Markov chains - review

(See EE6150 course homepage for detailed notes)

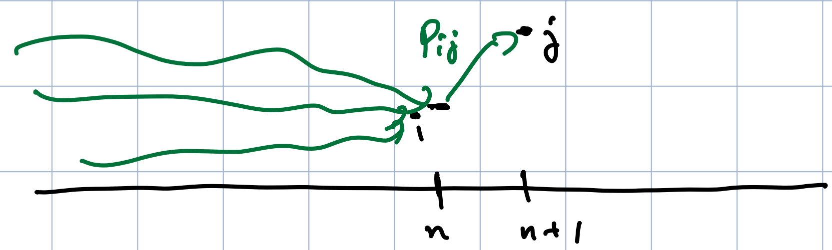
Def: A stochastic process $\{X_n, n \geq 0\}$ with a countable state space S is a DTMC if

(i) $X_n \in S, \forall n \geq 0$

(ii) $\forall n \geq 0, \quad i, j \in S,$

Markov property
$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, \ldots, X_0 = i_0)$$
$$= P(X_{n+1} = j \mid X_n = i)$$

"The future is conditionally independent of the past, given the present".



Def: A DTMC with countable state space S is time-homogeneous if

$$P(X_{n+1} = j \mid X_n = i) = P_{i,j} \quad \forall n \geq 0, \ \forall i, j \in S$$

transition probability

If the state space is finite, we can form the "transition probability matrix" P as follows:

Let $|S| = m$

$$P = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ m \end{array} \begin{array}{c} \begin{array}{ccc} 1 & \cdots & m \end{array} \\ \begin{bmatrix} P_{1,1} & & P_{1,m} \\ \vdots & & \\ P_{m,1} & \cdots & P_{m,m} \end{bmatrix} \end{array}$$

Q: Is the transition probability matrix (t.p.m.), say enough to derive the finite dimensional distribution i.e.,

$$\boxed{P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n)} \text{ ? "NO"}$$

Simple case: $P(X_0 = i_0, X_1 = i_1)$

$$= P(X_1 = i_1 \mid X_0 = i_0)\, P(X_0 = i_0)$$

$$= P_{i_0, i_1} \times P(X_0 = i_0)$$

Suppose we are given

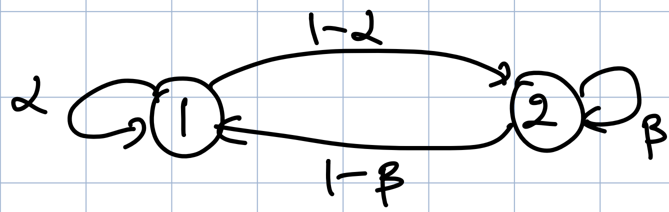
$$\left. a_i = P(X_0 = i), \quad \forall i \in S \right] \text{Initial distribution}$$

Fact: A DTMC $\{X_n, n \geq 0\}$ is completely specified by the initial distribution "a" & t.p.m. P.

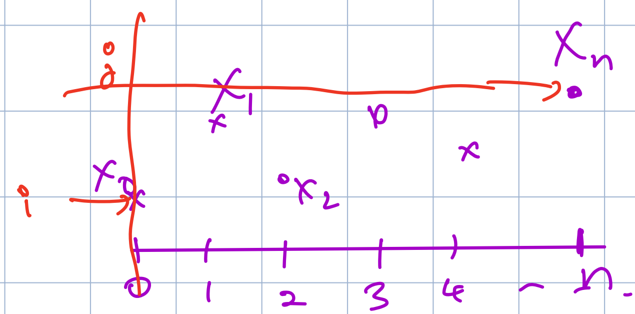
Example: Two-state DTMC

$$S = \{1, 2\}$$

$$P = \begin{bmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{bmatrix} \qquad 0 \le \alpha, \beta \le 1$$



