6.262: Discrete Stochastic Processes 2/21/11

Lecture 6: From Poisson to Markov

Outline:

- Joint conditional densities for Poisson
- Definition of finite-state Markov chains
- Classification of states
- Periodic states and classes
- Ergodic Markov chains

Recall (Section 2.2.2) that the joint density of interarrivals $X_1, \ldots, X_n$ and arrival epoch S_{n+1} is

$$\mathsf{f}_{X_1 \cdots X_n S_{n+1}}(x_1, \ldots x_n, s_{n+1}) = \lambda^{n+1} \exp(-\lambda s_{n+1})$$

Conditional on $S_{n+1} = t$ (which is Erlang),

$$\mathsf{f}_{X_1 \cdots X_n | S_{n+1}}(x_1, \ldots x_n | t) = \frac{\lambda^{n+1} e^{-\lambda t}}{\left[\frac{\lambda^{n+1} t^n e^{-\lambda t}}{n!}\right]} = \frac{n!}{t^n} \qquad (1)$$

Similarly (from Eqn. 2.43, text)

$$\mathsf{f}_{X_1 \cdots X_n | N(t)}(x_1, \ldots x_n | n) = \frac{\lambda^{n+1} e^{-\lambda t}}{\left[\frac{\lambda^{n+1} t^n e^{-\lambda t}}{n!}\right]} = \frac{n!}{t^n} \qquad (2)$$

Both equations are for $0 < x_1, \ldots, x_n$ and $\sum x_k < t$.

Both say the conditional density is uniform over the constraint region.

Why are the two equations the same? If we condition $X_1, \ldots, X_n$ on both $N(t) = n$ and $S_{n+1} = t_1$ for any $t_1 > t$, (2) is unchanged.

By going to the limit $t_1 \to t$, we get the first equation. The result, $n!/t^n$, is thus the density conditional on n arrivals in the open interval $(0, t)$ and is unaffected by future arrivals.

This density, and its constraint region, is symmetric in the arguments $x_1, \ldots, x_n$. More formally, the constraint region (and trivially the density in the constraint region) is unchanged by any permutation of $x_1, \ldots, x_n$.

Thus the marginal distribution, $\mathsf{F}_{X_k|N(t)}(x_k|n)$ is the same for $1 \le k \le n$. From analyzing $S_1 = X_1$, we then know that $\mathsf{F}^{\mathsf{c}}_{X_k|N(t)}(x_k|n) = (t - x_n)^n/t^n$ for $1 \le k \le n$.

For the constraint $S_{n+1} = t$, we have analyzed $X_1, \ldots, X_n$, but have not considered X_{n+1}, the final interarrival interval before t.

The reason is that $\sum_{k=1}^{n+1} X_k = S_{n+1} = t$, so that these variables do not have an $n{+}1$ dimensional density.

The same uniform density as before applies to each subset of n of the $n{+}1$ variables, and the constraint is symmetric over all $n + 1$ variables.

This also applies to the constraint $N(t) = n$, using $X^*_{n+1} = t - S_n$

Definition of finite-state Markov chains

Markov chains are examples of integer-time stochastic processes, $\{X_n; n \geq 0\}$ where each X_n is a rv.

A finite-state Markov chain is a Markov chain in which the sample space for each rv X_n is a fixed finite set, usually taken to be $\{1, 2, \ldots, \mathsf{M}\}$.

Any discrete integer-time process is characterized by $\Pr\{X_n = j \mid X_{n-1}{=}i, X_{n-2}{=}k, \ldots, X_0{=}m\}$ for $n \geq 0$ and all $i, j, k, \ldots, m$, each in the sample space.

For a finite-state Markov chain, these probabilities are restricted to be

$$\Pr\{X_n = j \mid X_{n-1}{=}i, X_{n-2}{=}k \ldots X_0{=}m\} = P_{ij}$$

where P_{ij} depends only on i, j and $\mathsf{p}_{X_0}(m)$ is arbitrary.

