ECE 6010
Lecture 10 – Markov Processes

Basic concepts

A Markov process $\{X_t\}$ one such that

$$P(X_{t_{k+1}} = x_{k+1}|X(t_k) = x_k, X(t_{k-1}) = x_{k-1}, \ldots, X(t_1) = x_1) = P(X_{t_{k+1}} = x_{k+1}|X_{t_k} = x_k)$$

(for a discrete random process) or

$$f(x_{t_{k+1}}|X_{t_k} = x_k \ldots, x_{t_1} = x_1) = f(x_{t_{k+1}}|X_{t_k} = x_k)$$

(for a continuous random process). The most recent observation determines the state of the process, and prior observations have no bearing on the outcome if the state is known.

Example 1 Let X_i be i.i.d. and let $S_n = X_1 + \cdots + x_n = S_{n-1} + X_n$. Then

$$P[S_{n+1} = s_{n+1}|S_n = s_n, \ldots, S_1 = s_1] = P[X_{n+1} = S_{n+1} - S_n] = P[S_{n+1} = s_{n+1}|S_n = s_n].$$

$\Box$

Example 2 Let $N(t)$ be a Poisson process.

$$P(N(t_{k+1}) = n_{k+1}|N(t_k) = n_k, \ldots, N(t_1) = n_1)] = P[j-i \text{ events in } t_{k+1}-t_k] = P[N(t_{k+1}) = n_{k+1}|N(t_k) = n_k]$$

$\Box$

Let $\{X_t\}$ be a Markov r.p. The joint probability has the following factorization:

$$P(X_{t_3} = x_3, X_{t_2} = x_2, X_{t_1} = x_1) = P(X_{t_3} = x_3|X_{t_2} = x_2)P(X_{t_2} = x_2|X_{t_1} = x_1)P(X_{t_1} = x_1)$$

(Why?)

Discrete-time Markov Chains

Definition 1 An integer-valued Markov random process is called a Markov chain. $\Box$

Let the time-index set be the set of integers. Let

$$p_j(0) = P[X_0 = j]$$

be the initial probabilities. (Note that $\sum_j p_j(0) = 1$.) We can write the factorization as

$$P[X_n = i_n, \ldots, X_0 = i_0] = P[X_n = i_n|X_{n-1} = i_{n-1}] \cdots P[X_1 = i_1|X_0 = i_0]P[X_0 = i_0].$$

If the probability $P[X_{n+1} = j|X_n = i]$ is does not change with n, then the r.p. X_n is said to have **homogeneous transition probabilities**. We will assume that this is the case, and write

$$p_{ij} = P[X_{n+1} = j|X_n = i].$$

Note: $\sum_j P[X_{n+1} = j|X_n = i] = 1$. That is, $\sum_j p_{ij} = 1$.

We can represent these transition probabilities in matrix form:

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ \vdots & & & \end{bmatrix}$$

The rows of P sum to 1. This is called a **stochastic matrix**.

Frequently discrete-time Markov chains are modeled with state diagrams.

Example 3 Two light bulbs are held in reserve. After a day, the probability that we need a light bulb is p. Let Y_n be the number of new light bulbs at the end of day n.

$$P = \begin{bmatrix} 1 & 0 & 0 \\ p & 1-p & 0 \\ 0 & p & 1-p \end{bmatrix}$$

Draw the diagram. □

Now let us look further ahead. Let

$$p_{ij}(n) = P[X_{n+k} = j | x_k = i]$$

If the r.p. is homogeneous, then $p_{ij}(n) = P[X_k = j | X_i = i]$.

Let us develop a formula for the case that $n = 2$.

$$\begin{aligned} P[X_2 = j, X_1 = l | x_0 = i] &= \frac{P[X_2 = j, X_1 = l, X_0 = i]}{P[X_0 = i]} \\ &= \frac{P[X_2 = j | X_1 = l] P[X_1 = l | X_0 = i] P[X_0 = i]}{P[X_0 = i]} \\ &= p_{il}(1) p_{lj}(1) = p_{il} p_{lj} \end{aligned}$$

Now marginalize:

$$P[X_2 = j | X_0 = i] = \sum_l P[X_2 = j, X_1 = l | X_0 = i] = \sum_l p_{il} P_{lj}.$$

