Chapter 3

$\{X_t\}$ is a discrete time Markov Chain with TPM P

Goal: $\Pr(X_n = j)$ for some large n

① $(\alpha P^n)_j$ ⇒ numerical difficulties

② simulate ⇒ Monte Carlo error (random)

③ exact (ish) $\rightarrow \infty$

definition $\{X_t\}$ has a limiting distribution

$$\lim_{n\to\infty} (P^n)_{ij} = \lambda_j \text{ for all } i \text{ and all } j$$

This is equivalent to ① $\lim_{n\to\infty} P(X_n = j) = \lambda_j$

② $\lim_{n\to\infty} \alpha P^n \to \lambda$ for all α

③ $\lim_{n\to\infty} P^n = \Omega = \begin{pmatrix} \underset{\sim}{\lambda} \\ \vdots \\ \underset{\sim}{\lambda} \end{pmatrix}$ all rows are the same

NOT all TPM have limiting distribution

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

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

$P^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ doesn't converge

$$\Rightarrow P^n = \begin{cases} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & n \text{ odd} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & n \text{ even} \end{cases}$$

<Numerical Method>

e.g. Binary Markov Chain.

$$P = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \end{array} \quad \text{assume } 0 < p, q < 1$$

① If $p + q = 1$, then

$P = \begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix} \Rightarrow P^n = P$ for $n \geq 1$

Then $\lambda = (1-p, p)$

② If $p+q \neq 1$, consider $(P^n)_{11}$

$$(P^n)_{11} = (P^{n-1}P)_{11} \qquad \begin{pmatrix} (P^{n-1})_{11} & (P^{n-1})_{12} \\ (P^{n-1})_{21} & (P^{n-1})_{22} \end{pmatrix} \begin{pmatrix} (1-p) & p \\ q & 1-q \end{pmatrix}$$

$$= (P^{n-1})_{11}(1-p) + q(P^{n-1})_{12}$$

Note: $(P^{n-1})_{11} + (P^{n-1})_{12} = 1$

$$= (P^{n-1})_{11}(1-p) + q(1-(P^{n-1})_{11})$$

$$= q + (1-p-q)(P^{n-1})_{11}$$

$$= q + (1-p-q)[q + (1-p-q)(P^{n-2})_{11}]$$

$$= q[1+(1-p-q)] + (1-p-q)^2(P^{n-2})_{11}$$

$$\cdots \quad \cdots$$

$$= \sum_{k=0}^{n-1} q(1-p-q)^k + (1-p-q)^n \underbrace{(P^0)_{11}}_{=1}$$

identity matrix

$$= q\,\frac{(1-(1-p-q)^n)}{1-(1-p-q)} + (1-p-q)^n$$

If p and q are not both zero, nor both 1, then $|1-p-q|<1$.

$$\lim_{n\to\infty}(P^n)_{11} = q\,\frac{1}{1-(1-p-q)} = \frac{q}{p+q}$$

$$\Rightarrow \lim_{n\to\infty} P^n = \begin{pmatrix} \frac{q}{q+p} & \frac{p}{q+p} \\ \frac{q}{q+p} & \frac{p}{q+p} \end{pmatrix}$$

$$\underset{\sim}{\lambda} = \begin{pmatrix} \frac{q}{q+p} & \frac{p}{q+p} \end{pmatrix}$$

Choose the limiting distribution as initial distribution

$$(Pr(X_1=0) \;\; Pr(X_1=1)) = \alpha P = \underset{\sim}{\lambda} P = \begin{pmatrix} \frac{q}{q+p} & \frac{p}{q+p} \end{pmatrix} \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

$$Pr(X_1=0) = \frac{q}{q+p}(1-p) + \frac{p}{q+p}q = \frac{q}{q+p} = \lambda_1$$

$$Pr(X_1=1) = \frac{q}{q+p}p + \frac{p}{q+p}(1-q) = \frac{p}{q+p} = \lambda_2$$

If $\underset{\sim}{\pi} P = \underset{\sim}{\pi}$ for TPM P, then $\underset{\sim}{\pi}$ is a stationary distribution for P

Limiting distribution must be stationary.

$$\pi = \lim_{n\to\infty} \alpha P^n = (\lim_{n\to\infty} \alpha P^{n-1}) P = \pi P$$

However, $\underset{\sim}{\lambda} P = \underset{\sim}{\lambda}$ doesn't imply that $\underset{\sim}{\lambda}$ is a limiting distribution.