Marginal distributions



Let $\{X_n, n \geq 0\}$ be a DTMC with state space $S = \{0, 1, 2, \ldots\}$, t.p.m. "P", & initial distribution "a"

Want: $a_j^{(n)} = P(X_n = j)$ ← distribution of X_n

$$P(X_n = j) = \sum_{i \in S} P(X_n = j \mid X_0 = i)\, P(X_0 = i)$$

$$= \sum_{i \in S} P(X_n = j \mid X_0 = i)\, a_i$$

$$= \sum_{i \in S} a_i\, p_{i,j}^{(n)},$$

where $p_{i,j}^{(n)} = P(X_n = j \mid X_0 = i), \ \forall i, j \in S, \ n \geq 0$

n-step transition probability

In particular,

$$P_{i,j}^{(0)} = P(X_0 = j \mid X_0 = i) = \delta_{i,j}, \quad i,j \in S,$$

where $\delta_{i,j} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{else} \end{cases}$

$$P_{i,j}^{(1)} = P(X_1 = j \mid X_0 = i) = P_{i,j}, \quad i,j \in S$$

Q: How to calculate $P_{i,j}^{(n)}$, $n \geq 2$?

Chapman-Kolmogorov equations

$$P_{i,j}^{(n)} = \sum_{r \in S} P_{i,r}^{(k)} P_{r,j}^{(n-k)}, \quad i,j \in S$$

where $0 \leq k \leq n$

$$\begin{tikzcd} i \arrow[r, bend left, "\text{"k steps"}"] & r \arrow[r, bend left, "\text{"n-k steps"}"] & j \end{tikzcd}$$

Suppose $|S| = m < \infty$ (finite state space)

$$P^{(n)} = \begin{bmatrix} P_{1,1}^{(n)} & \cdots & P_{1,m}^{(n)} \\ \vdots & & \\ P_{m,1}^{(n)} & \cdots & P_{m,m}^{(n)} \end{bmatrix}$$

n-step t.p.m.

$$P^{(n)} = \left[p^{(n)}_{i,j} \right]_{i,j=1\ldots m}$$

Lecture 29*

Chapman-Kolmogorov equation:

$$P^{(n)} = P^{(k)} \, P^{(n-k)}, \quad 0 \le k \le n$$

Fact: $P^{(n)} = P^{n}$

nth power of t.p.m. P

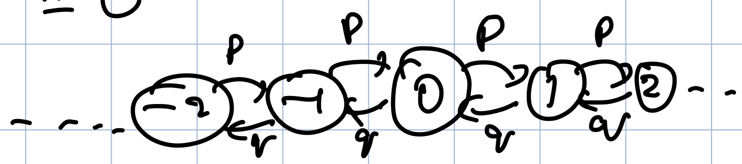
EXAMPLE: Random walk

$P_{i,i+1} = p, \quad P_{i,i-1} = q \; (=1-p), \quad -\infty < i < \infty$

$0 < p < 1$

Find $P^{(n)}_{0,0} = P(X_n = 0 \mid X_0 = 0)$

If n is odd, $P^{(n)}_{0,0} = 0$



If n is even, say $n = 2k$, then one has to take k steps to right & k steps to left, in any order.

$$P^{(2k)}_{0,0} = \frac{(2k)!}{k!\,k!} \; p^{k} q^{k}$$

ways (distinct) of k right &

prob of k left × prob k right steps

k left steps

EXAMPLE:

DTMC with $S = \{1, 2, 3, 4\}$

$$a = \{\tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}, \tfrac{1}{4}\}$$

$$P = \begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.25 & 0.25 & 0.5 & 0 \\ 0.5 & 0 & 0.1 & 0.4 \\ 0 & 0 & 0.4 & 0.6 \end{bmatrix}$$

$$a^{(4)} = a\, P^4 = \text{H.W.},$$

where $a^{(4)} = \left[P(X_4 = 1) \;\; P(X_4 = 2) \;\; \cdots \;\; P(X_4 = 4) \right]$

