# DISCRETE STOCHASTIC PROCESSES

## Lecture 19

**Review:** Markov Processes

Kolmogorov Differential Equations

Backward Kolmogorov Equation

Forward Kolmogorov Equation

**Review:** Reversibility in Markov Processes

Burke's Theorem for M/M/k Queues

# MARKOV PROCESSES

A Markov process is a semi-Markov process with exponentially distributed inter-transition intervals (rate $\nu_i$) that are independent of the next state. There can be a different rate of departure for each state the chain may be in.

Now the future is conditionally independent of the past, given the present state $X(t)$. Therefore the process has the **Markov Property**

For $s > t$, $P\left[X(s) = j \mid X(t) = i, \{X(\tau); \tau < t\}\right] = P\left[X(s) = j \mid X(t) = i\right]$

A Markov process is defined to be ***irreducible*** if the embedded chain is irreducible; assume this throughout. For now, we also assume $P_{ii} = 0$, i.e., there are no self transitions allowed in the embedded chain.

We have defined a Markov process in terms of an exponential departure rate for each state ($\nu_i$) and embedded Markov chain probabilities ($P_{ij}$).

## Alternate Description of a Markov Process

Given I'm in state *i*, assume that there are a bunch of exponential processes racing. If process $i \to j$ fires first I go to state *j*. (For example, this corresponds to an M/M/1 queue with a positive number of customers having exponential interarrivals increasing the state by 1 and independent different rate exponential interdepartures decreasing the state by 1.)

Assume that, for each *j*, $q_{ij}$ is the rate of the exponential intertransition time racing to move the process to state *j*. Then the intertransition time of leaving state state *i* is exponential with rate $\nu_i = \sum_{j \neq i} q_{ij}$ and the probability of going from *i* to *j* is

$$P_{ij} = \frac{q_{ij}}{\sum_{j \neq i} q_{ij}} = \frac{q_{ij}}{\nu_i}.$$

This recovers our original description of a Markov process.
Starting from the original description we can get the racing rates by $q_{ij} = P_{ij} \nu_i$, i ≠ j,
and of course

$$q_{ii} = -\sum_{j \neq i} q_{ij}$$

gives the expected rate of leaving state i.

## Steady State Probabilities

The embedded chain is positive recurrent iff the **steady state equations**,

$$\pi_j = \sum_i \pi_i P_{ij}, \pi_j \geq 0 \text{ for all } j \text{ and} \sum_i \pi_i = 1$$

have a solution. Then $\pi_j > 0$ for all *j*, the solution is unique, and $\pi_j$ is the fraction of the total **transitions** that go into state *j*.

If the steady state equations have a solution and if $\sum_i \pi_i / \nu_i < \infty$, (so we don't get $p_i = 0$ for all $i$) then

$$p_i = \frac{\pi_i / \nu_i}{\sum_k \pi_k / \nu_k}$$

gives the fraction ***of* time** spent in state *i*. $\sum_i p_i = 1$ and $p_i$ can be interpreted as the probability of being in state *i* at time *t*.

Multiplying this equation by $\nu_i$ and summing over *i* gives $\sum_i p_i \nu_i = \dfrac{1}{\sum_k \pi_k / \nu_k}$ so

$$\pi_i = \frac{p_i \nu_i}{\sum_k p_k \nu_k}$$

Subbing this into $\vec{\pi} = \vec{\pi}P$ we get the steady-state equations in the following theorem. This equation says that the rate of transitions into a state equals the rate of transitions out of a state in steady-state.

## Steady State Equations

**Theorem 1:** Recall that $q_{ij} = \upsilon_i P_{ij}$ and we have assumed $q_{ii} = P_{ii} = 0$. If $\{p_i\}$ satisfies

$$p_j \nu_j = \sum_i p_i \nu_i P_{ij} = \sum_i p_i q_{ij}; \quad p_j \geq 0, \sum_j p_j = 1$$

and $\sum_i p_i \nu_i < \infty$, then the solution is unique, $p_i > 0$, $p_i = \lim_{t\to\infty} P\big(X(t) = i\big)$, the embedded chain is positive recurrent and $\pi_i = \dfrac{p_i \nu_i}{\sum_k p_k \nu_k}$

Thus we have two guessing theorems: Guess the $\vec{\pi}$ (along with $\sum_i \pi_i / \nu_i < \infty$) or guess the $\vec{p}$ (along with $\sum_i p_i \nu_i < \infty$) and either way, we get a unique solution.

## Two Strange Cases

You can get the $\vec{\pi}$ but $\sum_i \pi_i / \nu_i = \infty$. This means that the rates of leaving high ordered states get small, so that the expected time to return to any state is infinite (although return is certain).