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

$$\underset{\sim}{\pi} P = \underset{\sim}{\pi}$$

$$(\pi_1 \quad 1-\pi_1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (\pi_1 \quad 1-\pi_1)$$

$$\Rightarrow \quad 1-\pi_1 = \pi_1$$

$$\Rightarrow \quad \pi_1 = \frac{1}{2}$$

$$\left(\tfrac{1}{2} \quad \tfrac{1}{2}\right) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \left(\tfrac{1}{2} \quad \tfrac{1}{2}\right)$$

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

The chain simply stay forever in its starting state – has no limiting distribution

However, every probability vector is a stationary distribution

$$xP = x$$

Summary

$\{X_t\}$ $t \in \{0, 1 \cdots\}$ is a Markov chain if

$$Pr(X_t = j \mid X_{t-1} = i) = Pr(X_{t-1} = j \mid X_{t-1} = i, X_{t-2} \cdots X_0)$$

It is time-homogeneous if

$$Pr(X_t = j \mid X_{t-1} = i) = Pr(X_s = j \mid X_{s-1} = i) \quad \forall s, t, i, j$$

$\{X_t\}$ is characterized by TPM P, where $P_{ij} = Pr(X_t = j \mid X_{t-1} = i)$

$$Pr(X_n = j \mid X_0 = i) = (P^n)_{ij}$$

$$Pr(X_n = j) = (\alpha P^n)_j \quad \text{where } \alpha = (P(X_0 = 1) \cdots P(X_0 = k))$$

$\{X_t\}$ or P has limiting distribution if

$$\lim_{n \to \infty} P^n \to \Omega = (\underset{\sim}{\lambda} \cdots \underset{\sim}{\lambda})^T$$

If P ($\{X_t\}$) has a limiting distribution. Then for some $\underset{\sim}{\lambda}$

$$\underset{\sim}{\lambda} P = \underset{\sim}{\lambda}$$

More generally, we say that $\underset{\sim}{\pi}$ is a stationary distribution for P

$$\underset{\sim}{\pi} P = \underset{\sim}{\pi}$$

Stationary is a necessary condition for a limiting distribution, but not sufficient

Note that if I choose λ as the initial distribution and λ is the limiting distribution for $\{X_t\}$, then $X_0, X_1 \cdots$ are identically marginally distributed.

$$Pr(X_t = i) = \lambda_i \quad \forall t \in \{0, 1 \cdots\}$$

$$Pr(X_t = i) = (\alpha P^t)_i = (\lambda P^t)_i = ((\lambda P) P^{t-1})_i = (\lambda P^{t-1})_i \cdots = \lambda_i$$

But they're not independent!

Positive TPM is a TPM where all entries are larger than 0 $(P>0)$

Now, lots of P's are not positive. However, there might be an n such that $P^n > 0$. In this case, we call P a **regular** TPM

If P is a regular TPM, then there exist a unique $\underset{\sim}{\pi}$ that is stationary distribution for P and $\underset{\sim}{\pi}$ will also be a limiting distribution.

P is possibly NOT regular if we can find P^n and P^{n+k} that have zero entries in the same place.

Finding Stationary Distribution $\underset{\sim}{\pi} P = \underset{\sim}{\pi}$

Note for a constant c, $c\underset{\sim}{\pi} P = c\underset{\sim}{\pi}$

① Let $\underset{\sim}{x} = (1, x_2 \cdots x_k)$

② Solve $\underset{\sim}{x} P = \underset{\sim}{x}$
$$\begin{cases} P_{11} + x_2 P_{21} + \cdots + x_k P_{k1} = 1 \\ \vdots \\ P_{1k} + x_2 P_{2k} + \cdots + x_k P_{kk} = x_k \end{cases}$$