DTMCs: First passage times

Let $\{X_n, n \geq 0\}$ be a DTMC on $S = \{0, 1, 2, \ldots\}$

$$T = \underbrace{\min\{n \geq 0 \mid X_n = 0\}}_{\text{First passage time into "0".}}$$

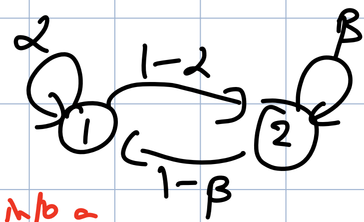
Quantities related to T:

① Complementary cdf $P(T > n)$, $n \geq 0$

② Probability of eventually hitting "0" $P(T < \infty)$

EXAMPLE:

Two-state DTMC



$$P = \begin{bmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{bmatrix}$$

to obtain the first passage into a certain state

$$T = \min\{n \geq 0 \mid X_n = 1\}$$

$$v_2(n) = P(T > n \mid X_0 = 2) = \beta^n$$

$$P(T = n \mid X_0 = 2) = v_2(n-1) - v_2(n)$$
$$= \underbrace{\beta^{n-1}(1-\beta)}_{\text{Geometric distribution}}$$

Occupancy times

$\{X_n, n \geq 0\}$ DTMC with state space S, t.p.m. P

$V_j^{(n)}$ = # visits to "j" upto time n (including 0)

Note: $V_j^{(0)} = 1$ if $X_0 = j$

Define $\underbrace{M_{i,j}^{(n)} = E\left(V_j^{(n)} \mid X_0 = i\right)}_{\text{Occupancy time of "}j\text{" upto "}n\text{", starting in "}i\text{"}}$, $i, j \in S, n \geq 0$

Occupancy matrix $\rightarrow M^{(n)} = \left[M^{(n)}_{i,j} \right]_{i,j \in S}, \quad |S| < \infty$

Fact : $M^{(n)} = \sum_{r=0}^{n} P^r, \quad n \geq 0$

$\boxed{\begin{aligned} P^0 &= I \\ P^1 &= P \end{aligned}}$

EXAMPLE:

3-State DTMC, $S = \{A, B, C\}$

t.p.m. $P = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.4 & 0.4 \\ 0.1 & 0.3 & 0.6 \end{bmatrix} \end{array}$

$$M^{(9)} = \sum_{r=0}^{9} P^r = \begin{bmatrix} 2.14 & 2.74 & 5.12 \\ 1.26 & 3.95 & 4.78 \\ 1.15 & 2.85 & 6 \end{bmatrix}$$

$M^{(9)}_{(A,A)} = 2.14 =$ if customer buys A on day 0 (initial distribution), then the expected # of purchases of A = 2.14 over 10 days

$M^{(9)}_{(A,C)} = 5.12$

Remark: Marginal distributions & passage times → these characterize the transient behaviour of Markov chains.

Markov chains - Limiting behaviour

(I) Recall $P^{(n)} \to$ n-step t.p.m.

$P^{(n)}_{i,j} \to$ prob. of going from i to j in "n" steps

$$P^{(n)} = P^n$$

Q: Does $P^{(n)}$ converge as $n \to \infty$?

(II) Occupancy matrix $M^{(n)} = \sum_{r=0}^{n} P^r$

$M^{(n)}_{i,j} \to$ # visits to j starting in i up to time n

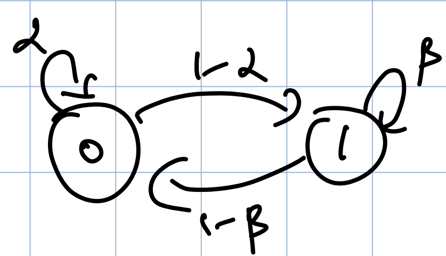
$M^{(n)} \to$ row sums are "$n+1$"

Q: Does $\frac{M^{(n)}}{n+1}$ converge as $n \to \infty$?

