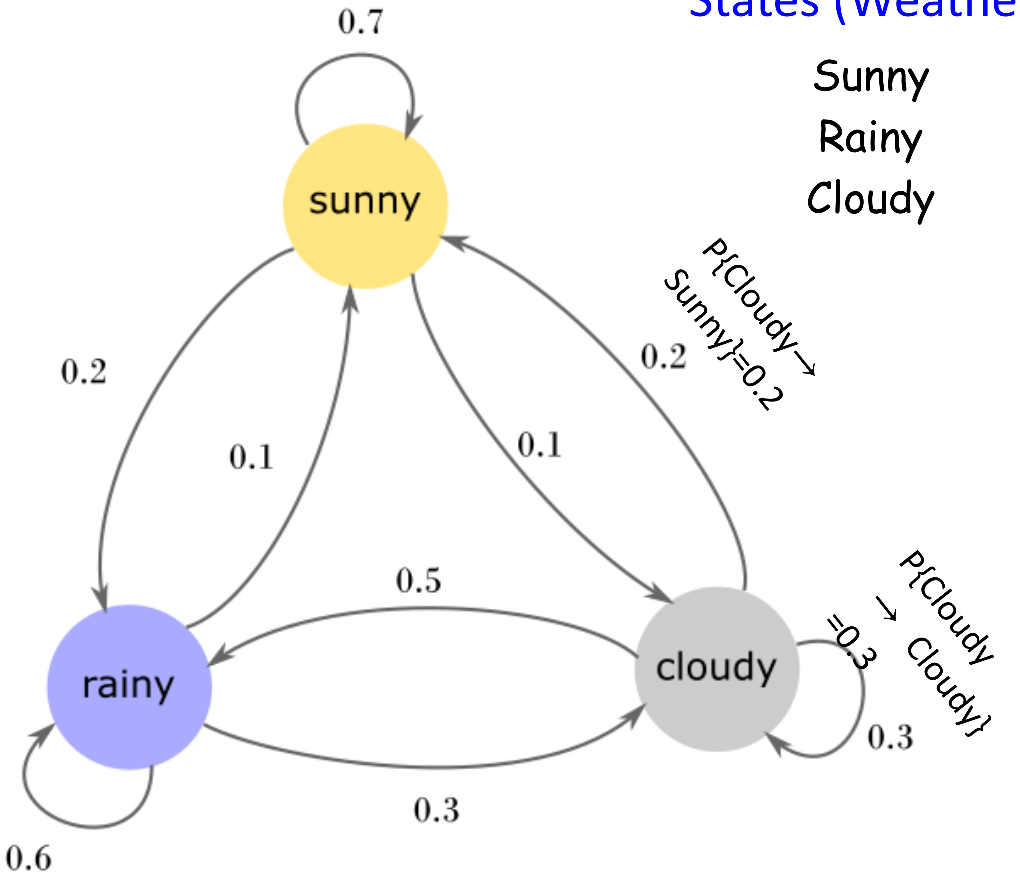
Discrete Time Markov Chains

Mini Project for Module –

Automatic Prediction of Control Laws of an Aircraft using the Viterbi Algorithm

Markov Chain (Definition)

A Markov Chain (or equivalently, a Markov Process) is a stochastic (i.e., random) process describing a series of possible events, where the probability of an event at a given instant depends **only** on the outcome of the previous event.

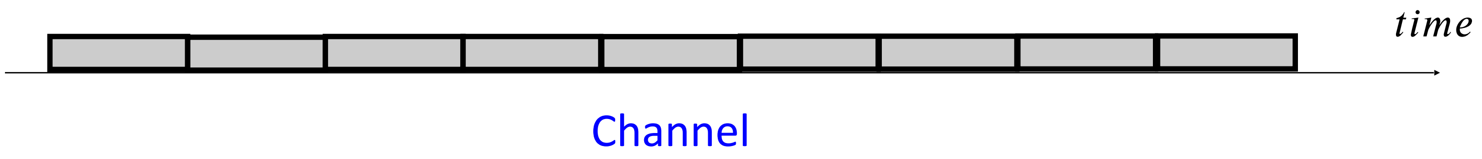


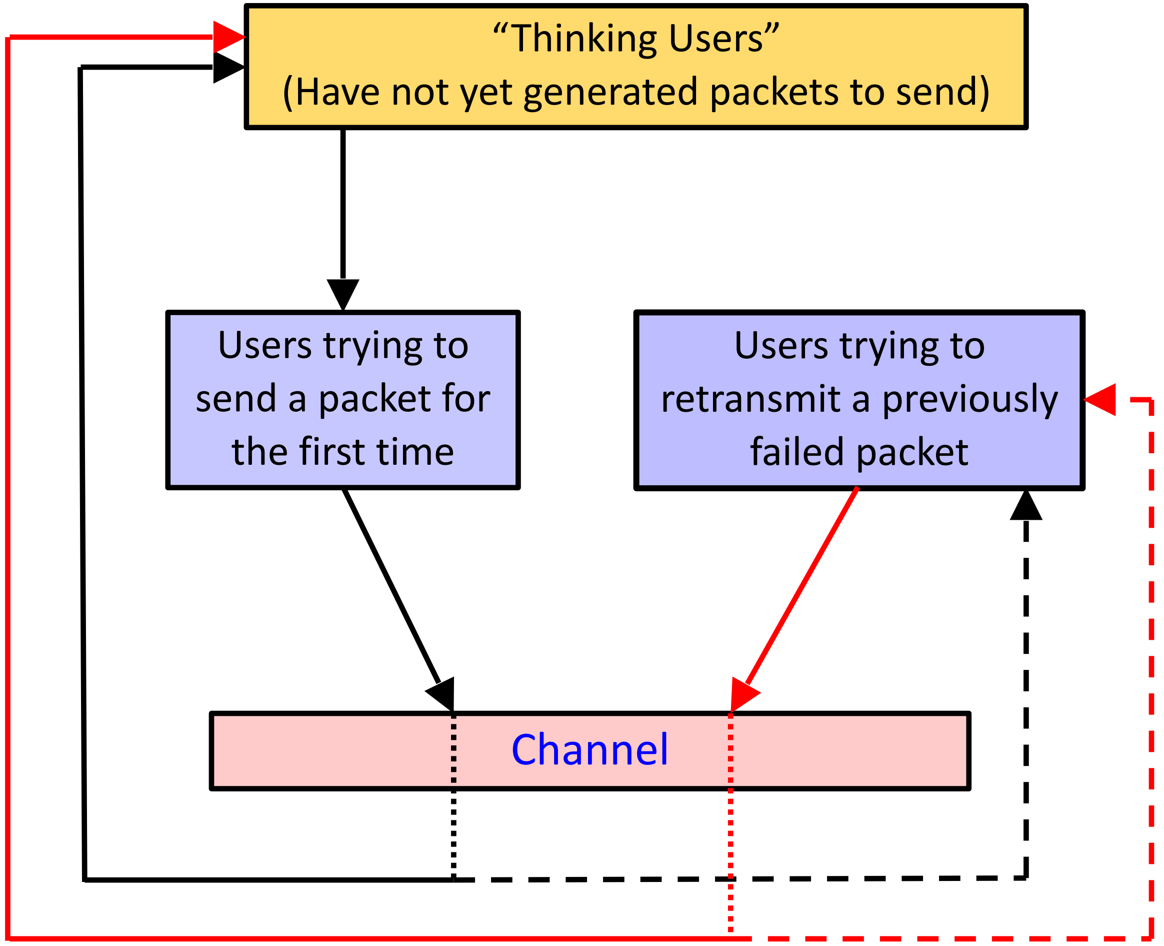
States (Weather)

