CSE 107 6-2-26

- Final: Th 6/11, 4-6 pm
- HW7: Tonight 10 pm.

7.1 Discrete Markov Chains

7.1 Discrete Markov Chains

The Bernoulli and Poisson Processes were memoryless: the future does not depend on, and is independent of the Past

We now consider Processes in which the future depends, in a limited way, on the Past. The effect of the Past will be embodied in a state, which changes over time.

The set S of states can be anything in general, but we will always take it to be the finite set

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

of integers.

Let $X_0, X_1, \ldots, X_n, \ldots$ be a sequence of random variables taking values in S, and consider

$$X_n = \text{state of Process at time } n.$$

we define the transition probabilities p_{ij} to be

$$p_{ij} = P(X_{n+1}=j \mid X_n=i)$$

note: does not depend on n

for any $i,j \in S$ and $n \geq 0$.

Such a process is said to satisfy the <u>Markov Property</u> iff

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

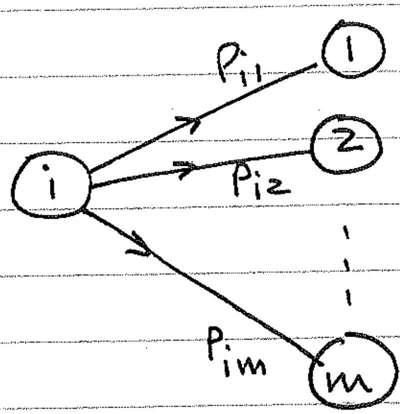
$$= P(X_{n+1}=j \mid X_n=i) = p_{ij}$$

<u>i.e. the state depends on the past only throug the immediatly preceding state, not on the entire history of states.</u>

The state space S, together with transition probabilities, and X_n satisfying the Markov Property, is called a <u>Markov chain model.</u>

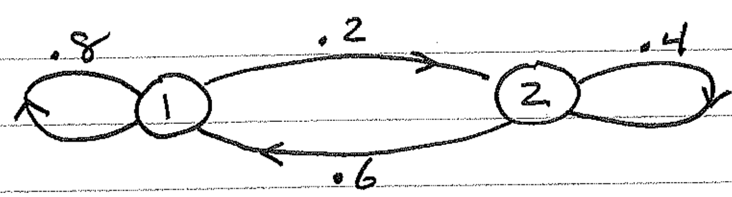
Note the transition probabilities necessarily satisfy

$$\sum_{j=1}^{m} P_{ij} = 1 \quad \text{for all } i \in S$$



Ex. $m = 2$, $S = \{1, 2\}$

$$P_{11} = .8, \quad P_{12} = .2, \quad P_{21} = .6, \quad P_{22} = .4$$



state transition diagram

Suppose $X_0 = 1$. Find $P(X_2 = 2)$.

By the total probability theorem

$$P(X_2 = 2) = P(X_2 = 2 \mid X_1 = 1) \cdot P(X_1 = 1)$$
$$\qquad + P(X_2 = 2 \mid X_1 = 2) \cdot P(X_1 = 2)$$

$$= p_{12} \cdot P(X_1=1) + p_{22}\, P(X_1=2).$$

note

$$P(X_1=1) = P(X_1=1 \mid X_0=1) \cdot \overbrace{P(X_0=1)}^{1}$$
$$+ P(X_1=1 \mid X_0=2) \cdot \overbrace{P(X_0=2)}^{0}$$

$$= p_{11}$$

Also

$$P(X_1=2) = P(X_1=2 \mid X_0=1) = p_{12}$$

Thus

$$P(X_2=2) = p_{12} \cdot p_{11} + p_{22} \cdot p_{12} = p_{11}p_{12} + p_{12}p_{22}$$
$$= (.8)(.2) + (.2)(.4) = .16 + .08 = \boxed{.24}$$

Similarly

$$P(X_2=1) = p_{11} \cdot p_{11} + p_{12}\, p_{21} = (.8)^2 + (.2)(.6) =$$
$$= (.64) + (.12) = \boxed{.76}$$

Defn

A probability matrix is a square ($n \times n$) matrix in which each row sums to 1.
i.e. if $A = (a_{ij})$, then

$$\sum_{j=1}^{n} a_{ij} = 1 \quad \text{for all } 1 \le i \le n$$

Exercise

Prove that the product of two probability matrices is also a probability matrix.

