{jj}(r)p_{ji}(m) = h\, p_{jj}(r) \tag{11}$$

Sum over r to deduce

$$\sum_k p_{ii}(k) \geq \sum_r p_{ii}(n+m+r) \geq h \sum_r p_{jj}(r) \tag{12}$$

Therefore either both sums are finite or both are infinite, hence either both states are transient or both are persistent.

If all states intercommunicate we say that the Markov chain is *irreducible*.

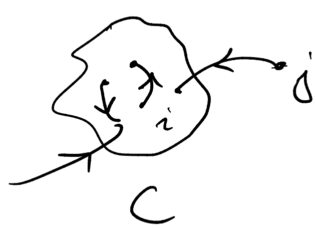


$p_{ij}(n) > 0$ some $n \geq 1$

$p_{ji}(m) > 0$ some $m \geq 1$.

$\Rightarrow$ either both persistent or both transient.

A set of states C is <u>closed</u> if



$i \in C$, $j \in C^c$

$\Rightarrow p_{ij} = 0$.

A set of states C is <u>irreducible</u> if all states intercommunicate in C.

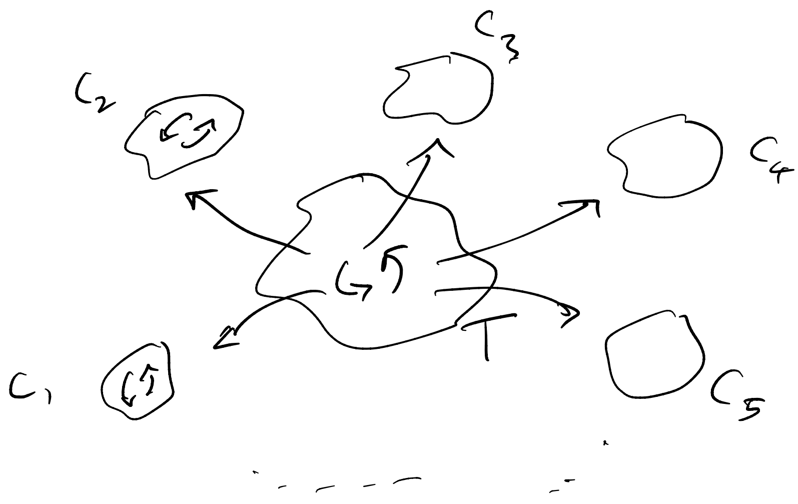
A class of states C in Ω is called *closed* if $p_{ij} = 0$ whenever $i \in C$ and $j \notin C$.

There is a fairly obvious decomposition of the state space.
The state space Ω can be partitioned uniquely as

$$\Omega = T \cup C_1 \cup C_2 \cup \cdots \tag{13}$$

where T is the set of all transient states, and each class C_i is closed and irreducible.

If the chain starts with $X_0 \in C_i$ then it stays in C_i forever. If it starts with $X_0 \in T$ then eventually it enters one of the classes C_i and stays there forever. So the irreducible classes determine the *long-time behavior.*



Lemma. Every finite chain has at least one persistent state.

Proof.
1) Pigeonhole principle: chain must visit some state infinitely often. This state must be persistent.