Sunny
Rainy
Cloudy

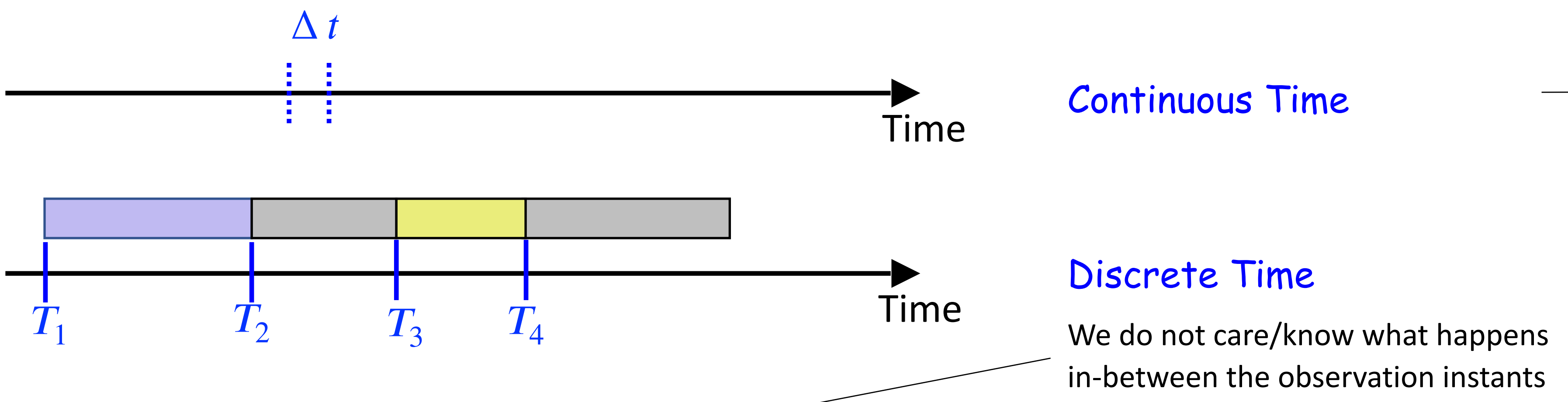
Knowing the weather right now, we can use a model like this to make a prediction about what the weather is likely to be the next time we check it

A (Simplified) Networking Example





- Slot can carry one packet of data correctly; transmission fails if multiple users transmit in a slot
- Users with a packet of data to send randomly decide whether to transmit in a slot
- Transmissions which fail are retried in another randomly chosen slot until success
- A user does not generate a new packet until the previous one is transmitted



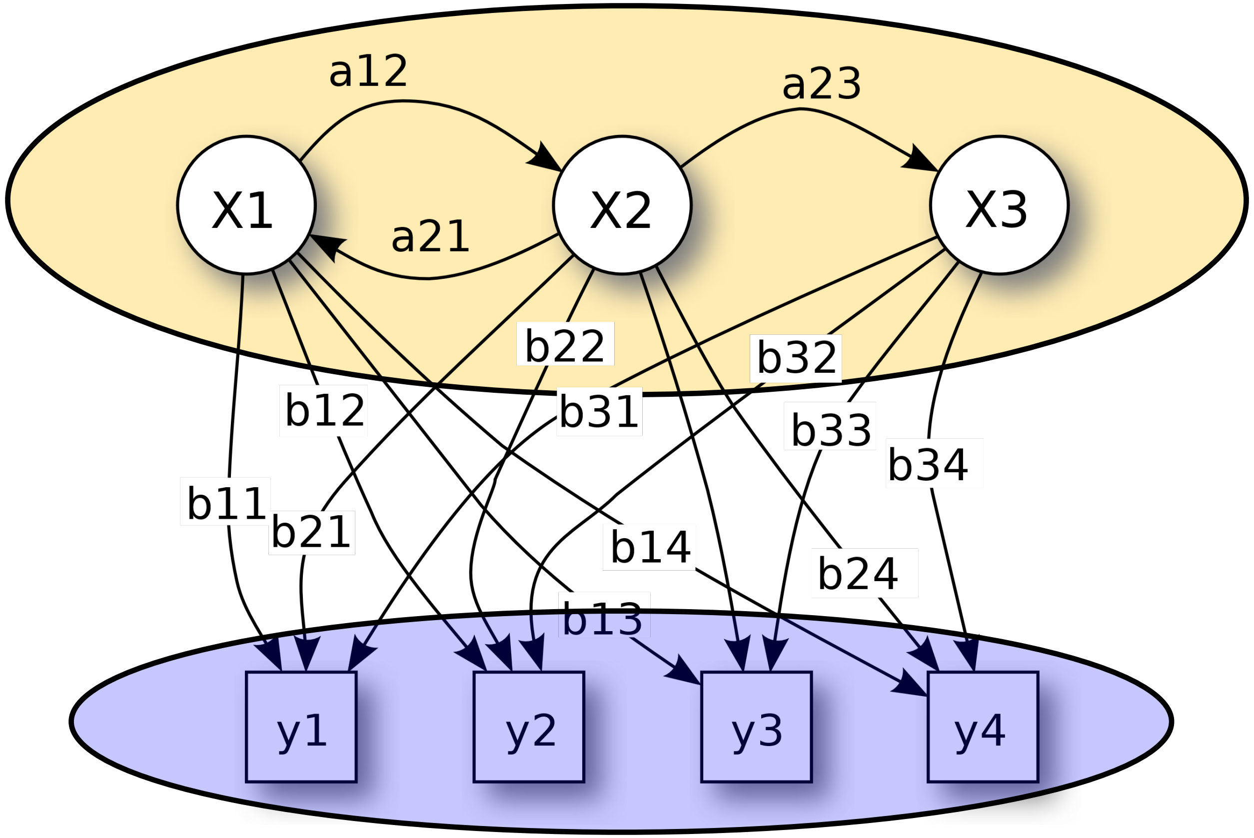
We consider only Discrete Time Markov Chains **in this course**, where the outcomes are events occurring at discrete instants of time. The occurrence of a particular event at any instant of time depends only on the previous event (at the previous instant of time) and does not depend directly on any earlier event.

One can also have *Continuous Time Markov Chains* where states move continuously through time rather than in discrete time steps

Something Interesting!