Example

① Two-state DTMC:

$\alpha + \beta < 2$



By an induction argument, it can be shown that

$$P^n = \frac{1}{2-\alpha-\beta}\begin{bmatrix} 1-\beta & 1-\alpha \\ 1-\beta & 1-\alpha \end{bmatrix} + \frac{(\alpha+\beta-1)^n}{2-\alpha-\beta}\begin{bmatrix} 1-\alpha & \alpha-1 \\ \beta-1 & 1-\beta \end{bmatrix}, \forall n$$

H.W. Check case $n=1$ $P^1 = P = \begin{bmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{bmatrix}$

induction hypothesis for $k=1 \ldots n$ & prove for '$n+1$' using $P^{n+1}=P^nP$

$$\lim_{n\to\infty} P^n = \frac{1}{2-\alpha-\beta}\begin{bmatrix}1-\beta & 1-\alpha\\ 1-\beta & 1-\alpha\end{bmatrix}$$

Another induction argument gives

$$M^{(n)} = \frac{n+1}{2-\alpha-\beta}\begin{bmatrix}1-\beta & 1-\alpha\\ 1-\beta & 1-\alpha\end{bmatrix} + \frac{1-(\alpha+\beta-1)^{n+1}}{(2-\alpha-\beta)^2}\begin{bmatrix}1-\alpha & \alpha-1\\ \beta-1 & 1-\beta\end{bmatrix}$$

$$\lim_{n\to\infty}\frac{M^{(n)}}{n+1} = \frac{1}{2-\alpha-\beta}\begin{array}{cc} & \begin{array}{cc}0 & \quad 1\end{array}\\ \begin{array}{c}0\\1\end{array} & \begin{bmatrix}1-\beta & 1-\alpha\\ 1-\beta & 1-\alpha\end{bmatrix}\end{array}$$

P^n & $\dfrac{M^{(n)}}{n+1}$ Converged to the same limit.

Classification of states

"Accessibility"

A state j is said to be accessible from a state i if $\exists\, n\geq 0$ s.t. $P_{i,j}^{(n)}>0$.

If j is accessible from i, we write $i\to j$

$i\to j \Rightarrow \exists$ a directed path from i to j in the transition diagram

Communication: States i and j communicate if $i \to j$ and $j \to i$

Denote communication by "$\leftrightarrow$"

Claim:

(i) $i \leftrightarrow i$ (reflexive)

(ii) If $i \leftrightarrow j$ then $j \leftrightarrow i$ (symmetric)

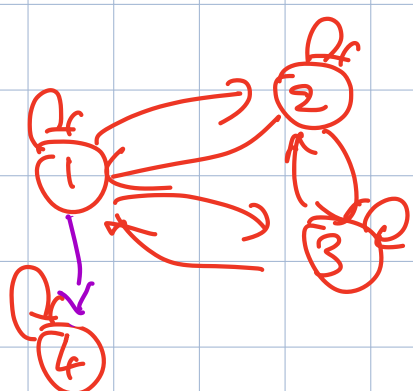
(iii) If $i \leftrightarrow j$, $j \leftrightarrow k$, then $i \leftrightarrow k$ (transitive)

Communicating class:

A set $C \subset S$ is a communicating class if

(i) $i \in C,\ j \in C \Rightarrow i \leftrightarrow j$

(ii) $i \in C,\ i \leftrightarrow j \Rightarrow j \in C$ $\leftarrow$ makes C maximal



$1 \leftrightarrow 2$ $2 \leftrightarrow 3$

$\{1,2\}$ is not a communicating class

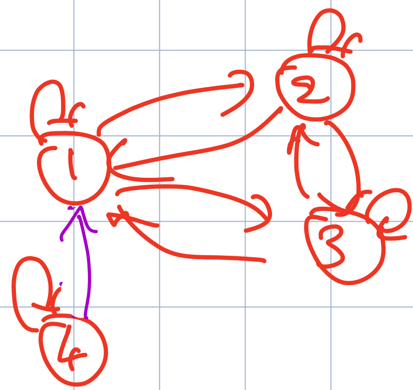
but $\{1,2,3\}$ is.