③ Renormalize to find $\underset{\sim}{\pi}$
$$\underset{\sim}{\pi} = \frac{1}{1+\sum_{j=2}^{k} \tilde{x}_j} (1, \tilde{x}_2 \cdots \tilde{x}_k)$$

e.g.
$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 1-p & p \\ 1 & q & 1-q \end{array} \qquad \underset{\sim}{x} = (1, x_2)$$

$$(1-p) + x_2 q = 1 \Rightarrow x_2 = \frac{p}{q}$$
$$\underset{\sim}{x} = \frac{1}{1+p/q} \left(1, \frac{p}{q}\right) = \left(\frac{q}{q+p}, \frac{p}{q+p}\right)$$

Linear Algebra

W is a square matrix $(k \times k)$, $\underset{\sim}{v}$ is a vector $(k \times 1)$

If $W\underset{\sim}{v} = \lambda \underset{\sim}{v}$, then we say that $\underset{\sim}{v}$ is a right eigenvector of W with eigenvalue λ.

Construct a matrix of eigenvectors $\underset{\sim}{v_1}, \cdots \underset{\sim}{v_k}$ (V) where λ_j is the eigenvalue associated with $\underset{\sim}{v_j}$. Then

$$W = V \Omega V^{-1} \quad \text{where} \quad \Omega = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}$$

Eigenvalue Decomposition of W

Note $WV = (V \Omega V^{-1}) V = V \Omega$

$(W\underset{\sim}{v_1} \cdots W\underset{\sim}{v_k}) = (\lambda_1 \underset{\sim}{v_1} \cdots \lambda_1 \underset{\sim}{v_k})$

$$\underset{\sim}{\pi} P = \underset{\sim}{\pi} \Rightarrow P^T \underset{\sim}{\pi}^T = \underset{\sim}{\pi}^T$$

The right eigenvector of P^T corresponding to an eigenvalue of 1.

$$\underset{\sim}{\pi} = \frac{\underset{\sim}{v}_j}{\sum_{\ell=1}^{k} v_\ell}$$

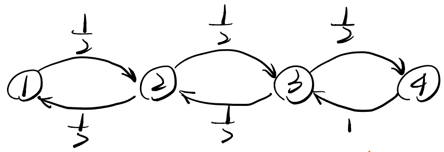
$$P = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \end{array} \begin{array}{c} \begin{array}{ccccc} a & b & c & d & e \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 2/3 & 1/3 \\ 1/2 & 0 & 1/6 & 0 & 1/3 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 3/4 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}, \quad Q = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \end{array} \begin{array}{c} \begin{array}{cccccc} a & b & c & d & e & f \end{array} \\ \begin{pmatrix} 1/6 & 1/3 & 0 & 0 & 1/2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3/4 & 1/4 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 4/5 & 0 & 0 & 1/5 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 \end{pmatrix} \end{array}.$$

$\{a, d, c, f\}, \{e\}$ $\quad$ $\{a, d, e\}, \{b\}$ and $\{c, f\}$

$$P = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \end{array} \begin{array}{c} \begin{array}{cccccc} a & b & c & d & e & f \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 0 & 1/4 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

irreducible

e.g.



$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

irreducible

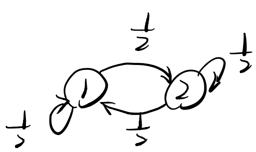
$\underset{\sim}{x} = (1, x_2, x_3, x_4)$

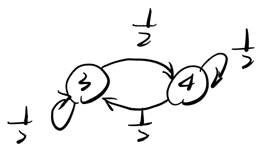
$\underset{\sim}{x} P = \underset{\sim}{x}$

$$\begin{cases} \frac{1}{2}x_2 = x_1 \\ 1 + \frac{1}{2}x_3 = x_2 \\ \frac{1}{2}x_2 + x_4 = x_3 \\ \frac{1}{2}x_3 = x_4 \end{cases} \Rightarrow \underset{\sim}{x} = (1, 2, 2, 1)$$

$\Rightarrow \underset{\sim}{\pi} = (\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6})$

e.g.





$$P = \begin{array}{c} \begin{array}{cc} \{1,2\} & \{3,4\} \end{array} \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \end{array}$$