Let $P(2)$ be the matrix of two-step transition probabilities. Then we have

$$P(2) = P(1)P(1) = P^2.$$

In general (by induction) we have

$$P(n) = P^n.$$

Let

$$\mathbf{p}(n) = \begin{bmatrix} P(X_n = 0) \\ P(X_n = 1) \\ \vdots \end{bmatrix}$$

(or whatever the outcomes are). Then

$$p_j(n) = P(X_n = j) = \sum_i P(X_n = j | X_{n-1} = i) P(X_{n-1} = i) = p_{ij} p_i(n-1).$$

Stacking these up, we obtain the equation

$$\mathbf{p}(n) = \mathbf{p}(n-1)P.$$

or

$$\mathbf{p}(n) = \mathbf{p}(0)P^n.$$

We we run the Markov r.p. for a long time what happens to the probabilities? That is, what is $p_j(n)$ as $n \to \infty$? Let us denote

$$\pi_j = \lim_{n \to \infty} p_j(n).$$

If there is a limit, the probability vector $\boldsymbol{\pi}$ should satisfy

$$\boldsymbol{\pi} = \boldsymbol{\pi} P.$$

or

$$P^T \boldsymbol{\pi}^T = \boldsymbol{\pi}^T.$$

This is an eigenvalue problem!

Continuous-Time Markov Processes

Let is still deal with discrete outcomes. If X_t is homogeneous, then

$$P(X(s+t) = j | X(s) = j) = P(X(t) = j | X(0) = i).$$

Let $p_{ij}(t) = P(X(t) = j | X(0) = i)$, and form a matrix $P(t) = [p_{ij}(t)]$, with $P(0) = I$.

Example 4 Suppose $X(t)$ is a Poisson counting process,

$$p_{ij}(t) = P[j - i \text{ events in } t \text{ seconds}] = \frac{(\lambda t)^{j-i} e^{-\lambda t}}{(j-i)!} \qquad j > i.$$

Then

$$P = \left[e^{-\lambda t} \quad \lambda t e^{\lambda t} \quad (\lambda t)^2 e^{-\lambda t}/2 \quad \cdots \quad e^{-\lambda t} \quad \lambda t e^{\lambda t} \quad (\lambda t)^2 e^{-\lambda t}/2 \quad \cdots \vdots \right]$$

□

Let us now consider the question of how long the r.p. remains in a state. Let T_i be the time spent in a state i. The probability of spending more than t seconds in a state is

$$P(T_i > t).$$

Suppose that the process has been in state i already for s seconds. What is the probability that it remains for t more seconds:

$$P(T_i > t + s | T_i > s) = P[T_i > t + s | X(a) = i, 0 \leq a \leq s].$$

But recall that $X(t)$ is Markov:

$$P(T_i > t+s | T_i > s) = P[T_i > t+s | X(a) = i, 0 \leq a \leq s] = P(T_i > t+s | X(s) = i) = P(T_i > t).$$

Such a process is said to be memoryless.

Let us look at these computations again

$$P(T_i > t + s | T_i > s) = \frac{P(T_i > t + s, T_i > s)}{P(T_i > s)} = \frac{P(T_i > t + s)}{P(T_i > s)}.$$

We have seen that this probability must be $P(T_i > t)$:

$$\frac{P(T_i > t + s)}{P(T_i > s)} = P(T_i > t).$$

There is thus a sort of cancellation that takes place. The only distribution which has this property is the exponential,

$$P(T_i > t) = e^{-\lambda_i t}.$$

Using this we have

$$\frac{e^{-\lambda_i (t+s)}}{e^{-\lambda_i s}} = e^{-\lambda_i t}.$$

So the waiting time for a Poisson r.p. is exponential. (We have derived this another way in the homework.)

This result has the following rather curious interpretation: The amount of additional time you have to wait does not depend on the amount of time you have already waited.

We can describe the operation of a continuous-time, Markov chain as follows:

1. Enter a state i.
2. Wait a random amount of time T_i. (this random variable is continuous)

3. Select a new state according to a discrete-time Markov chain with transition probabilities we will call $\tilde{q}_{ij}$
4. Repeat.