2) some state i: $\sum_j p_{ij} = 1$

$$\Rightarrow \sum_j p_{ij}(n) = 1.$$

$$\Rightarrow \sum_{n=1}^{\infty} \sum_{j} p_{ij}(n) = \infty$$

$$\Rightarrow \sum_{j} \sum_{n=1}^{\infty} p_{ij}(n) = \infty.$$

$$\Rightarrow \text{some } j \quad \sum_{n=1}^{\infty} p_{ij}(n) = \infty$$

recall $\hat{P}_{ij}(1) = \hat{f}_{ij}(1)\, \hat{P}_{jj}(1) \quad (i \neq j).$

$$\infty = \hat{f}_{ij}(1)\, \hat{P}_{jj}(1)$$

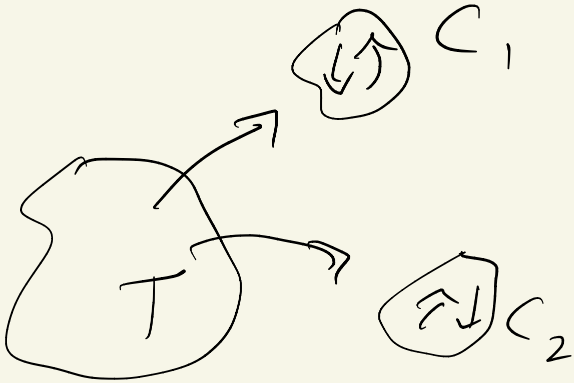
between 0 and 1.

$$\Rightarrow \hat{P}_{jj}(1) = \infty$$

$$\Rightarrow \sum_{n=1}^{\infty} p_{jj}(n) = \infty$$

$\Rightarrow j$ is persistent.

If chain is irreducible and finite, then all states are persistent.



C is closed, irreducible class.

Then C by itself is an irreducible Markov chain.

$\Rightarrow$ if C is finite, then all states in C are persistent.

Stationary distribution

Let $\{\pi_i\}$ $(i \in \Omega)$ be a probability distribution on Ω. We say that π_i is *stationary* if

$$\boxed{\sum_i \pi_i \, p_{ij} = \pi_j} \qquad \text{for each state } j.$$

for all states $j \in \Omega$.

Lemma 1 *Suppose that $P(X_0 = j) = \pi_j$ for all $j \in \Omega$, where π is the stationary distribution. Then $P(X_n = j) = \pi_j$ for all $j \in \Omega$, and all $n \geq 1$.*

The proof is easy: note that

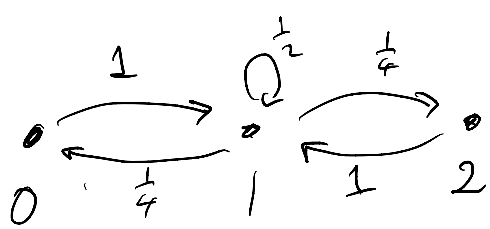
$$\begin{aligned} P(X_1 = j) &= \sum_i P(X_1 = j \,|\, X_0 = i)\, P(X_0 = i) \\ &= \sum_i p_{ij}\, \pi_i \\ &= \pi_j \end{aligned}$$

where we used the stationarity property in the last line.

Now repeat the argument for X_2 by conditioning on X_1, and so on.

This result explains why it is called a stationary distribution: once the chain enters this distribution it must remain there.

Ex Balls in boxes



uniform dist. on Ω

$\pi_0 = \frac{1}{3}$

$\pi_1 = \frac{1}{3}$

$\pi_2 = \frac{1}{3}$

Stationary equations:

$j=0$: $\pi_0 P_{00} + \pi_1 P_{10} + \pi_2 P_{20} = \pi_0$

$j=1$: $\pi_0 P_{01} + \pi_1 P_{11} + \pi_2 P_{21} = \pi_1$

$j=2$: $\pi_0 P_{02} + \pi_1 P_{12} + \pi_2 P_{22} = \pi_2$

$\pi_0 + \pi_1 + \pi_2 = 1$

Substitute for P_{ij}:

$j=0$: $\frac{1}{4}\pi_1 = \pi_0$

$j=1$: $\pi_0 + \frac{1}{2}\pi_1 + \pi_2 = \pi_1$

$j=2$: $\frac{1}{4}\pi_1 = \pi_2$

$\pi_0 = \pi_2 = \frac{1}{4}\pi_1$

$\pi_0 + \pi_1 + \pi_2 = 1$.