Not a limiting distribution

(initial state matters — 1 or 2 → 1 or 2; 3 or 4 → 3 or 4)

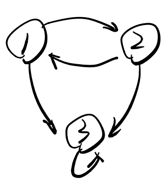
Stationary (not unique)

$$\begin{cases} \frac{1}{2} + \frac{1}{2}x_2 = 1 \\ \frac{1}{2} + \frac{1}{2}x_2 = x_2 \\ \frac{1}{2}x_3 + \frac{1}{2}x_4 = x_3 \\ \frac{1}{2}x_3 + \frac{1}{2}x_4 = x_4 \end{cases} \Rightarrow (1, 1, x, x)$$

$\underset{\sim}{\pi}_1 = (\frac{1}{2}, \frac{1}{2}, 0, 0)$

$\underset{\sim}{\pi}_2 = (0, 0, \frac{1}{2}, \frac{1}{2})$

e.g.



1 and 2 are not accessible from 3.

$\{1,2\}$ $\{3\}$

If we can get to state j from state i, then j is accesible from i.

$(P^n)_{ij} > 0$ for at least one $n \geq 0$

If j is accessible from i and i is accessible from j, then i commuates with j and j communates with i.

Communication class A class (or subclass) of states which all commuicate with each other.

Properties of communication

① Symmetric If i communicates with j, then j communicates with i.

② Reflexive Every state communicates with itself $P^0_{ii} = 1 > 0$

③ Transitive If i comm j, j comm k $\Rightarrow$ i comm k

$$(P^n)_{ik} = \sum_l (P^m)_{il} \, (P^{n-m})_{lk}$$

Take $l = j$, then $(P^m)_{ij} > 0$ for some m since i comm j

$$(P^{n-m})_{jk} > 0$$

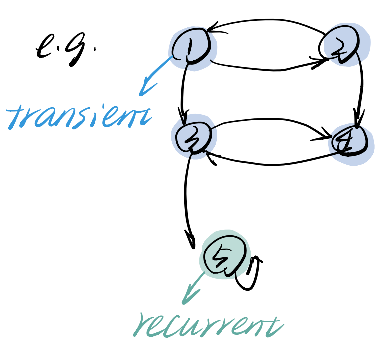
Then $(P^n)_{ik} > (P^m)_{ij} \, (P^{n-m})_{lj} > 0$

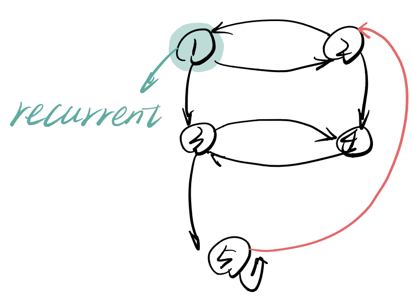
i.e. k is accessible from i

Reverse i and k and repeat.

Irreducibility If a Markov chain has only one communication class, we call the chain irreducible.

e.g.





There is only one communication class (irreducible)

Define $T_j = \min\{n > 0 : X_n = j \text{ if } X_0 = j\}$ First hitting time

$Pr(T_j < \infty \mid X_0 = j) = 1$, then state j is a recurrent state.
$Pr(T_j < \infty \mid X_0 = j) < 1$, then state j is a transient state

How many times will we visit state j ?

Let $I_n = \begin{cases} 1 & \text{if } X_n = j \\ 0 & \text{otherwise} \end{cases}$

$$E\left[\sum_{n=0}^{\infty} I_n\right] = \sum_{n=0}^{\infty} E(I_n) = \sum_{n=0}^{\infty} (P^n)_{jj}$$

- If j is recurrent, $\sum_{n=0}^{\infty} Pr(X_n = j \mid X_0 = j) = \infty$

  $f_j = Pr(T_j < \infty \mid X_0 = j)$
  If j is recurrent, then $f_j = 1$. Finally I get back to state j in finite time
  $\Rightarrow$ Repeat.

- If j is transient, let Y_j be a random variable equal to the number I hit j before I never hit it again.

  $Y_j \sim \text{Geometric}(1 - f_j)$ The probability of never hitting again.