$S = \{1,2,3\} \cup \{4\}$

Closed communicating class

A communicating class C is closed if for any $i \in C$ & $j \notin C$, we have $i \not\to j$

j is not accessible from i



$\{1, 2, 3\}$ is a closed communicating class

If $X_n \in C$ for some n, & C is closed, then $X_m \in C \quad \forall\, m \geq n$.

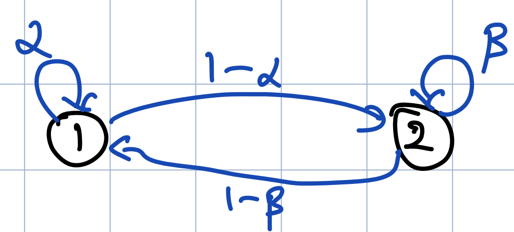
We can partition the state space as

$$S = \underbrace{C_1 \cup C_2 \cup \dots \cup C_k}_{\text{disjoint closed communicating classes}} \cup \underbrace{T}_{\text{left-over states (could include communicating classes that aren't closed)}}$$

Irreducibility: If the state space S is a single closed communicating class, then the DTMC is said to be irreducible.

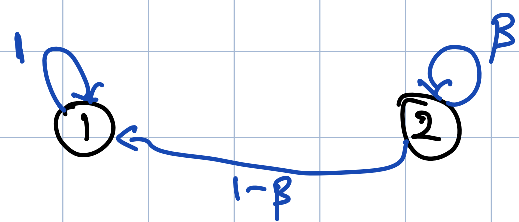
Else it is reducible.

Example:



If $0 \leq \alpha, \beta < 1$, then $\{1, 2\}$ is a closed communicating class & the DTMC is irreducible.

Suppose $\alpha = 1$ & $0 < \beta < 1$



$\{1\}$ is a closed communicating class

$\{2\}$ is not closed

Partition of state space $= \underbrace{\{1\}}_{C_1} \cup \underbrace{\{2\}}_{T}$

Case $\alpha = \beta = 1$: Partition: $\underbrace{\{1\}}_{C_1} \cup \underbrace{\{2\}}_{C_2}$ $\quad T = \phi$

Recurrence & transience:

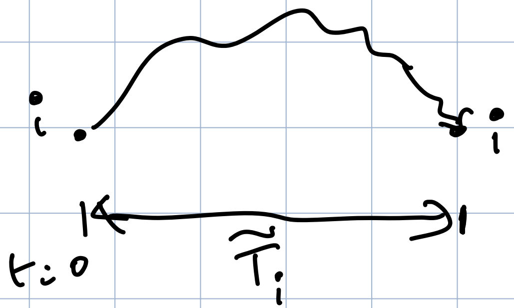
$$\tilde{T}_i = \min\{n > 0 \mid X_n = i\}, \; i \in S$$

$$\tilde{u}_i = P(\tilde{T}_i < \infty \mid X_0 = i) \rightarrow \text{Prob of returning to state } i$$

$$\tilde{m}_i = E(\tilde{T}_i \mid X_0 = i) \rightarrow \text{expected \# steps taken before returning}$$

Def: A state i is $\begin{cases} \text{recurrent if } \tilde{u}_i = 1 \\ \text{transient if } \tilde{u}_i < 1 \end{cases}$

Q: If $\underset{\text{state is transient}}{\tilde{u}_i < 1}$, then $\underset{\text{mean \# steps for return}}{\tilde{m}_i = \infty}$.