Hidden Markov Model (HMM)

Underlying Markov Model which is not directly observable (i.e., Hidden)



Observation Layer visible from outside

Hidden Markov models are known for their applications to thermodynamics, statistical mechanics, physics, chemistry, economics, finance, signal processing, information theory, pattern recognition etc..

Discrete Time Markov Chain

We denote a sequence of events as $\{X_n\}_{n\in I, n\geq 0}$, where the subscript n is a non-negative integer that indexes time, X_n is a random variable which may take values from a set of possible outcomes (events) $\mathcal{S} = \{x_0, x_1, \ldots\ldots\}$

Then $\{X_n\}$ describes a Markov Chain when

$$P(X_n = x_n \mid X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots\ldots, X_0 = x_0) = P(X_n = x_n \mid X_{n-1} = x_{n-1})$$

A Markov Process is therefore characterized to be a *Memoryless Process*

Current state depends only on the previous state and not on states earlier than that

Something useful (on Markov Chains)

$$P(X_n = x_n, X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots\ldots, X_0 = x_0)$$

$$= P(X_n = x_n \mid X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots\ldots, X_0 = x_0) P(X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots\ldots, X_0 = x_0)$$

$$= P(X_n = x_n \mid X_{n-1} = x_{n-1}) P(X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots\ldots, X_0 = x_0) \qquad \text{Markov Property}$$

$$= P(X_n = x_n \mid X_{n-1} = x_{n-1}) P(X_{n-1} = x_{n-1} \mid X_{n-2} = x_{n-2}) P(X_{n-2} = x_{n-2}, \ldots\ldots, X_0 = x_0)$$

....

....

$$= \left[\prod_{k=1}^{n} P(X_k = x_k \mid X_{k-1} = x_{k-1})\right] P(X_0 = x_0)$$

1. Probability of a Sequence of States (LHS) is the continued product of state transition probabilities multiplied by the initial state probability (RHS)
2. Easy to calculate this using the logarithms of each term

Stochastic Matrix (or Probability Transition Matrix)

$$\mathbb{P} = \{p_{ij}\}$$

$$p_{ij} = P(\text{transition from state } i \text{ to state } j \text{ in one step})$$
$$= P(X_{n+1} = j \mid X_n = i) \qquad \forall n$$

Property of **Stationarity** or **Time-Homogenity :** This implies that the (one step) transition probability depends only on the end-states i and j but not on when the transition actually occurs, i.e., it does not depend on *n.*

The probability axioms also imply that

$$\sum_{\forall j} p_{ij} = 1$$

since system **has to go** to some state *j* from state *i*

Example: Gambler's Ruin

Gambling game where in each step, the player wins Rs. 1 with probability p=0.4 or loses Rs. 1 with probability $p^{\mathbb{C}} = (1 - p) = 0.6$

The player's strategy is to quit when he/she makes a "fortune" of \$ 10. The player gets thrown out of the casino when he/she runs out of money

Let X_n=Amount of money the player has after n plays

Then for $i,\ i = \{1,2,\ \ldots\ldots 9\}$, we have

$$P\left(X_{n+1} = i+1 \mid X_n = i, X_{n-1} = i_{n-1}, \ldots\ldots, X_0 = i_0\right)$$
$$= P\left(X_{n+1} = i+1 \mid X_n = i\right) = p_{i,i+1} = 0.4$$

and similarly, $p_{i,i-1} = 0.6$

We also observe that -

$p_{0,0} = 1$ cannot play any more when no money is left

and $p_{10,10} = 1$ cannot accumulate more than Rs. 10

States 0 and 10 are Absorbing States

Absorbing States are states from which one cannot move to another state

Example: Gambler's Ruin ………… *continued…………*

The Stochastic Matrix (also called the **State Transition Matrix**) for this is given by -

$$
\mathbb{P} = \begin{array}{c} \\ i=0 \\ i=1 \\ i=2 \\ \cdot \\ \cdot \\ \cdot \\ i=9 \\ i=10 \end{array}
\begin{array}{c}
\begin{array}{ccccccc} j=0 & j=1 & j=2 & \cdot & \cdot & \cdot & j=10 \end{array} \\
\begin{pmatrix}
1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\
0.6 & 0 & 0.4 & \cdot & \cdot & \cdot & 0 \\
0 & 0.6 & 0 & 0.4 & \cdot & \cdot & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdot & \cdot & 0.6 & 0 & 0.4 \\
0 & 0 & \cdot & \cdot & 0 & 0 & 1
\end{pmatrix}
\end{array}.
$$

Note that these give the One-Step State Transition Probabilities from state i to state j for this Markov Chain

Absorbing State

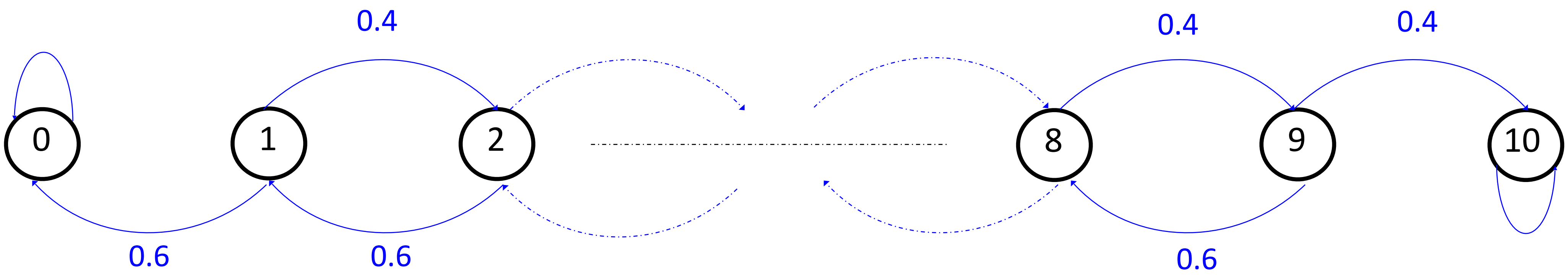
You can enter the state, but you can never leave it!

Sort of like the Eagles' "Hotel California" song!