In the same manner, if $X_0 = 2$, then

$$P(X_2 = 1) = P_{21}P_{11} + P_{22}P_{21} = (.6)(.8) + (.4)(.6)$$

$$= .48 + .24 = \boxed{.72}$$

and

$$P(X_2 = 2) = P_{21}P_{12} + P_{22}P_{22} = (.6)(.2) + (.4)^2$$

$$= .12 + .16 = \boxed{.28}$$

We can summarize all these calculations with the <u>transition probability matrix</u>.

$$R = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} = \begin{pmatrix} .8 & .2 \\ .6 & .4 \end{pmatrix}$$

observe that

$$R^2 = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} = \begin{pmatrix} P_{11}^2 + P_{12}P_{21} & P_{11}P_{12} + P_{12}P_{22} \\ P_{21}P_{11} + P_{22}P_{21} & P_{21}P_{12} + P_{22}^2 \end{pmatrix}$$

$$= \begin{pmatrix} .76 & .24 \\ .72 & .28 \end{pmatrix}$$

Thus the ij^{th} element in R^2 is the probability of going from state i to state j in exactly 2 steps.

Defn

The <u>n-step transition probabilities</u> are

$$r_{ij}(n) = P(X_n = j \mid X_0 = i)$$

note: same as $P(X_{n+k} = j \mid X_k = i)$

for any $i, j \in S$ and $n \geq 0$. Thus $r_{ij}(n)$ is the probability of going from state i to state j in n steps.

Observe that $r_{ij}(1) = p_{ij}$. For $n \geq 2$ we have the following recurrence relation.

$$\boxed{r_{ij}(n) = \sum_{k=1}^{m} r_{ik}(n-1)\, p_{kj} \qquad (i, j \in S,\ n \geq 2)}$$

These are the <u>Chapman-Kolmogorov equations</u>

<u>Proof</u>: By total probability, condition on the event $X_0 = i$

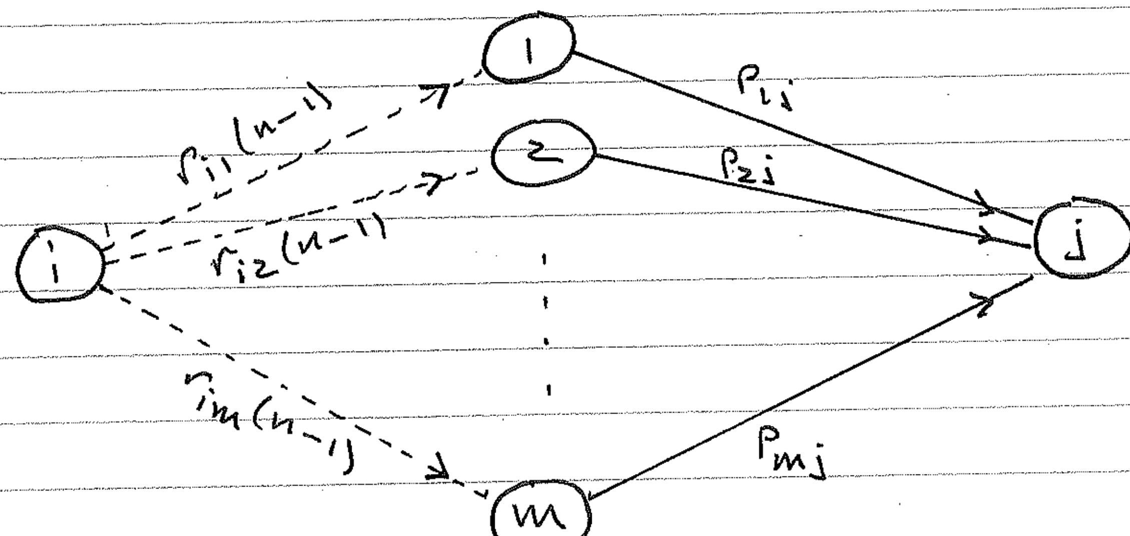
$$\begin{aligned} r_{ij}(n) &= P(X_n = j \mid X_0 = i) \\ &= \sum_{k=1}^{m} P(X_n = j \mid X_{n-1} = k, X_0 = i) \cdot P(X_{n-1} = k \mid X_0 = i) \\ &= \sum_{k=1}^{m} P(X_{n-1} = k \mid X_0 = i)\, P(X_n = j \mid X_{n-1} = k) \end{aligned}$$

Hence

$$r_{ij}(n) = \sum_{k=1}^{m} r_{ik}(n-1)\, p_{kj}$$

as required. ∎

In this recurrence, $r_{ij}(n)$ is computed by conditioning on every possible predecessor $k \in S = \{1, 2, \ldots, m\}$ to j.



Notice that the Chapman-Kolmogorov eqns. say that $r_{ij}(n)$ is the ij^{th} entry in the n^{th} power of R. i.e.

$$\boxed{\left(r_{ij}(n)\right)_{i,j \in S} = R^{n-1} R = R^{n}}$$

Matrix form of C-K eqns.