In discrete time we have the probability update $\mathbf{p}(k+1) = \mathbf{p}(k)P$. We will develop an analogous result for continuous time. Instead of a set of couple difference equations, we will get a set of coupled differential equations.

Let δ be a small time increment.

$$P(T_i > \delta) = e^{-\nu_i \delta} \approx 1 - \nu_i \delta + o(\delta).$$

The probability that we remain in the same state at time δ is

$$p_{ii}(\delta) = P(T_i > \delta) = 1 - \nu_i \delta + o(\delta)$$

or $1 - p_{ii}(\delta) = \nu_i \delta + o(\delta)$.

Now consider the transition. When leaving state i, we move to state j with probability $\tilde{q}_{ij}$:

$$p_{ij}(\delta) = \underbrace{(1 - p_{ii}(\delta))}_{\text{leave state}} \tilde{q}_{ij} = \nu_i \delta \tilde{q}_{ij} + o(\delta).$$

Let $\gamma_{ij} = \nu_i \tilde{q}_{ij}$:

$$p_{ij}(\delta) = \gamma_{ij}\delta + o(\delta).$$

We say that γ_{ij} is the **rate** at which $X(t)$ enters state j from state i. Define $\gamma_{ii} = -\nu_i$, so that

$$1 - p_{ii}(\delta) = -\gamma_{ii}\delta + o(\delta)$$

or

$$p_{ii}(\delta) - 1 = \gamma_{ii}\delta + o(\delta).$$

Summarizing what we have so far:

$$\begin{aligned} p_{ii}(\delta) - 1 &= \gamma_{ii}\delta + o(\delta) \\ p_{ij}(\delta) &= \gamma_{ij}\delta + o(\delta). \end{aligned}$$

Divide by δ and take the limit:

$$\lim_{\delta \to 0} \frac{p_{ii}(\delta) - 1}{\delta} = \gamma_{ii}$$

$$\lim_{\delta \to 0} \frac{p_{ij}(\delta)}{\delta} = \gamma_{ij}$$

Now define $p_j(t) = P(X(t) = j)$. Then we have

$$p_j(t+\delta) = P(X(t+\delta) = j) = \sum_i P(X(t+\delta) = j | X(t) = i) P(X(t) = i) = \sum_i p_{ij}(\delta) p_i(t).$$

and

$$\begin{aligned} p_j(t+\delta) - p_j(t) &= \sum_i p_{ij}(\delta)p_i(t) - p_j(t) \qquad = \sum_{i \neq j} p_{ij}(\delta)p_i(t) + p_{jj}(t)p_j(t) - p_j(t) \\ &= \sum_{i \neq j} p_{ij}(\delta)p_i(t) + (p_{jj}(t) - 1)p_j(t). \end{aligned}$$

Divide both sides by δ and take the limit:

$$p_j'(t) = \sum_{i \neq j} \gamma_{ij} p_i(t) + \gamma_{jj} p_j(t) = \sum_i \gamma_{ij} p_i(t)$$

Example 5 Let us model a two-state system, having an idle state and a busy state. In the idle state, the system is waiting for work to arrive. Assume that the waiting time is an exponential r.v. with mean $1/\alpha$. In the busy state, the machine works for a random amount of time, with an exponential distribution having mean $1/\beta$.

We can think of the rate of motion from idle to busy as α, and the rate of motion from busy to idle as β.

The underlying discrete-time Markov chain has

$$\tilde{Q} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(when the transition occurs, there is no ambiguity). We find that

$$\gamma_{00} = -\alpha \qquad \gamma_{01} = \alpha q_{01} = \alpha \qquad \gamma_{10} = \beta q_{10} = \beta \qquad \gamma_{11} = -\beta.$$

We obtain the coupled differential equations

$$\begin{bmatrix} p_0(t) \\ p_1(t) \end{bmatrix}' = \begin{bmatrix} -\alpha & \beta \\ \alpha & -\beta \end{bmatrix} \begin{bmatrix} p_0(t) \\ p_1(t) \end{bmatrix}$$

or

$$\mathbf{p}'(t) = A\mathbf{p}(t).$$

We also have an auxiliary equation $p_0(t) + p_1(t) = 1$, or $\mathbf{p}(t)^T\mathbf{1} = 1$.