State Transition Matrix

$$\mathbb{P} = \begin{array}{c} \\ i=0 \\ i=1 \\ i=2 \\ \cdot \\ \cdot \\ \cdot \\ i=9 \\ i=10 \end{array} \begin{array}{c} \begin{array}{ccccccc} j=0 & j=1 & j=2 & \cdot & \cdot & \cdot & j=10 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0.6 & 0 & 0.4 & \cdot & \cdot & \cdot & 0 \\ 0 & 0.6 & 0 & 0.4 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & 0.6 & 0 & 0.4 \\ 0 & 0 & \cdot & \cdot & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

The state transition matrix can also be represented by the following **State Transition Diagram**



This should bring back fond memories of "The Misadventures of Squeaky"

Question: Consider a simplified version of the earlier matrix (*Maximum Amount of Win*=3)

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 \\ 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

One Step Transition Probability Matrix

$$P^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.6 & 0.24 & 0 & 0.16 \\ 0.36 & 0 & 0.24 & 0.4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Two Step Transition Probability Matrix

$$P^3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.744 & 0 & 0.096 & 0.16 \\ 0.36 & 0.144 & 0 & 0.496 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Three Step Transition Probability Matrix

See next slide for the corresponding state diagrams

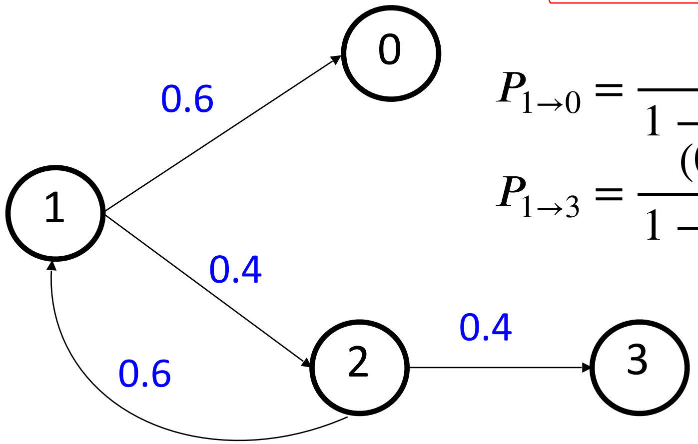
…………..

$$\begin{array}{c|cccc} & 0 & 1 & 2 & 3 \\ \hline 0 & 1 & 0 & 0 & 0 \\ 1 & 0.6 & 0 & 0.4 & 0 \\ 2 & 0 & 0.6 & 0 & 0.4 \\ 3 & 0 & 0 & 0 & 1 \end{array}$$

States 0 and 3 are absorbing states

System in those states will stay there for ever

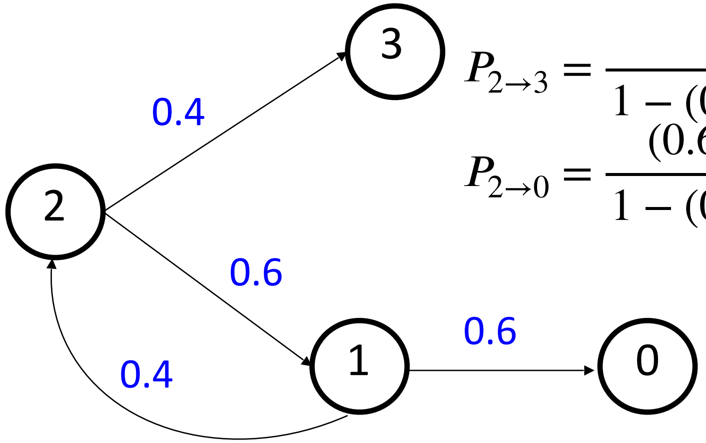
System starting in States 1 or 2 will move from one state to another but will eventually land in an absorbing state



$$(0.6) + (0.4\times 0.6)(0.6) + (0.4\times 0.6)^2(0.6)\ldots\ldots$$

$$P_{1\to 0} = \frac{0.6}{1-(0.6)(0.4)} = 0.7895$$

$$P_{1\to 3} = \frac{(0.4)(0.4)}{1-(0.6)(0.4)} = 0.2105$$



$$P_{2\to 3} = \frac{0.4}{1-(0.6)(0.4)} = 0.5263$$

$$P_{2\to 0} = \frac{(0.6)(0.6)}{1-(0.6)(0.4)} = 0.4737$$

If I start in state 1 (state 2), what will be the average number of steps for me to reach the absorbing state 0 (state 3) ? ………… *think about how you can calculate these!*

The answers are $T_{1\to 0} = 1.2881$ $T_{1\to 3} = 0.5540$ $T_{2\to 0} = 1.2465$ $T_{2\to 3} = 0.8587$

Multi-Step Transition Probabilities may also be given as the probabilities of transition between states in more than one step,

e.g., the probability of transition from state i to state j in m ($m > 1$) steps -

$$p_{i,j}^{(m)} = p^{(m)}(i, j) = P\left\{X_{n+m} = j \mid X_n = i\right\}; \quad m > 1$$

$(i,\ j)^{th}$component of $\mathbb{P}^m$

In general,

$\mathbb{P} = \left\{p_{ij}\right\}$ p_{ij}=P{going from i to j in 1 step}

$\mathbb{P}^{(2)} = \left\{p_{ij}^{(2)}\right\}$ $p_{ij}^{(2)}$=P{going from i to j in 2 steps}

…….

…….

$\mathbb{P}^{(n)} = \left\{p_{ij}^{(n)}\right\}$ $p_{ij}^{(n)}$=P{going from i to j in n steps}

…….

…….

Example: *Social Motility*

Let X_n represent the social class of a family in the n^{th} generation where we consider that there are broadly three social groups based on income, viz., lower=1, middle=2, upper=3.

Based on a certain demographic and per capita income analysis, the motility within this society was captured succinctly by the following stochastic matrix.

$$\mathbb{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix} \end{array}$$

State Transition Matrix

Example: *Social Motility …… continued…….*

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix} \end{array}$$

If Ginny's parents (X_0) were of the *middle income* (2) class, what is the probability that Ginny herself (X_1) belongs to the *upper income* (3) class and her children (X_2) belong to the *lower income* (1) class. We essentially need to find $P(X_2 = 1, X_1 = 3 \mid X_0 = 2)$

$$P(X_2 = 1, X_1 = 3 \mid X_0 = 2)$$

………… applying Baye's Rule ………….. $P\{C, B \mid A\} = P\{C \mid B, A\} P\{B \mid A\}$

$$= P(X_2 = 1 \mid X_1 = 3, X_0 = 2) P(X_1 = 3 \mid X_0 = 2)$$

………… using the Markov Property …………..

$$= P(X_2 = 1 \mid X_1 = 3) P(X_1 = 3 \mid X_0 = 2)$$

$$= p_{31} p_{23}$$

Note that this may be found directly by multiplying the probabilities of the successive transitions 2→3 and 3 →1, i.e. 0.2x0.2 = 0.04

Read your online textbook for a somewhat longer but more detailed explanation

$$\mathbb{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix} \end{array}$$

Example: *Social Motility …… continued……*

$$\mathbb{P}^2 = \begin{pmatrix} 0.57 & 0.28 & 0.15 \\ 0.4 & 0.39 & 0.21 \\ 0.34 & 0.4 & 0.26 \end{pmatrix}$$

We can directly use the elements of this matrix to find the probability of a particular transition happening in TWO Generations

Example:

(a) Probability of going from Lower (1) to Upper (3) in two generations is given by the (1, 3) element of $\mathbb{P}^2$, i.e. 0.15

$$p_{1,2}p_{2,3} + p_{1,1}p_{1,3} + p_{1,3}p_{3,3}$$
$$= 0.2\times0.2 + 0.7\times0.1 + 0.1\times0.4$$
$$= 0.15$$