$\frac{1}{4}\pi_1 + \pi_1 + \frac{1}{4}\pi_1 = 1 \Rightarrow \pi_1 = \frac{2}{3}$

$\pi_0 = \pi_2 = \frac{1}{6}$

$\Rightarrow$ Stat. dist is $(\pi_0, \pi_1, \pi_2) = \left(\frac{1}{6}, \frac{2}{3}, \frac{1}{6}\right)$.

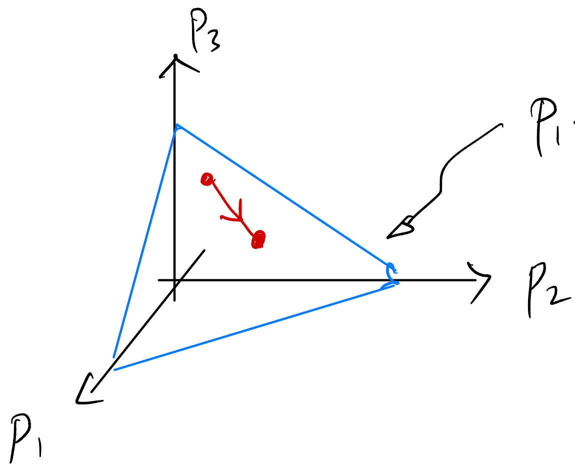
So if we regard the Markov chain as a map on probability distributions, via the mapping

$$p_j \to \sum_i p_i \, p_{ij}$$

then the stationary distribution is a fixed point for this map. Several immediate questions:

- is there always a stationary distribution?
- if it exists, is the stationary distribution unique?
- does the chain always converge to the stationary distribution?

In general the answer is 'no', but under certain assumptions the answer is yes.



$p_1 + p_2 + p_3 = 1$

- probability dist. for $\Omega = \{0, 1, 2\}$.

If $\mathbb{P}(X_0 = i) = p_i$ all i

then $\mathbb{P}(X_1 = j) = \sum_i p_i \, p_{ij}$

$\boxed{\textbf{Irreducible and regular}}$

Notation: for a matrix T write $T \geq 0$ if $T_{ij} \geq 0$ for all i, j and $T > 0$ if $T_{ij} > 0$ for all i, j.

Definition 2 *Let P be the transition matrix of a Markov chain.*

(1) The Markov chain is irreducible *if for all states i, j there is an integer $n(i, j)$ such that $p_{ij}(n(i, j)) > 0$.*

(2) The Markov chain is regular *if there is an integer n such that $P^n > 0$.*

Example 4 *Recall the balls in boxes model:*

$$\begin{array}{cc} P = \begin{pmatrix} 0 & 1 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1 & 0 \end{pmatrix} & \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \end{array} \tag{14}$$

Since

$$P^2 = \begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 1/8 & 3/4 & 1/8 \\ 1/4 & 1/2 & 1/4 \end{pmatrix} \tag{15}$$

it follows that P is regular.

<u>Irreducible</u>: for every i,j there is $n \geq 1$ such that $p_{ij}(n) > 0$.

<u>Regular</u>. There is $n \geq 1$ such that for every i,j, $p_{ij}(n) > 0$

Regular $\Rightarrow$ Irreducible.

<u>But</u> converse not true.

b/c counterexample

$$P = \cancel{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}$$

not irreducible

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Check: P is irred.? ✓

P not regular? $P^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$$P^3 = P$$

$$P^4 = I$$

$$\vdots$$

$$P^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$P^{2n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

always zeros in matrix

$\Rightarrow$ not regular.

Note: sufficient to check if P^d is strictly positive where d is dimension of P.

For irreducible, check if

$P + P^2 + P^3 + \cdots + P^d$ is strictly positive.