Ex

Returning to the previous example, we have

$$R^3 = \begin{pmatrix} .76 & .24 \\ .72 & .28 \end{pmatrix}\begin{pmatrix} .8 & .2 \\ .6 & .4 \end{pmatrix} = \begin{pmatrix} .752 & .248 \\ .744 & .256 \end{pmatrix}$$

so $P(X_3=1 \mid X_0=2) = r_{21}(3) = .744$. One calculates with some effort

$$R^{10} = \begin{pmatrix} .7500000256 & .2499999744 \\ .7499999232 & .2500000768 \end{pmatrix}$$

so that $P(X_{10}=1 \mid X_0=2) = r_{21}(10) = .7499999232$.

It can be shown that

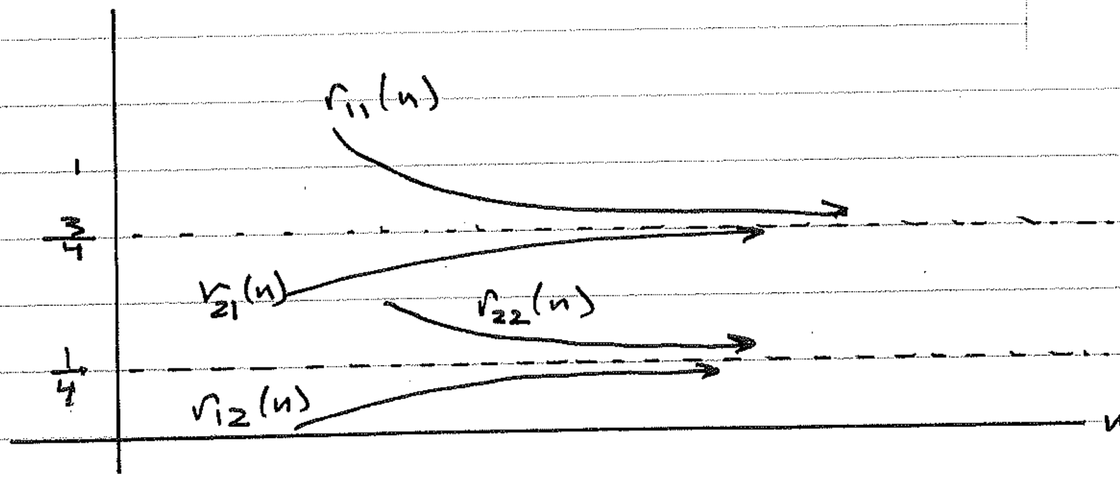
$$\lim_{n\to\infty} R^n = \begin{pmatrix} .75 & .25 \\ .75 & .25 \end{pmatrix}$$

i.e.

$$r_{11}(n) \to \frac{3}{4}^{+} \qquad r_{12}(n) \to \frac{1}{4}^{-}$$

$$r_{21}(n) \to \frac{3}{4}^{-} \qquad r_{22}(n) \to \frac{1}{4}^{+}$$

as $n \to \infty$.

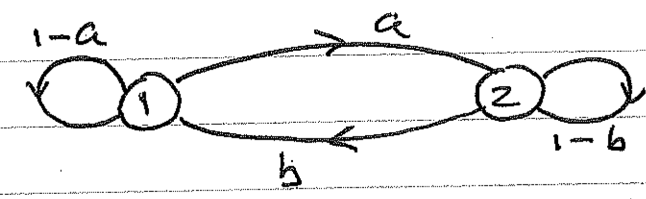


Thus, if we wait a long time, then

$$P(X_n = 1) \approx \frac{3}{4} \text{ and } P(X_n = 2) \approx \frac{1}{4}$$

independently of the initial state X_0. These are called <u>steady state probabilities</u>. Such behavior is not unusual in a Markov Chain.

<u>Ex.</u>



$$R = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

one can show (using linear algebra)

$$\lim_{n\to\infty} P^n = \begin{pmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{pmatrix}$$

and hence

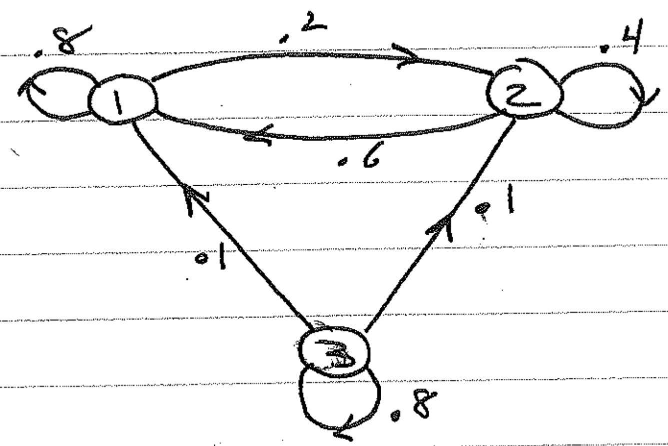
$$\lim_{n\to\infty} P(X_n = 1) = \frac{b}{a+b}$$

and

$$\lim_{n\to\infty} P(X_n = 2) = \frac{a}{a+b}$$

again regardless of X_0.