(b) Probability of going from Upper (3) to Lower (1) in two generations is given by the (3, 1) element of $\mathbb{P}^2$, i.e. 0.34

$$p_{3,2}p_{2,1} + p_{3,1}p_{1,1} + p_{3,3}p_{3,1}$$
$$= 0.4\times0.3 + 0.2\times0.7 + 0.4\times0.2$$
$$= 0.34$$

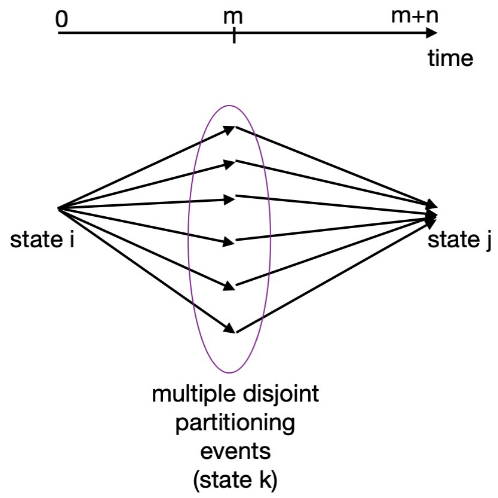
Chapman-Kolmogorov Equation

Multi-Step Transition Probabilities for a Markov Chain

$$p_{i,j}^{m+n} = \sum_{\forall k} p_{i,k}^{m} p_{k,j}^{n}$$

Probability of transitioning from state i to state j by passing through an intermediate state k.

The state k can be any state of the system.



$$\begin{aligned}
p_{i,j}^{m+n} &= P\left(X_{m+n} = j \mid X_0 = i\right) \\
&= \sum_{k \in S} P\left(X_{m+n} = j, X_m = k \mid X_0 = i\right) \\
&= \sum_{k \in S} P\left(X_{m+n} = j \mid X_m = k, X_0 = i\right) P\left(X_m = k \mid X_0 = i\right) \\
&= \sum_{k \in S} p_{k,j}^{n} p_{i,k}^{m} \\
&= \sum_{k \in S} p_{i,k}^{m} p_{k,j}^{n}
\end{aligned}$$

This is exactly what we saw happen in the example of a 2-step probability calculation that we did in the last slide

Chapman-Kolmogorov Equation

These may also be written in the following forms which have special names –

Forward C-K Equation

$$p_{ij}^{(n+1)} = \sum_k p_{ik}^{(n)} p_{kj} \quad \text{for } n=1,2\ldots\ldots$$

i to k in n steps and k to j in *one* step

Backward C-K Equation

$$p_{ij}^{(n+1)} = \sum_k p_{ik} p_{kj}^{(n)} \quad \text{for } n=1,2\ldots\ldots$$

i to k in *one* steps and k to j in n steps

Distribution of the States of a Markov Chain

Given (a) A Markov model with a stochastic matrix $\mathbb{P}$ (i.e. the state transition matrix),
and (b) The initial state probability distribution

We may be interested in finding the state probability distribution
(a) At some later instant of time (i.e., after n steps, $n \geq 10$)
and/or (b) After a long time (i.e., $n \to \infty$)

Consider a Markov Chain with k states $\{s1,\ s2,\ \ldots\ldots\ldots,\ s_k\}$ and initial state distribution

$$\mu^{(0)} = \left(\mu_1^{(0)}, \mu_2^{(0)}, \ldots\ldots, \mu_k^{(0)}\right) = \left(P(X_{01} = s_1), P(X_{02} = s_2), \ldots\ldots, P(X_{0k} = s_k)\right) \qquad \sum_{i=1}^{k} \mu_i^{(0)} = 1$$

Normalization Condition

Here, the 0's represent the initial time instant and the indices $i = \{1,2,\ldots,k\}$ refer to the states

Then, $\vec{\mu}^{(n)} = \vec{\mu}^{(0)} \mathbb{P}^n$ is the state distribution after n steps and by taking the limit $n \to \infty$, we can find the state distribution after a long time.

Note that the state distribution after a long time, $\mathrm{i.e.}$, $n \to \infty$, will also be the system state at equilibrium, $\vec{\mu}^{(\infty)}$

Therefore, $$\vec{\mu}^{(\infty)} = \vec{\mu}^{(\infty)} \mathbb{P}$$

This equation can be directly solved to find the state distribution after a long time, i.e., when the system has reached its equilibrium

Example: *A Simple Weather Model*

The weather is either RAINY (R) or SUNNY (S). It stays the same tomorrow as it is today with probability 0.75 and changes with probability 0.25.

Therefore,

$P(R\rightarrow R)=P(S\rightarrow S)=0.75$

and $P(R\rightarrow S)=P(S\rightarrow R)=0.25$

This gives us the stochastic matrix

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{cc} s & r \end{array} \\ \begin{array}{c} s \\ r \end{array} & \begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix} \end{array}$$

We can then calculate -

$$\vec{\mu}^{(1)} = \vec{\mu}^{(0)}\mathbb{P} = (1\ \ 0)\begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix} = (0.75\ \ 0.25)$$

$$\vec{\mu}^{(2)} = \vec{\mu}^{(0)}\mathbb{P}^2 = \vec{\mu}^{(1)}\mathbb{P} = (0.625\ \ 0.375)$$

.

.

.

$$\vec{\mu}^{(\infty)} = \vec{\mu}^{(0)}\mathbb{P}^{\infty} = (0.5\ \ 0.5)$$

This is what happens as $t \rightarrow \infty$.
We can also calculate this by using the results of the last slide!

Example: *Markov Model of a Badminton Game (Deciding the Winning Strategy)*

Consider a Markov model of a game of badminton. For simplicity, let us consider that a player chooses to play one of three shots, viz., smash (S), drop (D), and lift (L). The objective is to devise a winning strategy given a match situation. Based on the data generated over several games, the following table lists the probability (P) of a return shot played by a player given a certain type of shot played by their opponent.

Played Shot	Return Shot	P
D	D	1/3
D	L	1/3
D	S	0
L	D	1/5
L	L	1/5
L	S	2/5
S	L	2/5
S	D	1/5
S	S	0

Questions:

1. Identify an appropriate state space for the Markov Model
2. Construct the stochastic matrix $\mathbb{P}$
3. Given a *lifted* serve, what are the chances that there is a winner in THREE shots
4. In a rally, if a player receives a lift from their opponent, which shot option maximizes his chance of winning the rally in the return shot?