Recurrent i : $P_i(\tilde{T}_i < \infty) = 1$

Transient i : $P_i(\tilde{T}_i < \infty) < 1$

Def: A recurrent state i is

positive recurrent if $\tilde{m}_i < \infty$

null-recurrent if $\tilde{m}_i = \infty$

expected return time

Example: If i is an absorbing state, then

$$P(X_1 = i \mid X_0 = i) = 1$$

$$\tilde{u}_i = 1, \quad \tilde{m}_i = 1$$

Thus, i is positive recurrent.

Recurrence & transience are (communicating) class properties

Facts:

1. i is recurrent, $i \leftrightarrow j \Rightarrow j$ is recurrent
2. i transient, $i \leftrightarrow j \Rightarrow j$ is transient.

Next, positive/null recurrence are class properties

More Facts: ① $i \leftrightarrow j$, i is positive recurrent

$\Rightarrow$ j is positive recurrent

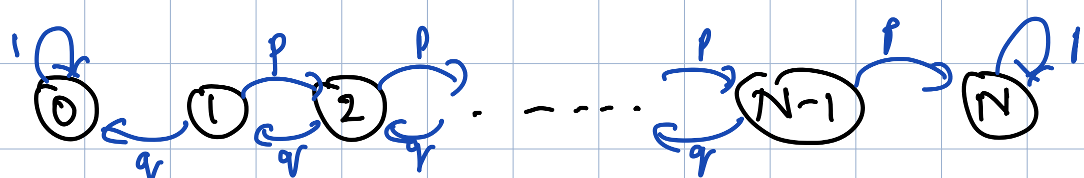
② $i \leftrightarrow j$, i is null recurrent
$\Rightarrow$ j is null recurrent

(I) A communicating class is called
(i) "transient" if all its states are transient
(ii) "positive recurrent" if all its states are positive recurrent
(iii) "null recurrent" if all its states are null recurrent

(II) An irreducible DTMC is positive (null recurrent, transient) if all its states are positive (null recurrent, transient).

Example: Random walk

$$P_{0,0} = P_{N,N} = 1, \quad P_{i,i+1} = p, \quad P_{i,i-1} = q, \quad 1 \le i \le N-1$$



$$0 < p, q < 1, \quad p + q = 1$$

State 0 : recurrent
N : ——"——
$\{1, \dots, N-1\}$: transient

Lecture-30

Markov chains taken to their limit

Fact: Finite irreducible Markov chain

$\Rightarrow$ recurrent & also positive recurrent

Stationary distributions

(aka steady-state / invariant distribution)

Def: For a discrete-time Markov chain with t.p.m. P, the vector $\pi = (\pi_i, i \in S)$ is called a stationary distribution if

(i) $\pi_i \geq 0 \ \forall i$, $\sum_i \pi_i = 1$ $\leftarrow$ π is a distribution

(ii) $\pi = \pi P$ $\leftarrow$ π is stationary

Remark: If the initial distribution is π, then what is the distribution of X_n: πP^n

$$= \pi P . P^{n-1}$$
$$= \pi P^{n-1}$$
$$\vdots$$

dist^n of X_n $= \pi$

Main result. Consider an irreducible Markov chain

(a) $\exists$ a stationary distribution π if and only if some state is positive recurrent.

(i.e., If DTMC is irreducible + transient/null-recurrent, then NO stationary distribution)

(b) If there exists a stationary distribution π, then (i) every state is positive recurrent

(ii) $\boxed{\pi_i = \frac{1}{m_i}}$ $\forall i \in S$, where

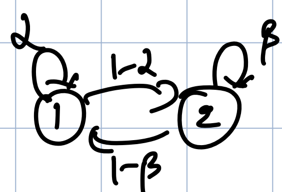
mean return time $m_i := E(T_i \mid X_0 = i)$, with $T_i := \min\{n \geq 1 \mid X_n = i\}$

[i.e., $\frac{1}{m_i} = \sum_j \frac{1}{m_j} P_{j,i}$, since $\pi_i = \sum_{j \in S} \pi_j P_{j,i}$ $\Leftrightarrow$ $\pi = \pi P$]

(iii) π is unique.