You can get the $\vec{p}$ but $\sum_i p_i \nu_i = \infty$. This means that the rates of leaving high ordered states get large. The embedded chain is either null recurrent or transient. The transient case is called "irregular" and means that the process can transit through an infinite number of states in a finite time.
In an irregular process the $\vec{p}$ is meaningless and has no steady state connotation.

**Definition:** We will usually limit our attention to irreducible embedded chain where, in addition, you can get the $\vec{p}$ and $\sum_i p_i \nu_i < \infty$. Such processes are called **ergodic**. (This implies the $\pi$'s exist, and therefore the embedded chain is positive recurrent. Thus an ergodic Markov process has an irreducible, positive recurrent embedded chain, ($\pi$'s exist, are unique and are strictly positive), and the p's exist, are unique and are strictly positive (because $\sum_k \pi_k / \upsilon_k < \infty$.)

# Comparison Between Steady-State Equations for a Markov Chain and a Markov Process

## Markov Chain

$$\vec{\pi} = \vec{\pi} P$$

$$\pi_j = \sum_{all\ i} \pi_i P_{ij} = \sum_{i \neq j} \pi_i P_{ij} + \pi_j P_{jj}$$

$$\pi_j (1 - P_{jj}) = \sum_{i \neq j} \pi_i P_{ij}$$

$$\pi_j (\sum_{k \neq j} P_{jk}) = \sum_{i \neq j} \pi_i P_{ij}$$

Expected # transitions out of j per step (≤1) = expected # transitions into j per step.

## Markov Process

$$q_{ij} = \upsilon_i P_{ij}$$

$$q_i = P_{ii} = 0$$

$$p_j \upsilon_j = \sum_{i \neq j} p_i q_{ij}$$

$$p_j \upsilon_j \delta = \sum_{i \neq j} p_i q_{ij} \delta$$

Expected # transitions out of j in time step $\delta$ = expected # of transitions into j in time step $\delta$ .

## Kolmogorov Differential Equations for $P_{ij}(t)$

Assume Markov process is ergodic and define $P_{ij}(t)$ as:

$$P_{ij}(t) = P\left(X(t)=j \mid X(0)=i\right) = \sum_k P\left\{X(s)=k \mid X(0)=i\right\} P\left\{X(t)=j \mid X(s)=k\right\}$$
$$= \sum_k P_{ik}(s) P_{kj}(t-s)$$

$$\boxed{P_{ij}(t) = \sum_k P_{ik}(s) P_{kj}(t-s)}$$ (Chapman-Kolmogorov)

Letting *s* approach 0 or *t*, we get the **Kolmogorov differential equations.** The **Backward Equation** results as $s \to 0$. Assume $q_{ii} = 0$.

$$\boxed{P_{ij}(t) = \sum_k P_{ik}(s) P_{kj}(t-s)}$$

$P_{ik}(s) = q_{ik} s + o(s)$ for $k \neq i$;

$$P_{ii}(s) = 1 - \sum_{k \neq i} q_{ik} s + o(s) = 1 - v_i s + o(s)$$

$$P_{ij}(t) = \sum_{k \neq i} q_{ik} s \, P_{kj}(t-s) + (1 - v_i s) P_{ij}(t-s) + o(s)$$

$$\frac{P_{ij}(t) - P_{ij}(t-s)}{s} = \sum_{k \neq i} q_{ik} \, P_{kj}(t-s) - v_i P_{ij}(t-s) + \frac{o(s)}{s}$$

Going to the limit $s \to 0$

$$\frac{dP_{ij}(t)}{dt} = \sum_{k \neq i} q_{ik} \, P_{kj}(t) - v_i \, P_{ij}(t)$$

$$\left( = \sum_{k \neq i} q_{ik} \left( P_{kj}(t) - P_{ij}(t) \right) \right)$$

$$\frac{dP_{ij}(t)}{dt} = \sum_{k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t)$$
$$= \sum_{k \neq i} q_{ik} \left( P_{kj}(t) - P_{ij}(t) \right)$$

For initial state i, the rate of change in probability of occupancy of state j is the expected rate of transitions into j per unit time from states into which transitions may have occurred at the beginning ( $q_{ik} P_{kj}$ ) minus the expected rate of transitions out of j ($-\upsilon_i P_{ij}$ ).

The backward equation fixes j and has variable i (!)
For any final state *j*, this gives us a family of first order linear differential equations (same as linear electrical circuit or control theory).