Example 5 *Define the two-state swapping chain:*

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{16}$$

Then $P^2 = I$ *is the identity, hence for all* $n \geq 1$

$$P^{2n} = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P^{2n+1} = P \tag{17}$$

So P *is irreducible but not regular.*

0.1 Perron-Frobenius Theorem

Let e denote the vector in $\mathbb{R}^n$ with all entries 1, so

$$e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix}^T \tag{18}$$

Theorem 2 [Perron-Frobenius] *Suppose P is a regular $n \times n$ transition matrix. Then there is a unique strictly positive vector $w \in \mathbb{R}^n$ such that*

$$w^T P = w^T \tag{19}$$

$\leftarrow$ stationary equ. for w.

and such that

$$P^k \to e\,w^T \quad \text{as } k \to \infty \tag{20}$$

$$\underbrace{w^T P = w^T}$$

$w^T = (w_1, w_2, \ldots, w_n)$ prob. dist. on Ω.

$$(w_1 \; w_2 \; \cdots \; w_n) \begin{pmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \cdots \\ \vdots & & \end{pmatrix} = (w_1 \; w_2 \; \cdots \; w_n)$$

$$w_1 p_{11} + w_2 p_{21} + \cdots + w_n p_{n1} = w_1$$

$$\Rightarrow \sum_i w_i p_{i1} = w_1$$

stationary equation.

$$e w^T = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} (w_1 \; w_2 \; \cdots \; w_n) = \begin{pmatrix} w_1 & w_2 & \cdots & w_n \\ w_1 & w_2 & \cdots & w_n \\ \vdots & & & \\ w_1 & w_2 & \cdots & w_n \end{pmatrix}$$

$(n \times 1)$ $(1 \times n)$ $\qquad$ $n \times n$

$$\lim_{k \to \infty} P^k = e w^T$$

(i,j) entry: $(P^k)_{ij} = P_{ij}^{(k)}$

$$= \mathbb{P}(X_k = j \mid X_0 = i)$$

$$\Rightarrow \lim_{k \to \infty} \mathbb{P}(X_k = j \mid X_0 = i) = w_j.$$

$\Rightarrow$ initial state $X_0 = i$ is irrelevant

the Markov chain "forgets" about the initial state.

so $\mathbb{P}(X_k = j)$ converges to the stationary distribution w_k.

So the stationary distribution is the long-run probability for the chain to be in each state.

<u>Example</u> Balls in boxes.

$$P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 \end{pmatrix}$$

$$P^{30} = \begin{pmatrix} \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{pmatrix}$$

Long-run fraction of time the chain spends in state 0 is

$w_0 = \frac{1}{6}$

Given chain started in state 0, what is the prob. that it is in state 2 after

- 1 step Ans 0
- 2 steps Ans. $\frac{1}{4}$
- 30 steps Ans $\frac{1}{6}$
- 31 steps Ans. $\frac{1}{6}$

Proof of Perron-Frobenius: We show that for all vectors $y \in \mathbb{R}^n$,

$$P^k y \to e\, w^T y \tag{21}$$

which is a positive multiple of the constant vector e. This implies the result.

Suppose first that $P > 0$ so that $p_{ij} > 0$ for all $i, j \in S$. Let $d > 0$ be the smallest entry in P (so $d \le 1/2$). For any $y \in \mathbb{R}^n$ define

$$m_0 = \min_j\, y_j, \quad M_0 = \max_j\, y_j \tag{22}$$

and

$$m_1 = \min_j\, (Py)_j, \quad M_1 = \max_j\, (Py)_j \tag{23}$$

Consider $(Py)_i = \sum_j p_{ij} y_j$. This is maximized by pairing the smallest entry m_0 of y with the smallest entry d of p_{ij}, and then taking all other entries of y to be M_0. For any i,

$$\begin{aligned}
(Py)_i &= \sum_j p_{ij} y_j \\
&= p_{ik} m_0 + \sum_{j \ne k} p_{ij} y_j && (24) \\
&\le p_{ik} m_0 + (1 - p_{ik}) M_0 && (25) \\
&\le d m_0 + (1 - d) M_0 && (26)
\end{aligned}$$