  $E(Y_j) = \dfrac{1}{1 - f_j} < \infty$

$$\begin{cases} \text{State } j \text{ is recurrent iff } \sum_{n=0}^{\infty} P_{jj}^n = \infty \\ \text{State } j \text{ is transient iff } \sum_{n=0}^{\infty} P_{jj}^n < \infty. \end{cases}$$

Communication class C

Let $j \in C$ and let j be recurrent. What about $i \in C$?

$$\sum_{n=0}^{\infty} (P^{r+n+t})_{ii} = \sum_{n=0}^{\infty} \sum_{l} (P^r)_{il} (P^n)_{ll} (P^t)_{li}$$

$$\geq \sum_{n=0}^{\infty} (P^r)_{ij} (P^n)_{jj} (P^t)_{ji}$$

$$= (P^r)_{ij} \underbrace{\sum_{n=0}^{\infty} (P^n)_{jj}}_{j \text{ recurrent}} (P^t)_{ji}$$

$$= \infty$$

$\Rightarrow$ i recurrent

All states in a communication class, they're either transient or recurrent

For a finite irreducible Markov chain, all states are recurrent.

Assume all states are transient. Then state 1 will be visited for a finite amount of time, after which it is never hit again, similarly with other states. Since there are finitely many states, it follows that none of the states will be visited after some finite amount of time, which is not possible.

Canonical decomposition of the chain

closed communication class $\quad P_{ij} = 0 \quad \forall\ i \in C,\ j \notin C$

A communication class is closed if it consists of all recurrent states

A finite communication class is closed only if it consists of all recurrent states.

$$\Rightarrow \quad S = T \cup \underbrace{R_1 \cup R_2 \cup \cdots \cup R_m}_{\text{closed communication class of recurrent state}}$$

T is a class of transient state

$$P = \begin{array}{c} \\ T \\ R_1 \\ \vdots \\ R_m \end{array} \begin{array}{c} \begin{array}{cccc} T & R_1 & \cdots & R_m \end{array} \\ \begin{pmatrix} \square & ? & ? & ? \\ & P_1 & & 0 \\ & & \ddots & \\ 0 & & & P_m \end{pmatrix} \end{array}$$

Because R_j's are closed recurrent. P_k's must be regular and irreducible

$$\lim_{n\to\infty} P^n = \begin{pmatrix} * & * & \cdots & * \\ 0 & \lim_{n\to\infty} P_1^n & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & & & \lim_{n\to\infty} P_m^n \end{pmatrix}$$

e.g. $\quad P = \begin{pmatrix} R_1 & 0 \\ 0 & R_2 \end{pmatrix} = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}$

if π_1 stationary for P_1.

$$(\underset{\sim}{\pi}, 0)\, P = (\underset{\sim}{\pi}_1 P_1, \underset{\sim}{0}) = (\underset{\sim}{\pi}_1, 0)$$

$$(0, \underset{\sim}{\pi})\, P = (\underset{\sim}{0}, \underset{\sim}{\pi}_2 P_2) = (0, \underset{\sim}{\pi}_2)$$

$T_j = \min\{n>0 : X_n = j \text{ if } X_0 = j\}$

Let $\mu_j = E(T_j \mid X_0 = j)$

Limit Theorem for Finite Irreducible Markov Chain

Let $X_0, X_1, \dots$ be a finite irreducible Markov chain. Then

$\mu_j < \infty$ and $\exists$ a unique, positive stationary distribution π s.t.

$$\pi_j = \frac{1}{\mu_j} \quad \forall j$$

doesn't imply π_j is limiting

Furthermore, for all state i

$$\pi_j = \lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} P_{ij}^{m}$$

Note: If $|S| = \infty$, then $E(T_j \mid X_0 = j)$ could be infinite, even if j is recurrent.