Example: *Markov Model of a Badminton Game ……….. continued………..*

1. State Space, $\mathscr{S} = \{D,\ L,\ S,\ W\}$ where *W* is the *winning* shot

2. The stochastic matrix $\mathbb{P}$ may be computed using the data table

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{cccc} D & L & S & W \end{array} \\ \begin{array}{c} D \\ L \\ S \\ W \end{array} & \begin{bmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 1/5 & 1/5 & 2/5 & 1/5 \\ 1/5 & 2/5 & 0 & 2/5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

3. Looking at the matrix $\mathbb{P}^3$ the probability of the $L \rightarrow W$ transition (in three shots) is 0.6142

$$\mathbb{P}^3 = \begin{matrix} 0.1215 & 0.1481 & 0.0711 & 0.6593 \\ 0.1316 & 0.1476 & 0.1067 & 0.6142 \\ 0.1102 & 0.1422 & 0.0587 & 0.6889 \\ 0 & 0 & 0 & 1 \end{matrix}$$

4. Starting from $X_0=L$, we want to find the choice of X_1 such that the the probability of $X_2=W$ is the highest

   i.e. X_1 such that $P(X2 = W, X1 \mid X0 = L)$ is the highest

   We find that $(p_{LD} \times p_{DW},\ p_{LS} \times p_{SW},\ p_{LL} \times p_{LW}) = (0.06667, 0.16, 0.04)$

   Therefore, playing a “Smash” in return for the “Lift” would be the best strategy to win this rally

Recurring Events in a Markov Chain

In a finite state system with stochastic transitions between the states, some states (i.e., events) of the Markov Chain may be visited multiple times, repeatedly, where the number of steps between successive visits may itself be a random number.

Example:

Consider the teller in a bank where customers queue for service and are served one by one. The customers arrive randomly and require a random amount of time to be serviced. The state at any instant of time is given by the number of customers in the system at that time.

The manager of the bank would be particularly interested in the probability of the teller being idle (because there are no customers in the system). This state 0 is then a recurrent state whose recurrence times indicates the times when the teller is left idle (until the arrival of the next customer.) *The teller may ask for a raise if he or she is being worked too hard, but he/she may be getting paid too much if not worked hard enough!*

We should point out that instead of a *Discrete Time Markov Chain*, it would be better to model this as a *Continuous Time Markov Chain* which is something that we will discuss later.

Hitting Probability for a State (Definition)

Let $\{X_n\}_{0,n\in\{0,\,\mathbb{Z}^+\}}$ represent a Markov Chain with state space $\mathcal{S}$. Let $A \subset \mathcal{S}$. We further define –

$T_A := \min\{n \geq 0 \mid X_n \in A\}$ = first time the chain hits **set** $\boldsymbol{A}$ starting from a **state outside (or inside)** $\boldsymbol{A}$

with $T_A = 0 \text{ if } X_0 \in A$ and $T_A = \infty \text{ if } \{n \geq 0 \,|\, X_n \in A\} = \{\}$ (return to A *in an infinite number of steps*)

Then the *probability of hitting state* l *for the first time, after time* T_A, *starting from state* k at time 0, is

$$g_k(l) = P\left(X_{T_A} = l \mid X_0 = k\right)$$

k=initial state, l=final state

For the example of the bank teller, if we find that $g_s(0) = 0$ for any $s > 0$, then the likelihood of the teller becoming idle starting from any state is 0. *The teller is definitely overworked, and the manager needs to increase the number of tellers.*

Calculating the "hitting probability" iteratively

Let $k \in \mathcal{S} \backslash A$ where "$\mathcal{S} \backslash A \equiv \mathcal{S} - \{A\}$" i.e., the set $\mathcal{S}$ minus the contents of set A

For $T_A \geq 1$ given $X_0 = k$, we have -

$$g_k(l) = P\left\{X_{T_A} = l \mid X_0 = k\right\} = \sum_{m \in S} P\left\{X_{T_A} = l, X_1 = m \mid X_0 = k\right\}$$ Sum over partitioning events

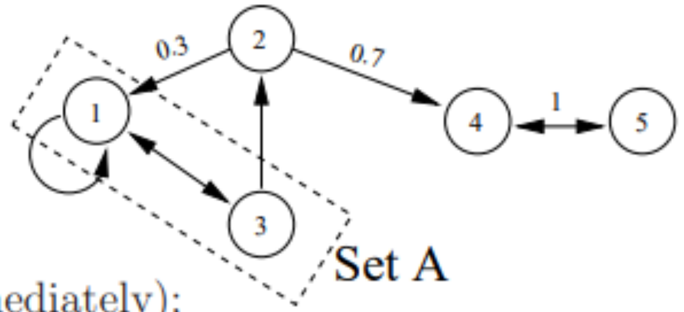
$$= \sum_{m \in S} P\left\{X_{T_A} = l \mid X_1 = m, X_0 = k\right\} P\left\{X_1 = m \mid X_0 = k\right\}$$ Applying the law of total probability

$$= \sum_{m \in S} P\left\{X_{T_A} = l \mid X_1 = m\right\} P\left\{X_1 = m \mid X_0 = k\right\}$$ Markov Property

$$= \sum_{m \in S} g_m(l) p_{km} = \sum_{m \in S} p_{km}\, g_m(l) \qquad k \in S \setminus A, l \in A$$ Rearranging the order of the terms

Transition from state k to state m and then from state m to state l to go from state k to state l

Example: Let set $A = \{1, 3\}$ as shown.



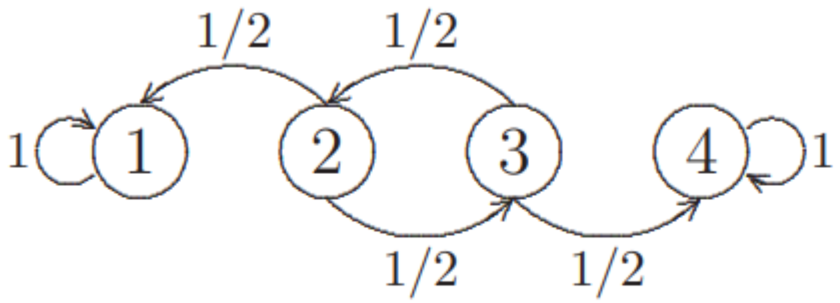
The hitting probability for set A is:

- *1 starting from states 1 or 3*
  (We are starting in set A, so we hit it immediately);

- *0 starting from states 4 or 5*
  (The set $\{4, 5\}$ is a closed class, so we can never escape out to set A);

- *0.3 starting from state 2*
  (We could hit A at the first step (probability 0.3), but otherwise we move to state 4 and get stuck in the closed class $\{4, 5\}$ (probability 0.7).)

Example:



$g_1(4) = 0$ state 4 unreachable from state 1

$$g_2(4) = \frac{1}{2}g_1(4) + \frac{1}{2}g_3(4)$$

$$g_3(4) = \frac{1}{2} + \frac{1}{2}g_2(4)$$

$g_4(4)$=1

Solve, to get

$$g_2(4) = \frac{1}{3}$$

$$g_3(4) = \frac{2}{3}$$

Note that states 1 and 4 are **absorbing states.**

An **Absorbing State** is one where, once the system enters such a state, it cannot subsequently move to any other state.