We can solve this as

$$\mathbf{p}(t) = \exp[At]\mathbf{p}(0).$$

Somewhat more explicitly:

$$s\mathbf{p}(s) - \mathbf{p}(0) = A\mathbf{p}(s)$$

$$(sI - A)\mathbf{p}(s) = \mathbf{p}(0)$$

$$\mathbf{p}(s) = (sI - A)^{-1}\mathbf{p}(0).$$

$$(sI - A)^{-1} = \begin{bmatrix} s+\alpha & -\beta \\ -\alpha & s+\beta \end{bmatrix}^{-1} = \frac{1}{(s+\alpha)(s+\beta)} \begin{bmatrix} s+\beta & \beta \\ \alpha & s+\alpha \end{bmatrix}.$$

so

$$\mathbf{p}(s)\frac{1}{(s+\alpha)(s+\beta)} \begin{bmatrix} (s+\beta)p_1(0) + \beta p_2(0) \\ \alpha p_1(0) + (s+\alpha)p_2(0) \end{bmatrix}.$$

Now it is a matter of straightforward (but careful) computation to show that

$$p_0(t) = \frac{\beta}{\alpha+\beta} + (p_0(0) - \frac{\beta}{\alpha+\beta})e^{-(\alpha+\beta)t}$$

$$p_1(t) = \frac{\alpha}{\alpha+\beta} + (p_1(0) - \frac{\alpha}{\alpha+\beta})e^{-(\alpha+\beta)t}.$$

In the limit

$$p_0(t) \to \frac{\beta}{\alpha+\beta} \qquad p_1(t) \to \frac{\alpha}{\alpha+\beta}.$$

□

What are the steady-state conditions in general?

$$0\ sum_i \gamma_{ij} p_i$$

Since $\gamma_{jj} = -\nu_j$ we can write

$$\nu_j p_j = \sum_{i \neq j} \gamma_{ij} p_i$$

and since

$$\nu_j = \sum_{i \neq j} \gamma_{ji}$$

we can write

$$p_j(\sum_{i \neq j} \gamma_{ji}) = \sum_{i \neq j} \gamma_{ij} p_i.$$

Classes of States

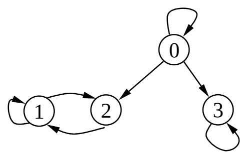
Definition 2 State j is **accessible** from state i if $p_{ij}(n) \geq 0$ for some n. More informally, a path from i to j exists in the state diagram.

States i and j **communicate** if they are accessible to each other.

Two states are said to be in the same **class** if they communicate with each other.

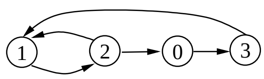
A Markov chain with a single class is **irreducible**. □

Example 6 3 classes.



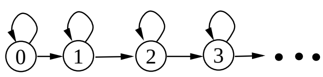
□

Example 7 One class: irreducible.



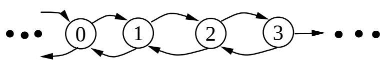
□

Example 8 Classes: $\{0\}$, $\{1\}$, $\{2\}$, $\ldots$



□

Example 9 One class = $\{\mathbb{Z}\}$.



□

Definition 3 A class is **recurrent** if the process returns to the state with probability 1. □

Let $f_i = P[\text{ever returning to state } i]$. Then state i is recurrent if $f_i = 1$.

If $f_i < 1$, then state i is said to be **transient**.

- If started in a transient state, then the state does not recur an infinite number of times.

- If in a recurrent state, then the state recurs an infinite number of times.

Let X_n denote the Markov chain with $X_0 = i$. Let

$$I_i(x) = \begin{cases} 1 \text{if } X = i \\ 0 \text{otherwise} \end{cases}.$$

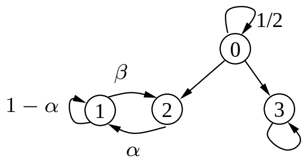
Then

$$E[\text{number of returns to state } i] = E\left[\sum_i I_i(X_n) | X_0 = i\right] = \sum_{n=1}^{\infty} p_{ii}(n).$$

We see that recurrent means that $\sum_{n=1}^{\infty} p_{ii}(n) = \infty$.

Transient means that $\sum_{n=1}^{\infty} p_{ii}(n) < \infty$.