and therefore

$$M_1 \le d m_0 + (1 - d) M_0$$

By similar reasoning,

$$m_1 = \min_i (Py)_i \ge (1 - d) m_0 + d M_0 \tag{27}$$

Subtracting these bounds gives

$$M_1 - m_1 \le (1 - 2d)(M_0 - m_0) \tag{28}$$

Now we iterate to give

$$M_k - m_k \le (1 - 2d)^k\, (M_0 - m_0) \tag{29}$$

where again

$$M_k = \max_i (P^k y)_i, \quad m_k = \min_i (P^k y)_i \tag{30}$$

Furthermore the sequence $\{M_k\}$ is decreasing since

$$M_{k+1} = \max_i (PP^k y)_i = \max_i \sum_j p_{ij} (P^k y)_j \leq M_k \tag{31}$$

and the sequence $\{m_k\}$ is increasing for similar reasons. Therefore both sequences converge as $k \to \infty$, and the difference between them also converges to zero. Hence we conclude that the components of the vector $P^k y$ converge to a constant value, meaning that

$$P^k y \to m\, e \tag{32}$$

for some m.

We can pick out the value of m with the inner product

$$m(e^T e) = e^T \lim_{k\to\infty} P^k y = \lim_{k\to\infty} e^T\, P^k y \tag{33}$$

Note that for $k \geq 1$,

$$e^T\, P^k y \geq m_k(e^T e) \geq m_1(e^T e) = \min_i (Py)_i (e^T e)$$

Since P is assumed positive, if $y_i \geq 0$ for all i it follows that $(Py)_i > 0$ for all i, and hence $m > 0$.

Now define

$$w_j = \lim_{k \to \infty} P^k e_j / (e^T e) \tag{34}$$

where e_j is the vector with entry 1 in the j^{th} component, and zero elsewhere. It follows that $w_j > 0$ so w is strictly positive, and

$$P^k \to e w^T \tag{35}$$

By continuity this implies

$$\lim_{k \to \infty} P^k P = e w^T P \tag{36}$$

and hence $w^T P = w^T$. This proves the result in the case where $P > 0$.

Now turn to the case where P is regular. Since P is regular, there exists integer N such that

$$P^N > 0 \tag{37}$$

Hence by the previous result there is a strictly positive $w \in \mathbb{R}^n$ such that

$$P^{kN} \to ew^T \tag{38}$$

as $k \to \infty$, satisfying $w^T P^N = w^T$.

It follows that $P^{N+1} > 0$, and hence there is also a vector v such that

$$P^{k(N+1)} \to ev^T \tag{39}$$

as $k \to \infty$, and $v^T P^{N+1} = v^T$. Considering convergence along the subsequence $kN(N+1)$ it follows that $w = v$, and hence

$$w^T P^{N+1} = v^T P^{N+1} = v^T = w^T = w^T P^N \tag{40}$$

and so

$$w^T P = w^T \tag{41}$$

The subsequence $P^{kN}y$ converges to $ew^T y$ for every y, and we want to show that the full sequence $P^m y$ does the same. For any $\epsilon > 0$ there is $K < \infty$ such that for all $k \geq K$ and all probability vectors y

$$\|(P^{kN} - ew^T)y\| \leq \epsilon \tag{42}$$

Let $m = kN + j$ where $j < N$, then for any probability vector y

$$\|(P^m - ew^T)y\| = \|(P^{kN+j} - ew^T)y\| = \|(P^{kN} - ew^T)P^j y\| \leq \epsilon \tag{43}$$

which proves convergence along the full sequence.

$\boxed{\text{QED}}$

Corollary of Perron-Frobenius Theorem

Note that as a corollary of the Theorem we deduce that the vector w is the unique (up to scalar multiples) solution of the equation

$$w^T P = w^T \tag{44}$$

Also since $v^T e = \sum v_i = 1$ for a probability vector v, it follows that

$$v^T P^n \to w^T \tag{45}$$

Chain forgets about initial configuration.

for any probability vector v.