Example:

Two state DTMC:



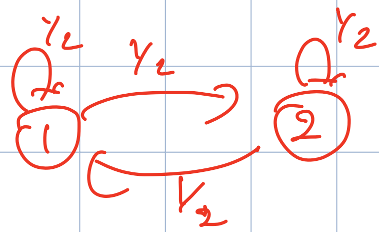
$\alpha + \beta < 2$

$\pi = \pi P \iff \left.\begin{aligned} \pi_1 &= \alpha \pi_1 + (1-\beta)\pi_2 \\ \pi_2 &= (1-\alpha)\pi_1 + \beta \pi_2 \end{aligned}\right\} \iff \pi = \pi P$

$\pi_1 + \pi_2 = 1$

If you solve, then

$$\pi_1 = \frac{1-\beta}{2-\alpha-\beta}, \quad \pi_2 = \frac{1-\alpha}{2-\alpha-\beta}.$$



$$\frac{1}{n}\sum_{m=1}^{n} I\{X_m=1\} \xrightarrow[n\to\infty]{} \frac{1}{2}$$

like a LLN

is the stationary dist^n value for state 1

$\pi_1 \to$ space avg.

LHS $\to$ time-average

In the limit, time-avg = space-avg

"ergodic".

Last Big Fact:

"Convergence in Cesaro sense":

$$\tilde{V}_n(j) = \sum_{m=1}^{n} I(X_m = j)$$ ← occupancy measure for state j

Time average $= \frac{1}{n} \tilde{V}_n(j)$

As $n \to \infty$, where does the "time average" converge for a irreducible, positive recurrent DTMC?

$$\frac{1}{n} \tilde{V}_n(j) \longrightarrow \pi_j \quad \text{as } n \to \infty \quad \text{a.s. (w.p. 1)}$$

or equivalently, $\frac{1}{n}\tilde{V}_n(j) \to \frac{1}{m_j}$ as $n \to \infty$ w.p. 1.

For a null recurrent state j,

$$\frac{1}{n} \tilde{V}_n(j) \longrightarrow 0 \quad \text{a.s. as } n \to \infty.$$

(the same for transient states)

Big Fact II: Consider a irreducible, recurrent DTMC. Then, for any $j \in S$,

a law of large numbers result for Markov chains →

$$\frac{\tilde{V}_n(j)}{n} \longrightarrow \frac{I\{T_j < \infty\}}{m_j} \quad \text{as } n \to \infty \quad \text{w.p. 1}$$

(almost surely = a.s.) ⇕ w.p. 1

for any start state $X_0 = i$

A few remarks:

① $$E\left(\frac{\tilde{V}_n(j)}{n}\,\Big|\,X_0=i\right) = \frac{1}{n}\sum_{m=1}^{n} E\left(I(X_m=j)\,|\,X_0=i\right)$$
$$= \frac{1}{n}\sum_{m=1}^{n} p^{(m)}(i,j)$$

$$E\left(\frac{\mathbb{I}\{T_j<\infty\}}{m_j}\,\Big|\,X_0=i\right) = \frac{P(T_j<\infty\,|\,X_0=i)}{m_j}$$

So,
$$\frac{1}{n}\sum_{m=1}^{n} p^{(m)}(i,j) \xrightarrow{n\to\infty} \frac{P(T_j<\infty\,|\,X_0=i)}{m_j} \quad \text{— (**)}$$

In (**), if j is null recurrent, then the limit is zero.
In (**), if j is positive recurrent, then the limit is positive.

② For a transient state j, $\sum_{m=1}^{\infty} p^m(i,j) < \infty$.

$$\frac{1}{n}\sum_{m=1}^{\infty} p^m(i,j) \xrightarrow{n\to\infty} 0$$

i.e., $$E\left(\frac{\tilde{V}_n(j)}{n}\,\Big|\,X_0=i\right) \xrightarrow{n\to\infty} 0$$

for transient j.