- Positive recurrent: recurrent j s.t. $E(T_j \mid X_0 = j) = \infty$
- Null recurrent: $E(T_j \mid X_0 = j) = \infty$

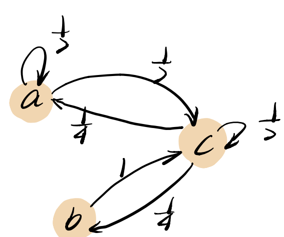
$$\sum_{n=0}^{\infty} (T_j = n \mid X_0 = j)\, n = \infty$$

transient
$Pr(T_j = \infty) > 0$

$Pr(T_j = n) > 0$ as $n \to \infty$
but $Pr(T_j = n)\cdot n \not\to 0$ as $n \to \infty$

First Step Analysis

$$P = \begin{array}{c|ccc} & a & b & c \\ \hline a & \frac{1}{2} & 0 & \frac{1}{2} \\ b & 0 & 0 & 1 \\ c & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{array}$$



Goal: $E(T_a = n \mid X_0 = a)$

Let $e_x = E(T_{xa} \mid X_0 = x)$ for T_{xa} = time to get to a from x

$$e_a = Pr(X_1 = a \mid X_0 = a)\cdot 1 + Pr(X_1 = c \mid X_0 = a)\cdot(e_c + 1)$$

1 step used

$$= \frac{1}{2} + \frac{1}{2}(e_c + 1)$$

$$e_c = \frac{1}{4}\cdot 1 + \frac{1}{4}(\underbrace{e_b}_{e_c+1} + 1) + \frac{1}{2}\cdot(e_c + 1)$$

$\Rightarrow e_c = 5 \quad e_a = 3.5$

Note: $E(X|X_0=x) = \sum_y E(X|X_1=y, X_0=x) Pr(X_1=y|X_0=x)$

$= \sum_y E(X|X_1=y) Pr(X_1=y|X_0=x)$ Markov

$= \sum_y [1+E(X|X_0=y)] Pr(X_1=y|X_0=x)$ TH

Periodicity

<u>definition:</u> Period of a state i

$$d(i) = \gcd\{n>0 : (P^n)_{ii} > 0\}$$

If $d(i)=1$, then i is aperiodic

3 If $d(i)=\infty$

Periodicity is also a class property. All states in a communication class have the same period.

Let $i, j \in C$. i, j commute $\Rightarrow \exists r, s$ s.t. $P^r_{ij} > 0$, $P^s_{ji} > 0$

Then $P^{r+s}_{ii} = \sum_k P^r_{ik} P^s_{ki} \geq P^r_{ij} P^s_{ji} > 0 \Rightarrow r+s$ is a return time for i.

Assume $P^n_{jj} > 0$ for some n.

$P^{r+n+s}_{ii} \geq P^r_{ij} P^n_{jj} P^s_{ji} > 0 \Rightarrow r+n+s$ is also a return time for i.

So $d(i)$ is a divisor of both $r+s$ and $r+n+s$

So $d(i)$ is a divisor of $\{n>0 : P^n_{jj}\}$ where
$d(j)$ is the biggest divisor of $\{n>0 : P^n_{jj} > 0\}$

$\Rightarrow d(i) \leq d(j)$ Reverse i and j, then $d(i) = d(j)$

A Markov chain is

- aperiodic if it is irreducible and $d(i) = 1 \ \forall i$
- periodic if it is irreducible and all states have period >1 $(d(i)>1)$

Note: Having a self loop ensures aperiodicity.

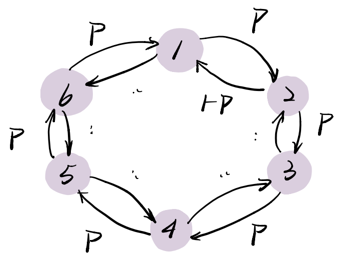
A Markov chain is ergodic if

1. irreducible
2. aperiodic
3. All states have finite return time (positive recurrent)

Let $X_0, X_1 \cdots$ be an ergodic Markov chain. Then $\exists$ a unique, positive stationary distribution which is the limiting distribution of the chain.

A chain is ergodic iff TPM is regular: $\exists\ n$ s.t. $P^n > 0$

e.g.



$$\begin{cases} p > \frac{1}{2} & \text{There is a bias to move clockwise} \\ p < \frac{1}{2} & \text{'' '' counterclockwise} \\ p = \frac{1}{2} & \text{no bias} \end{cases}$$

Observed sequence: 1 2 1 6 5 4 5 4 3 2 3 2 1 6

6 1 2 3 2 3 4 5 4 5 6 1 2 1 (reversed)

If I cannot tell whether a chain is going forwards or backwards we say the chain is time reversible

for a irreducible MC.