Example 6 *Recall the balls in boxes model:*

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1 & 0 \end{pmatrix} \tag{46}$$

We saw that P is regular. Solving the equation $w^T P = w^T$ yields the solution

$$w^T = (1/6, 2/3, 1/6) \tag{47}$$

Furthermore we can compute

$$P^{10} = \begin{pmatrix} 0.167 & 0.666 & 0.167 \\ 0.1665 & 0.667 & 0.1665 \\ 0.167 & 0.666 & 0.167 \end{pmatrix} \tag{48}$$

showing the rate of convergence.

$\boxed{\textbf{Meaning of Perron-Frobenius}}$

Concerning the interpretation of the result. Suppose that the distribution of X_0 is

$$P(X_0 = i) = \alpha_i \tag{49}$$

for all $i \in S$. Then

$$\begin{aligned} P(X_k = j) &= \sum_i P(X_k = j | X_0 = i) \, P(X_0 = i) \\ &= \sum_i (P^k)_{ij} \, \alpha_i \\ &= (\alpha^T P^k)_j \end{aligned}$$

where α is the vector with entries α_i. Using our Theorem we deduce that

$$P(X_k = j) \to w_j \tag{50}$$

as $k \to \infty$ for any initial distribution α. Furthermore if $\alpha = w$ then $\alpha^T P^k = w^T P^k = w^T$ and therefore

$$P(X_k = j) = w_j \tag{51}$$

for all k. So w is called the *equilibrium* or *stationary* distribution of the chain. The Theorem says that the state of the chain rapidly forgets its initial distribution and converges to the stationary value.

0.2 Finite and irreducible

Now suppose the chain is irreducible but not regular. Then we get a similar but weaker result.

Theorem 3 *Let P be the transition matrix of an irreducible Markov chain. Then there is a unique strictly positive probability vector w such that*

$$w^T P = w^T \quad \leftarrow \text{stationary equation} \tag{52}$$

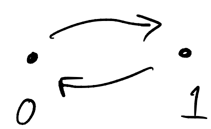
Furthermore

$$\frac{1}{n+1}\Big(I + P + P^2 + \cdots + P^n\Big) \to ew^T \rightarrow \begin{pmatrix} w_1 & w_2 & \cdots & w_n \\ w_1 & w_2 & \cdots & w_n \\ \vdots & & & \\ w_1 & w_2 & \cdots & w_n \end{pmatrix} \tag{53}$$

as $n \to \infty$.

This Theorem allows the following interpretation: for an irreducible chain, w_j is the long-run fraction of time the chain spends in state j.

Ex. $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$



Stationary eq.: $\pi_0 p_{00} + \pi_1 p_{10} = \pi_0$

$\pi_0 p_{01} + \pi_1 p_{11} = \pi_1$

$\pi_0 + \pi_1 = 1$

$$\Rightarrow \left.\begin{aligned} \pi_1 &= \pi_0 \\ \pi_0 &= \pi_1 \\ \pi_0 + \pi_1 &= 1 \end{aligned}\right\} \Rightarrow \pi_0 = \pi_1 = \frac{1}{2}.$$

$\left(\frac{1}{2}, \frac{1}{2}\right)$. Stat. dist.

$$\frac{1}{4}\left(I + P + P^2 + P^3\right) = \frac{1}{4}(2I + 2P) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \pi_0 & \pi_1 \\ \pi_0 & \pi_1 \end{pmatrix}$$

$\Rightarrow$ e.g. long-run fraction of time spent in state 0 is $\frac{1}{2}$.