Example 10



$$p_{00}(n) = (1/2)^n$$
$$\sum_{n=1}^{\infty} p_{00}(n) = 1 < \infty$$

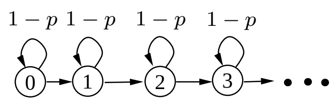
so state 0 is transient.

What if we start in state 1?

$$p_{11}(n) = \frac{\beta + \alpha(1 - \alpha - \beta)^n}{\alpha + \beta}$$

This sums to ∞. □

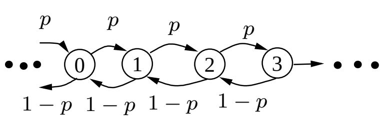
Example 11



$$p_{ii}(n) = (1 - p)^n.$$
$$\sum_n p_{ii}(n) < \infty$$

□

Example 12 Random walk.



Start in state 0. We can return if we make as many right-hand moves (with probability p) as left-hand moves (with probability $1-p$). The total number of moves (left and right) must be an even number; take the total number as $2n$.

There are $\binom{2n}{n}$ ways of making n RH moves:

$$p_{00}(n) = \binom{2n}{n} p^n (1-p)^n$$

Summing (to see if transient):

$$\sum_{n=1}^{\infty} p_{00}(n) = \sum_{n=1}^{\infty} \binom{2n}{n} p^n (1-p)^n$$

How to sum this? We can get a good approximation using Stirling's formula:

$$\boxed{n! \approx \sqrt{2\pi} n^n \sqrt{n} e^{-n}}$$

Then

$$\binom{2n}{n} \approx \frac{1}{\sqrt{\pi n}} 4^n$$

and

$$p_{00}(n) \approx \frac{[4p(1-p)]^n}{\sqrt{\pi n}}$$

Now sum:

$$\sum_{n=1}^{\infty} p_{00}(n) \approx \sum_{n=1}^{\infty} \frac{[4p(1-p)]^n}{\sqrt{\pi \nu}}$$

Still a little hard. But take the particular case of $p = 1/2$. Then we get

$$\sum_{n=1}^{\infty} \frac{1}{\sqrt{\pi n}} = \infty \text{ (why?)}$$

If $p \neq 1/2$, then $[4p(1-p)] < 1$, and the sum converges.

□

Observation: The states of an irreducible, finite-state Markov chain are all recurrent.

Limiting probabilities

If all states are transient, then all the state probabilities approach 0 as $n \to \infty$. If a M.C. has some transient classes and some recurrent classes, then eventually the process enters and remains in one of the recurrent classes. For limiting purposes, we can focus on individual recurrent classes.

Suppose a M.C. starts in a recurrent state i at time 0. Let $T_i(1), T_i(1) + T_i(2), \ldots$ denote the times when the process returns to state i, where $T_i(k)$ is the time that elapses between the $(k-1)$th and kth returns. The T_i form an i.i.d. sequence. The proportion of time spent in state i after k returns is

$$\frac{k}{T_i(1) + T_i(2) + \cdots + T_i(k)}$$

In the limit,

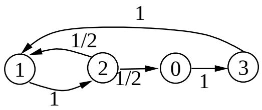
$$\text{proportion of time spent in state } i \to \frac{1}{E[T_i]} \triangleq \pi_i$$

(by the law of large numbers), where $E[T_i]$ is the mean recurrence time.

If $E[T_i] < \infty$, we say that state i is **positive recurrent**: $\pi_i > 0$.

If $E[T_i] = \infty$, we say that state i is **null recurrent**: $\pi_i = 0$.

Example 13



The M.C. returns to state 0 in two steps with probability 1/2, and in four steps with probability 1/2. The mean recurrence time is

$$E[T_0] = \frac{1}{2}2 + \frac{1}{2}4 = 3.$$

State 0 is positive recurrent: $\pi_0 = 1/3$. □

Example 14 In the random walk with $p = 1/2$, the process is recurrent. However, it can be shown that the mean recurrence time is ∞ in this case. This means that the process is null recurrent. □

We can find the π_js using $\boldsymbol{\pi} = \boldsymbol{\pi} P$ (and solve this if the number of states is finite).

Summary: The proportion of time spent in state j is π_j.