Iterative Formula for Mean Hitting Times and Mean Absorption Times

Consider the random variable giving the time duration (number of steps) from a state k to the set $A,\ A \subset S$. The mean of this random variable will be –

Mean Hitting Time from state k to A: $h_k(A) = E\left(T_A \mid X_0 = k\right)$ $h_k(A) = 0 \ \ \forall k \in A \subset S$

Further for $\forall k \in S \backslash A$

$$h_k(A) = E\left(T_A \mid X_0 = k\right) = \sum_{m \in S} E\left(T_A, X_1 = m \mid X_0 = k\right)$$

Sum over partitioning events

$$= \sum_{m \in S} E\left(T_A \mid X_1 = m, X_0 = k\right) P\left(X_1 = m \mid X_0 = k\right)$$

Applying the law of total expectation

$$= \sum_{m \in A} \underbrace{E\left(T_A \mid X_1 = m\right)}_{1} p_{km} + \sum_{m \in S \backslash A} \underbrace{E\left(T_A \mid X_1 = m\right)}_{1 + h_m(A)} p_{km}$$

Markov Property

Markov Chain has moved 1 step in both cases, and counting is reset once again for the second case

$$= \sum_{m \in A} p_{km} + \sum_{m \in S \backslash A} \left(\left(1 + h_m(A)\right)\right) p_{km} = \sum_{m \in S} p_{km} + \sum_{m \in S \backslash A} h_m(A)\, p_{km}$$

$$= 1 + \sum_{m \in S \backslash A} \left(h_m(A)\right) p_{km} + \sum_{m \in A} p_{km} h_m(A) \qquad h_m(A) = 0 \text{ for } m \in A \qquad \sum_{i \in S} p_i = 1$$

$$= 1 + \sum_{m \in S} p_{km} h_m(A) \quad \text{for all } k \in S \backslash A$$

First Return Time and its Mean

The time (in number of steps) taken to make a **FIRST RETURN** to a certain state $y \in S$ defined as -

$$T_y^r := \min\{n \geq 1 \mid X_n = y\};\ y \in S$$

with $T_y^r = \infty$ if $X_n \neq y\ \forall n \geq 1$ Note $T_y^r = T_y$ if $X_0 = y$

Using this, we can define the Mean Transit Time $\mu_x(y)$ from state x to state y as

$$\mu_x(y) = E\left(T_y^r \mid X_0 = x\right) \geq 1$$

and derive as before that

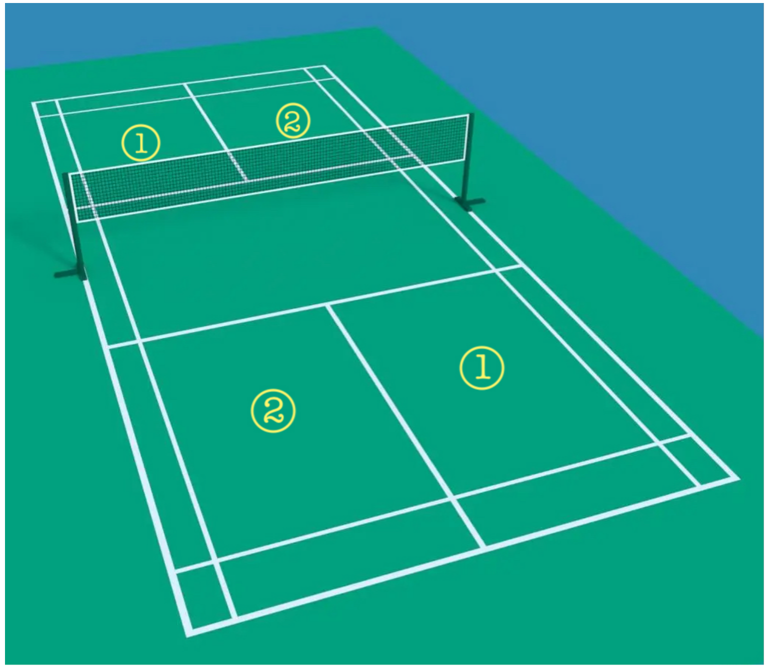
$$\mu_x(y) = 1 + \sum_{m \in S,\ m \neq y} p_{xm}\mu_m(y)$$

Note that when $x = y$, this would define the Mean First Return Time

Example: Return Times in a game of badminton

Let us divide each side of a badminton court into two quadrants (play zones) labelled 1 and 2 as shown. These quadrants may be regarded as states of a simple Markov model. Data collected over several games of badminton may enable us to populate a simple stochastic matrix as f

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix} \end{array}$$



1. During a rally, given a shot by a player from play zone 1 , after how many shots (on an average) does the shuttle return to quadrant 1 on either side of the net?

Now, let us re-define the Markov chain by constructing new states corresponding to the direction of shots played. For example, the states of the new model are 11 (corresponding to a shot 1 → 1) and so on.

2. Construct the stochastic matrix $\mathbb{P}_{new}$ for this new model.

3. Given a valid service 1 → 1, what is the average number of shots before either player can expect a shot 2 → 1 from their opponent?

Solution: Return Times in a game of badminton

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix} \end{array}$$

1. Using the one-step stochastic model $\mathbb{P}$, we can write the following based on our earlier equation $$\mu_x(y) = 1 + \sum_{m \in S,\, m \neq y} p_{xm}\mu_m(y)$$

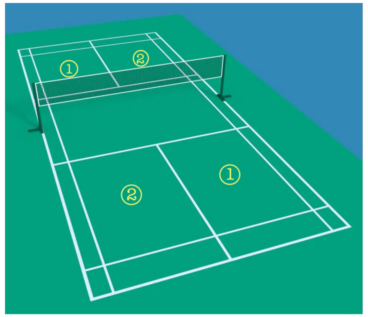
$$\mu_1(1) = 1 + p_{12}\mu_2(1) = 1 + \frac{1}{3}\mu_2(1)$$

$$\mu_2(1) = 1 + p_{22}\mu_2(1) = 1 + \frac{2}{3}\mu_2(1)$$

…. and from first principles…..

$$\mu_1(1) = p_{11}\left(1\right) + p_{12}\left(1 + \mu_2(1)\right) = 1 + p_{12}\mu_2(1)$$

$$\mu_2(1) = p_{21}\left(1\right) + p_{22}\left(1 + \mu_2(1)\right) = 1 + p_{22}\mu_2(1)$$



Solving, we get $\mu_2(1) = 3, \mu_1(1) = 2$

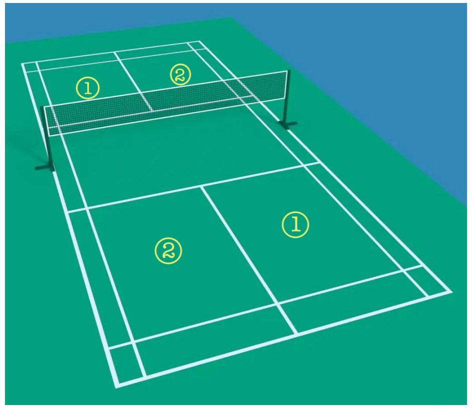
The result $\mu_1(1) = 2$ implies that, on an average, every second shot in a rally returns to 1, given a serve from 1.

Solution: Return Times in a game of badminton

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix} \end{array}$$

2. We now construct $\mathbb{P}_{new}$. The relevant states are 11 , 12 , 21 , and 22 .

For example, for a state transition 11 → 11 , the following sequence of shots must be played: 1 → 1 followed by 1 → 1.