$\pi_i P_{ij} = \pi_j P_{ji} \ \forall i, j$ for stationary $\underline{\pi}$

for large n, $Pr(X_n = i) Pr(X_{n+1} = j \mid X_n = i) = Pr(X_n = j) Pr(X_{n+1} = i \mid X_n = j)$

$\Rightarrow \quad Pr(X_n = i, X_{n+1} = j) = Pr(X_n = j, X_{n+1} = i)$

$\Rightarrow \quad Pr(X_0 = i_0 \cdots X_n = i_n) = Pr(X_0 = i_n \cdots X_n = i_0)$

If $\exists\ t = (t_1 \cdots t_k)$ s.t. $\sum t_j = 1$ and satisfies time reversibility

$t_i P_{ij} = t_j P_{ji} \ \forall i, j$,

then

$$\sum_j t_i P_{ij} = \sum_j t_j P_{ji}$$

$$t_i \sum_j P_{ij} = \sum_j t_j P_{ji}$$

$$t_i = (t_1 \cdots t_k) \begin{pmatrix} P_{1i} \\ \vdots \\ P_{ki} \end{pmatrix}$$

$$t_i = (tP)_i$$

$\Rightarrow$ t is a stationary process

Absorbing chain

State i is absorbing state if $P_{ii}=1$ recurrent in their own communication class

A MC is an absorbing chain if there exists at least one absorbing state.

Assume we have chain with t transient state and k absorbing states.

$$P = \begin{array}{c} \begin{array}{cc} t\times t & t\times k \end{array} \\ \begin{pmatrix} Q & R \\ 0 & I \end{pmatrix} \\ \begin{array}{cc} k\times t & k\times k \end{array} \end{array} \begin{array}{l} \text{transient} \\ \text{absorbing} \end{array} \qquad P^2 = \begin{pmatrix} Q^2 & QR+R \\ 0 & I \end{pmatrix} \quad \cdots \quad \cdots$$

$$P^n = \begin{pmatrix} Q^n & (Q^{n-1}+\cdots+Q+I)R \\ 0 & I \end{pmatrix}$$

lemma: Let A be a square matrix s.t. $\lim_{n\to\infty}(A^n)=0$, then $\sum_{n=0}^{\infty} A^n = (I-A)^{-1}$

Hewistic: $\sum_{k=0}^{\infty} r^k = (1-r)^{-1}$ where $0<r<1$

Note that Q indexes only transient states.

$\Rightarrow \lim_{n\to\infty}(Q^n)_{ij} = 0$ for all i,j transient

$$\Rightarrow \lim_{n\to\infty} P^n = \lim_{n\to\infty}\begin{pmatrix} Q^n & \overbrace{(Q^{n-1}+\cdots+Q+I)R}^{\sum_{k=0}^{n-1} Q^k R} \\ 0 & I \end{pmatrix}$$

$$= \begin{pmatrix} \lim_{n\to\infty} Q^n & \lim_{n\to\infty}\sum_{k=0}^{n-1} Q^k R \\ 0 & I \end{pmatrix}$$

$$= \begin{pmatrix} 0 & (I-Q)^{-1}R \\ 0 & I \end{pmatrix}$$ Not necessarily limiting

$(I-Q)^{-1}$ is called the fundamental matrix for absorbing chains.

Let F be a $(t\times t)$ matrix where F_{ij} contains the expected number of visits to the state j when starting in state i where i,j are transient. Then $F=(I-Q)^{-1}$

The ith row of $\lim_{n\to\infty}(P^n)$

$$\underset{1\times t}{\underset{\sim}{0}} \quad \underset{1\times k}{\left(\cdots, \sum_{j=1}^{k} F_{ij}R_{ji}, \cdots\right)}$$

proof

For a chain starting at transient i, define $I_n = \begin{cases} 1 & X_n = j \\ 0 & \text{otherwise} \end{cases}$

$\sum_{n=0}^{\infty} I_n = \#$ visits to j

$E(\sum_{n=0}^{\infty} I_n) = \sum_{n=0}^{\infty} E(I_n) = \sum_{n=0}^{\infty} Pr(X_n = j \mid X_0 = i) = \sum_{n=0}^{\infty} (P^n)_{ij}$

if j is transient, then $\sum_{n=0}^{\infty} (Q^n)_{ij} = (I-Q)^{-1}_{ij}$

For an absorbing Markov chain starting in transient state i, let a_i be the expected absorption time, the expected number of steps to reach some absorbing state.