The definition first says that X_n depends on the past only through X_{n-1}, and second says that the probabilities don't depend on n for $n \geq 1$.

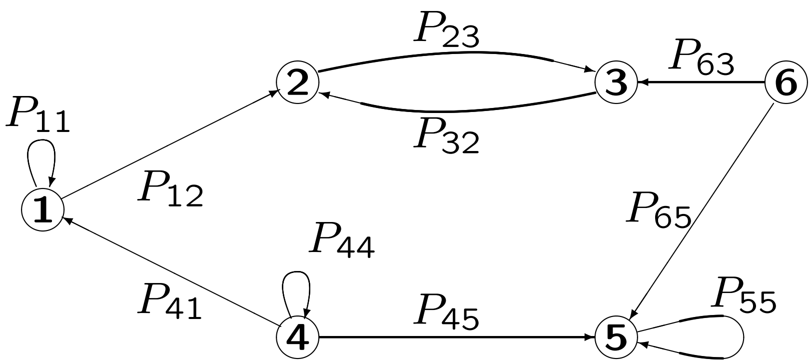
Some people call this a homogeneous Markov chain and allow P_{ij} to vary with n in general.

The rv's $\{X_n;\, n \geq 0\}$ are dependent, but in only a very simple way. X_n is called the state at time n and characterizes everything from the past that is relevant for the future.

A Markov chain is completely described by $\{P_{ij};\, 1 \leq i, j \leq \mathsf{M}\}$ plus the initial probabilities $\mathsf{p}_{X_0}(i)$.

We often take the initial state to be a fixed value, and often view the Markov chain as just the set $\{P_{ij};\ 1 \leq i, j \leq \mathsf{M}\}$, with the initial state viewed as a parameter.

Sometimes we visualize $\{P_{ij}\}$ in terms of a directed graph and sometimes as a matrix.



a) Graphical

$$[P] = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{16} \\ P_{21} & P_{22} & \cdots & P_{26} \\ \vdots & \vdots & \vdots\;\vdots\;\vdots & \vdots \\ P_{61} & P_{62} & \cdots & P_{66} \end{bmatrix}$$

b) Matrix

The graph emphasizes the possible and impossible (an edge from i to j explicitly means that $P_{ij} > 0$).

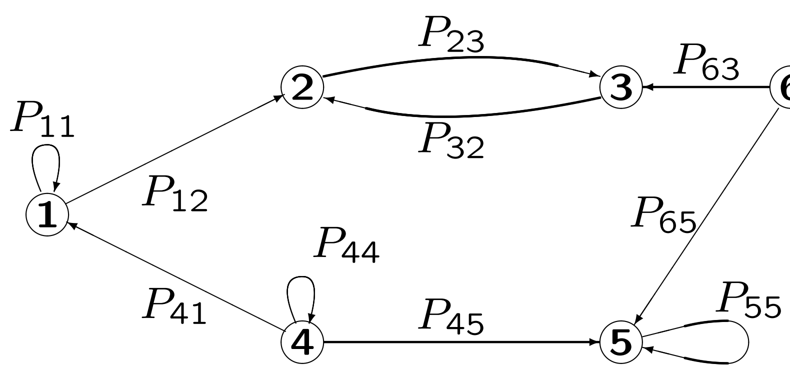
The matrix is useful for algebraic and asymptotic issues.

Classification of states

Def: An (n-step) walk is an ordered string of nodes (states), say $(i_0, i_1, \ldots i_n)$, $n \geq 1$, with a directed arc from i_{m-1} to i_m for each m, $1 \leq m \leq n$.

Def: A path is a walk with no repeated nodes.

Def: A cycle is a walk in which the last node is the same as the first and no other node is repeated.



Walk: (4, 4, 1, 2, 3, 2)
Walk: (4, 1, 2, 3)
Path: (4, 1, 2, 3)
Path: (6, 3, 2)
Cycle: (2, 3, 2)
Cycle: (5, 5)

It doesn't make any difference whether you regard $(2, 3, 2)$ and $(3, 2, 3)$ as the same or different cycles.