Ex



Thus

$$P = \begin{pmatrix} .8 & .2 & 0 \\ .6 & .4 & 0 \\ .1 & .1 & .8 \end{pmatrix}$$

one can show

$$\lim_{n\to\infty} R^n = \begin{pmatrix} .75 & .25 & 0 \\ .75 & .25 & 0 \\ .75 & .25 & 0 \end{pmatrix}$$

hence

$$P(X_n = 1) \to \frac{3}{4} \quad \text{as } n \to \infty$$

$$P(X_n = 2) \to \frac{1}{4} \quad \text{as } n \to \infty$$

$$P(X_n = 3) \to 0 \quad \text{as } n \to \infty$$

regardless of the initial state X_0.

All of these ~~limits~~ limits can be computed using theorems of linear algebra, but is there another way that does not use L.A. topics? yes.

7.2 Classification of states

<u>Defn</u>
We say that state j is <u>accessible</u> from state i (or that j is <u>reachable</u> from i) iff

$$r_{ij}(n) > 0 \quad \text{for some } n \geq 1$$

Equivalently, there exists a sequence of states $i, i_1, i_2, \ldots, i_{n-1}, j$ such that the transitions

$$i \to i_1 \to i_2 \to \cdots \to i_{n-1} \to j$$

all have positive probability.

<u>notation</u>: $i \dashrightarrow j$

<u>Defn</u>
Let $A(i)$ denote the set of states accessible from i, i.e.

$$A(i) = \{\, j \in S \mid i \dashrightarrow j \,\}$$

Defn

A state $i \in S$ is called *recurrent* if for every j accessible from i, we also have i accessible from j.

Equivalently

$$j \in A(i) \quad \text{iff} \quad i \in A(j)$$

i.e.

$$(i) \dashrightarrow (j) \quad \text{iff} \quad (j) \dashrightarrow (i)$$

Thus if we depart a recurrent state, there is a positive probability that we return, and so after a *long enough* time, we will return.

Hence a recurrent state, if it is visited at all, *will be visited infinitely often*.

Defn

A state i is called *transient* iff it is not recurrent. Thus i is transient iff there exists a state j with

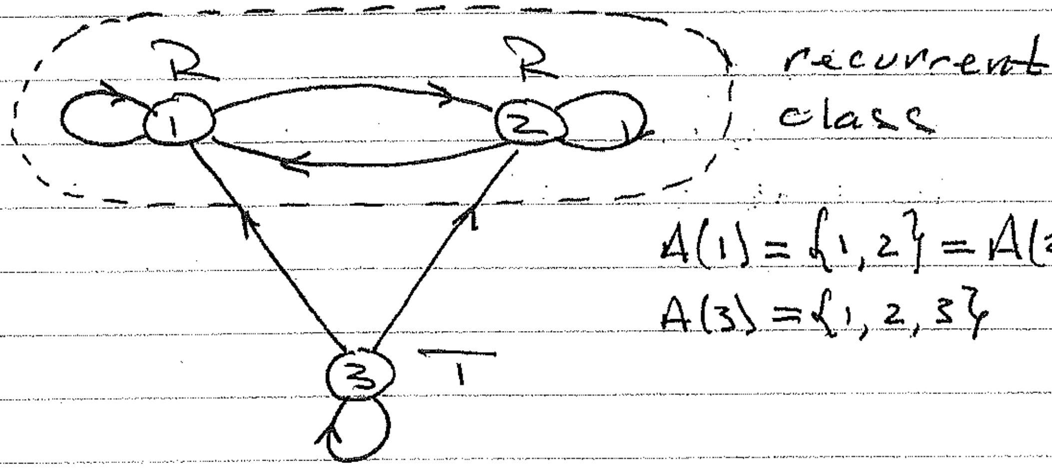
$$j \in A(i) \quad \text{and} \quad i \notin A(j)$$

Thus after a visit to a transient state i, there is a positive probability of entering such a state j. Given enough time, this is certain to happen.

Hence a transient state will only be visited a finite number of times.

<u>Note</u>: transience and recurrence are determined by the arcs in a state transition diagram, not by the actual probabilities p_{ij} on each arc.

Ex.



$A(1) = \{1, 2\} = A(2)$
$A(3) = \{1, 2, 3\}$

Ex.