Note that the naming convention of the states is set up in such a way that the last numeral of the current state must be the same as the first numeral of the next state.

This means transitions like 11 → 21 , 21 → 22 , etc. are not possible.

$$\mathbb{P}_{\text{new}} = \begin{array}{c c} & \begin{array}{cccc} 11 & 12 & 21 & 22 \end{array} \\ \begin{array}{c} 11 \\ 12 \\ 21 \\ 22 \end{array} & \begin{bmatrix} 2/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/3 & 2/3 \\ 2/3 & 1/3 & 0 & 0 \\ 0 & 0 & 1/3 & 2/3. \end{bmatrix} \end{array}$$

The transition 11 →11 means that after a shot from 1 to 1, the return shot is chosen to be from 1 to 1 with probability 2/3

$$\Longrightarrow P(11 \to 11) = 2/3.$$

The transition 12 →22 means that after a shot from 1 to 2, the return shot is chosen to be from 2 to 2 with probability 2/3

$$\Longrightarrow P(12 \to 22) = 2/3.$$

Solution: Return Times in a game of badminton

3. Given a valid service 1 → 1, what is the average number of shots before either player can expect a shot 2 → 1 from their opponent?

Using the equation $\mu_x(y) = 1 + \sum_{m \in S,\, m \neq y} p_{xm}\mu_m(y)$ once again with the states 11, 12, 21, 22 as defined earlier, we get the following system of equations

$$\mu_{11}(21) = 1 + \frac{2}{3}\mu_{11}(21) + \frac{1}{3}\mu_{12}(21)$$

$$\mu_{12}(21) = 1 + \frac{2}{3}\mu_{22}(21)$$

$$\mu_{22}(21) = 1 + \frac{2}{3}\mu_{22}(21)$$

These can be solved to get –

$$\mu_{22}(21) = 3$$
$$\mu_{12}(21) = 3$$
$$\mu_{11}(21) = 6$$

Therefore, given a valid serve 1 → 1, either player will have to wait 6 shots on an average before they may expect a shot 2 → 1 from their opponent

Classification of Markov States and Advanced Topics

The behavior of a Markov Chain is characterized by the properties of the stochastic matrix $\mathbb{P}$ and its states.

The states of a Markov Chain can be classified based on the entries of $\mathbb{P}$

Communicating States

A state $j \in \mathscr{S}$ is accessible from a state $i \in \mathscr{S}$, i.e., $i \to j$, if there exists a finite integer $n \geq 0$ such that $p_{ij}^n := P\left(X_n = j \mid X_0 = i\right) > 0$ ..... *i.e., it is possible to go from state* i to *state* j .

If $i \to j$ **and** $j \to i$, then $i \leftrightarrow j$, i.e., states i and j communicate.
When states communicate with each other, they are said to belong to the same class.
Example:

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 2/5 & 1/5 & 0 & 2/5 \\ 1/4 & 1/4 & 1/2 & 0 \end{bmatrix} \end{array}$$

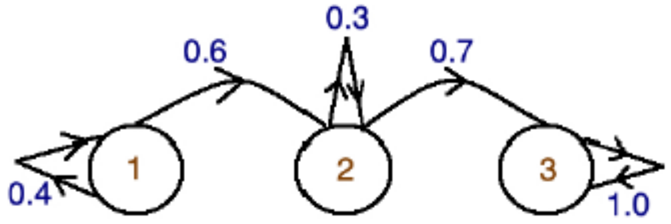
- $3 \leftrightarrow 3$ even though $p_{33} = 0$
- $2 \leftrightarrow 3$ even though $p_{23} = 0$ because 2→4 and 4→3 transitions can happen.
  3 $\to$ 2 can also happen

Irreducible and Reducible Markov Chains:

A Markov Chain is irreducible if all states belong to one class, i.e., if all states communicate with each other as in the example given below

$$\mathbb{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 2/3 & 0 & 0 & 1/3 \\ 1/3 & 0 & 0 & 0 & 2/3 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

An Irreducible Markov Chain is also Ergodic, i.e., *its time averages will be the same as its stochastic averages*



In this Markov model, state 3 is an absorbing state and does not communicate with states 1 and 2. This Markov Chain is not irreducible.

Mean Number of Returns to a State

Let $q_{ij} = p_{ij}^n = P\left(X_n = j \mid X_0 = i\right)$ for some $n \geq 1$ be the probability of return to state j in a finite time starting from state i.

We define the number of visits to state j by the chain $\left\{X_n\right\}_{n \in I, n \geq 0}$ as shown.

- The first visit to state j from state i must happen in $n \leq m$ steps with probability q_{ij}
- This must be followed by $m-1$ revisits to state j starting from state i with probability q_{jj}^{m-1}. This is true because the count for the re-visits to state j happens beginning with state j as the chain is reset as $X_0 = j$ after the first visit to state j.
- Since the summand of interest pertains to m visits to state j (and no more), we must account for the probability $\left(1 - q_{jj}\right)$ of no additional visits to state j after the m^{th}visit

$$E\left(R_j \mid X_0 = i\right) = \sum_{m=0}^{\infty} mP\left(R_j = m \mid X_0 = i\right)$$

$$= \sum_{m=1}^{\infty} m q_{ij} q_{jj}^{m-1}(1 - q_{jj}) = (1 - q_{jj}) q_{ij} \sum_{m=1}^{\infty} m q_{jj}^{m-1}$$

$$= (1 - q_{jj}) q_{ij} \frac{1}{\left(1 - q_{jj}\right)^2} \quad \text{using } \sum_{m=1}^{\infty} m r^{m-1} = \frac{1}{(1-r)^2} \quad \text{for } |r| < 1$$

$$= \frac{q_{ij}}{(1 - q_{jj})}$$

Recurrent States

State $i \in \mathcal{S}$ is recurrent if $q_{ii} = p_{ii}^{n} = 1$. Additionally,

1. State i is recurrent if and only if $E\left(R_i \mid X_0 = i\right) = \infty$
2. State i is recurrent if and only if $P\left(R_i = \infty \mid X_0 = i\right) = 1$

Transient States

State $i \in \mathcal{S}$ is transient when it is not recurrent, i.e., $P\left(R_i = \infty \mid X_0 = i\right) < 1$. Further,

1. $i \in \mathcal{S}$ is transient if and only if $E\left(R_i \mid X_0 = i\right) < \infty$
2. $i \in \mathcal{S}$ is transient if and only if $\sum_{n=1}^{\infty} p_{ii}^{n} < \infty$

Periodicity of a Markov Chain

The period of a state i is the greatest common divisor (denominator) of all integers $n > 0$ for which $p_{ii}^{n} > 0$.

A Markov Chain is aperiodic if it has period ONE.

Example:

$$\mathbb{P} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$

Here, $p_{ii} = 0$, $p_{ii}^{2} = 0$ but $p_{ii}^{3} = 1 > 0$ and so on.
Therefore, this chain has period THREE