Def: A state (node) j is accessible from i ($i \to j$) if a walk exists from i to j.

Let $P_{ij}^n = \Pr\{X_n = j \mid X_0 = i\}$. Then if i, k, j is a walk, $P_{ik} > 0$ and $P_{kj} > 0$, so $P_{ij}^2 \geq P_{ik}P_{kj} > 0$.

Similarly, if there is an n-step walk starting at i and ending at j, then $P_{ij}^n > 0$.

Thus if $i \to j$, there is some n for which $P_{ij}^n > 0$. To the contrary, if j is not accessible from i ($i \not\to j$), then $P_{ij}^n = 0$ for all $n \geq 1$.

$i \to j$ means that, starting in i, entry to j is possible, perhaps with multiple steps. $i \not\to j$ means there is no possibility of ever reaching j from i.

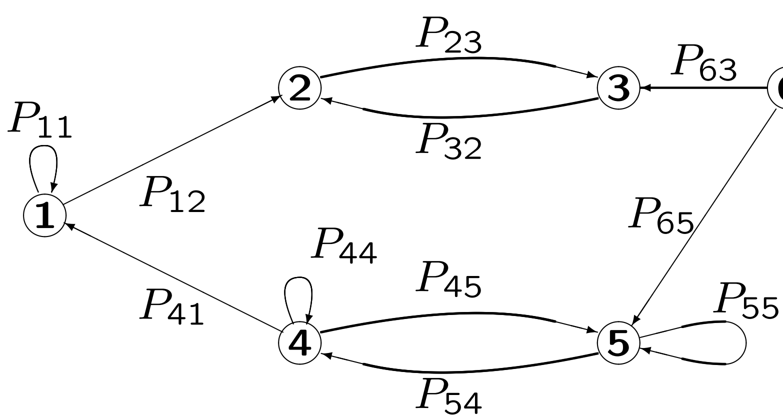
If $i \to j$ and $j \to k$, then $i \to k$. (Concatenate a walk from i to j with a walk from j to k.)

Def: States i and j communicate ($i \leftrightarrow j$) if $i \to j$ and $j \to i$.

Note that if ($i \leftrightarrow j$) and ($j \leftrightarrow k$), then ($i \leftrightarrow k$).

Note that if ($i \leftrightarrow j$), then there is a cycle that contains both i and j.

Def: A class $\mathcal{C}$ of states is a non-empty set of states such that each $i \in \mathcal{C}$ communicates with every other $j \in \mathcal{C}$ and communicates with no $j \notin \mathcal{C}$.



$\mathcal{C}_1 = \{2, 3\}$

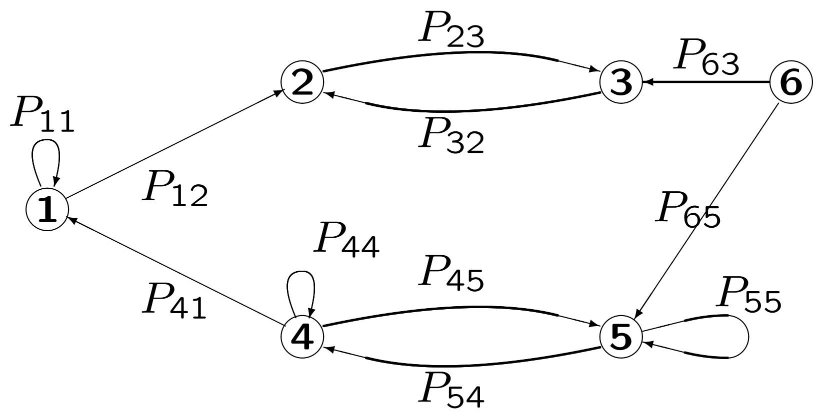
$\mathcal{C}_2 = \{4, 5\}$

$\mathcal{C}_3 = \{1\}$

$\mathcal{C}_4 = \{6\}$

Why is $\{6\}$ a class

Def: A state i is recurrent if $j \to i$ for all j such that $i \to j$. (i.e., if no state from which there is no return can be entered.) If a state is not recurrent, it is transient.