$a_i = \sum_{k \in T} F_{ik}$ row sums of the fundational matrix.

Assume I have an irreducible train. Want to find expected time to state k under TPM P for $S = \{1, \cdots k\}$

Let

$$P = \begin{array}{c} \begin{array}{cc} Q & R \end{array} \\ \left(\begin{array}{ccc|c} P_{11} & \cdots & P_{1\,k-1} & P_{1k} \\ \vdots & & \vdots & \vdots \\ P_{k-1,1} & \cdots & P_{k-1\,k-1} & P_{k-1\,k} \\ \hline 0 & \cdots & 0 & 1 \end{array}\right) \\ I_1 \end{array}$$

Then

$F = (I-Q)^{-1}$ and

$$Q = \begin{pmatrix} P_{11} & \cdots & P_{1\,k-1} \\ \vdots & & \\ P_{k-1\,1} & \cdots & P_{k-1\,k-1} \end{pmatrix}$$

$$P = \begin{pmatrix} Q & R \\ 0 & I \end{pmatrix} \qquad \lim_{n\to\infty} P^n = \begin{pmatrix} 0 & (I-Q)^{-1}R \\ 0 & I \end{pmatrix}$$

Let A be a square matrix s.t. $A^n \xrightarrow{\text{as } n\to\infty} 0$, then $\sum_{n=0}^{\infty} A^n = (I-A)^{-1}$

$$(I-A)(I+A+A^2+\cdots A^n) = (I+A+A^2+\cdots+A^n) - (A+A^2+\cdots+A^{n+1}) = I-A^{n+1}$$

If $I-A$ is invertible, then

$$I+A+A^2+\cdots A^n = \frac{I-A^{n+1}}{I-A}$$

$$\Rightarrow \sum_{n=0}^{\infty} A^n = (I-A)^{-1}.$$

Show $I-A$ is invertible

$$(I-A)\underset{\sim}{x} = \underset{\sim}{0}$$

$(I-A)$ is invertible iff $\underset{\sim}{x} = \underset{\sim}{0}$ is the only solution

$$\underset{\sim}{0} = (I-A)\underset{\sim}{x} = \underset{\sim}{x} - A\underset{\sim}{x}$$

$$\Rightarrow \underset{\sim}{x} = A\underset{\sim}{x} = A(A\underset{\sim}{x}) \cdots = A^n\underset{\sim}{x}$$

$$\Rightarrow \lim_{n\to\infty} \underset{\sim}{x} = \underset{\sim}{x} = \lim_{n\to\infty} A^n\underset{\sim}{x} = 0$$

$$\Rightarrow \underset{\sim}{x} = \underset{\sim}{0}$$

3.9 Regression

A random variable S is a stopping time for a Markov chain.
each possible value $S=s$, the event $\{S=s\}$ can be

$$Pr = (S/s \mid X_0, X_1 \cdots X_s) = \begin{cases} 1 \\ 0 \end{cases}$$

First hitting time $\quad T_i = \min\{n \geq 0 : X_n = i\}$

$$Pr(T_i = t_i \mid (X_0, X_1 \cdots X_i)) = \begin{cases} 1 & \text{if } X_{t_i} = i \text{ and } X_t \neq i \ \forall\, t < t_i \\ 0 & \text{otherwise} \end{cases}$$

Note: The last return time is not a stopping time.

Strong Markov Property

Let $X_0, X_1 \cdots X_n$ be a Markov chain with TPM P. Let S be a stopping time for $\{X_n\}$, then

$X_s, X_{s+1}, \cdots \cdots$ is also a Markov chain with TPM P.