$(I+P)^3 = I + 3P + 3P^2 + P^3$

Proof for finite state irreducible: define

$$Q = \frac{1}{2}I + \frac{1}{2}P \tag{54}$$

Then Q is a transition matrix. Also

$$2^n Q^n = \sum_{k=0}^{n} \binom{n}{k} \, P^k \tag{55}$$

Because the chain is irreducible, for all pairs of states i, j there is an integer $n(i,j)$ such that $(P^{n(i,j)})_{ij} > 0$. Let $n = \max n(i,j)$, then for all i, j we have

$$2^n (Q^n)_{ij} = \sum_{k=0}^{n} \binom{n}{k} \, (P^k)_{ij} \geq \binom{n}{n(i,j)} \, (P^{n(i,j)})_{ij} > 0 \tag{56}$$

by Peron-Frobenius

and hence Q is regular. Let w be the unique stationary vector for Q then

$$w^T Q = w^T \leftrightarrow w^T P = w^T \tag{57}$$

$w^T Q = w^T$

$w^T(\frac{1}{2}I + \frac{1}{2}P) = w^T$

$\frac{1}{2}w^T + \frac{1}{2}w^T P = w^T$

which shows existence and uniqueness for P.

Let $W = ew^T$ then a calculation shows that for all n

$$\Big(I + P + P^2 + \cdots + P^{n-1}\Big)(I - P + W) = I - P^n + nW \tag{58}$$

Note that $I - P + W$ is invertible: indeed if $y^T(I - P + W) = 0$ then

$$y^T - y^T P + (y^T e) w = 0 \tag{59}$$

Multiply by e on the right and use $Pe = e$ to deduce

$$y^T e - y^T P e + (y^T e)(w^T e) = (y^T e)(w^T e) = 0 \tag{60}$$

Since $w^T e = 1 > 0$ it follows that $y^T e = 0$ and so $y^T - y^T P = 0$. By uniqueness this means that y is a multiple of w, but then $y^T e = 0$ means that $y = 0$. Therefore $I - P + W$ is invertible, and so

$$I + P + P^2 + \cdots + P^{n-1} = (I - P^n + nW)(I - P + W)^{-1} \tag{61}$$

Now $WP = W = W^2$ hence

$$W(I - P + W) = W \implies W = W(I - P + W)^{-1} \tag{62}$$

Therefore

$$I + P + P^2 + \cdots + P^{n-1} = (I - P^n)(I - P + W)^{-1} + nW \tag{63}$$

and so

$$\frac{1}{n}\Big(I + P + P^2 + \cdots + P^{n-1}\Big) = W + \frac{1}{n}(I - P^n)(I - P + W)^{-1} \tag{64}$$

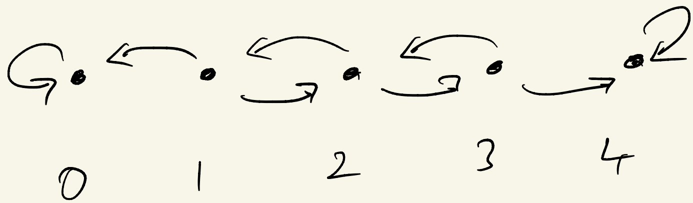
It remains to show that the norm of the matrix $(I - P^n)(I - P + W)^{-1}$ is bounded as $n \to \infty$, or equivalently that $\|(I - P^n)\|$ is uniformly bounded. This follows from the bound

$$\|P^n z\| \leq \sum_{ij} (P^n)_{ij} |z_j| = \sum_j |z_j| \tag{65}$$

Therefore $\frac{1}{n}(I - P^n)(I - P + W)^{-1} \to 0$ and the result follows,

$\boxed{\text{QED}}$

Example. Drunkard's Walk.



$$P = \begin{array}{c} \left(\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \end{array} \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array}$$

$$k \to \infty: \quad P^k \longrightarrow \left(\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 & 1-a \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 1-a & 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array}$$

<u>Row 1</u>: $\mathbb{P}(X_k = j \mid X_0 = 0)$ ✓

<u>Row 2</u>: $\mathbb{P}(X_k = j \mid X_0 = 1)$

This chain remembers the initial state.