2 and 3 are recurrent

4 and 5 are transient

$4 \to 1, 5 \to 1, 1 \not\to 4, 5$

6 and 1 also transient

Thm: The states in a class are all recurrent or all transient.

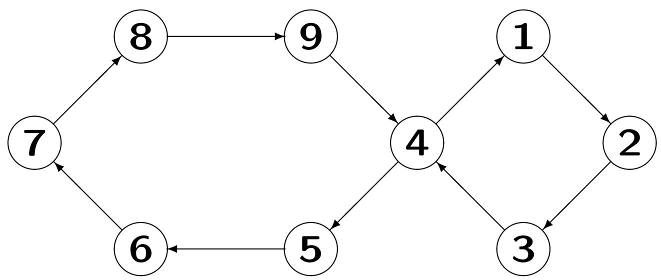
Pf: Assume i recurrent and let $\mathcal{S}_i = \{j : i \to j\}$. By recurrence, $j \to i$ for all $j \in \mathcal{S}_i$. Thus $i \leftrightarrow j$ if and only if $j \in \mathcal{S}_i$, so $\mathcal{S}_i$ is a class. Finally, if $j \in \mathcal{S}_i$, then $j \to k$ implies $i \to k$ and $k \to i \to j$, so j is recurrent.

Periodic states and classes

Def: The period, $d(i)$, of state i is defined as

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

If $d(i) = 1$, i is aperiodic. If $d(i) > 1$, i is periodic with period $d(i)$.



For example, $P_{44}^n > 0$ for

n = 4, 6, 8, 10; $d(4) = 2$

For state 7, $P_{77}^n > 0$ for

n = 6, 10, 12, 14; $d(7) = 2$

Thm: All states in the same class have the same period.

See text for proof. It is not very instructive.

A periodic class of states with period $d > 1$ can be partitioned into subclasses $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_d$ so that for $1 \leq \ell < d$, and all $i \in \mathcal{S}_\ell$, $P_{ij} > 0$ only for $j \in \mathcal{S}_{\ell+1}$. For $i \in \mathcal{S}_d$, $P_{ij} > 0$ only for $j \in \mathcal{S}_1$. (see text)

In other words, starting in a given subclass, the state cycles through the d subclasses.

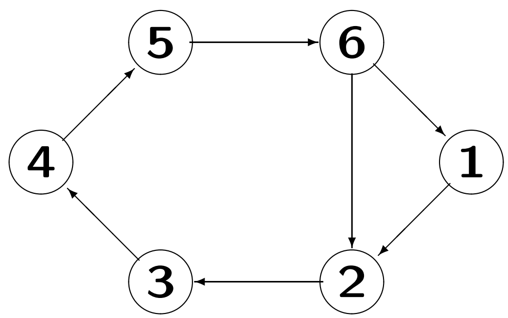
Ergodic Markov chains

The most fundamental and interesting classes of states are those that are recurrent and aperiodic. These are called ergodic. A Markov chain with a single class that is ergodic is an ergodic Markov chain.

Ergodic Markov chains gradually lose their memory of where they started, i.e., P_{ij}^n goes to a limit $\pi_j > 0$ as $n \to \infty$, and this limit does not depend on the starting state i.

This result is also basic to arbitrary finite-state Markov chains, so we look at it carefully and prove it next lecture.

A first step in showing that $P_{ij}^n \to \pi_j$ is the much weaker statement that $P_{ij}^n > 0$ for all large enough n. This is more a combinatorial issue than probabalistic, as indicated below.



Starting in state 2, the state at the next 4 steps is deterministic. For the next 4 steps, there are two possible choices then 3, etc.

This hints at the following theorem:

Thm: For an ergodic M state Markov chain, $P_{ij}^n > 0$ for all i, j, and all $n \geq (\mathsf{M} - 1)^2 + 1$.

6.262 Discrete Stochastic Processes
Spring 2011