Let $[P(t)]$ be a matrix with elements $P_{ij}(t)$. $[P(0)] = I$
Let [$Q$] be a matrix with terms $q_{ij}$ for $i \neq j$ and $q_{ii} = -\nu_i$. Note that [$Q$] is not a stochastic matrix; its rows sum to 0. It is a matrix of rates.

$$\frac{d[P(t)]}{dt} = [Q][P(t)]; \; [P(t)] = \sum_{m=0}^{\infty} \frac{t^m [Q]^m}{m!}$$

The **Forward Kolmogoroff Differential Equation** comes from letting $s \to t$ from below.

$$P_{ij}(t) = \sum_k P_{ik}(s) P_{kj}(t-s)$$

Then

$$P_{ij}(t) = \sum_{k \neq j} P_{ik}(s)(t-s)q_{kj} + P_{ij}(s)\left[1-(t-s)\nu_j\right] + o(t-s)$$

$$\frac{P_{ij}(t) - P_{ij}(s)}{t-s} = \sum_{k \neq j}\left[P_{ik}(s)q_{kj}\right] - P_{ij}(s)\nu_j$$

$$\frac{dP_{ij}(t)}{dt} = \sum_{k \neq j} P_{ik}(t)q_{kj} - P_{ij}(t)\nu_j$$

$$= \sum_{k \neq j} P_{ik}(t)q_{kj} - q_{jk}P_{ij}(t)$$

The forward equation uses only probabilities conditioned on starting in state i at time zero. The rate of change of the probability state j is occupied is the sum of the expected rate of transitions per unit time into j minus the expected rate of transitions per unit time out of j.

$$\frac{d[P(t)]}{dt} = [P(t)][Q]; \quad [P(t)] = \sum_{m=0}^{\infty} \frac{t^m [Q]^m}{m!}$$

These equations are helpful in seeing how steady state is approached. They can be analyzed with any of the techniques used for sets of linear first order differential equations.

Note that $[Q]$, with off diagonal elements $q_{ij}$ and diagonal elements $-\nu_i$, is related to the discrete time approximation matrix $[W]$ with off diagonal elements $q_{ij}d$ and diagonal elements $1-\nu_i d$.

$$[W] = d[Q] + [I]$$

## Discrete Time Markov Chain Approximation

Assume that $\nu_i \le b$ for all *i*, some bound *b*. Then we can look at the sampled time Markov chain approximation over a short time interval *d* with transition probabilities $W_{ij} = q_{ij}d = \nu_i P_{ij} d$, $W_{ii} = 1 - \nu_i d$. Then the steady state probabilities $\{w_j\}$ in this chain satisfy

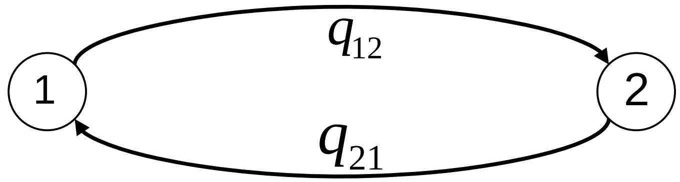
$$w_j = \sum_{i \neq j} w_i W_{ij} + w_j W_{jj} = \sum_{i \neq j} w_i \nu_i P_{ij} d + w_j \left(1 - \nu_j d\right)$$

$$\nu_j w_j = \sum_{i \neq j} w_i q_{ij}$$

Thus $w_i = p_i$; We get an exact solution for the steady state probabilities $p_i$ of being in state *i*, not the steady state embedded probabilities $\pi_i$.

Note that [*Q*], with off diagonal elements $q_{ij}$ and diagonal elements $-\nu_i$, is related to the discrete time approximation matrix [*W*] with off diagonal elements $q_{ij}d$ and diagonal elements $1 - \nu_i d$.

$$[W] = d[Q] + [I]$$




**Bkwd. Eqn.** $\dot{P}_{ij}(t)=\sum_{k\neq i} q_{ik}P_{kj}(t)-v_i P_{ij}(t)=\sum_{k\neq i} q_{ik}\left(P_{kj}(t)-P_{ij}(t)\right)$

$$\left.\begin{aligned}\dot{P}_{11}(t)&=q_{12}P_{21}-q_{12}P_{11}\\ \dot{P}_{21}(t)&=q_{21}P_{11}-q_{21}P_{21}\end{aligned}\right\}\quad \begin{aligned}&X(t)=1\\ &X(0)\text{ varies}\end{aligned}$$

$$\left.\begin{aligned}\dot{P}_{12}(t)&=q_{12}P_{22}-q_{12}P_{12}\\ \dot{P}_{22}(t)&=q_{21}P_{12}-q_{21}P_{22}\end{aligned}\right\}\quad \begin{aligned}&X(t)=2\\ &X(0)\text{ varies}\end{aligned}$$

**Fwd. Eqn.** $\dot{P}_{ij}(t)=\sum_{k\neq j} P_{ik}(t)q_{kj}-P_{ij}(t)v_j=\sum_{k\neq j}\left(P_{ik}(t)q_{kj}-P_{ij}(t)q_{jk}\right)$

$$\left.\begin{aligned}\dot{P}_{11}(t)&=q_{21}P_{12}-q_{12}P_{11}\\ \dot{P}_{12}(t)&=q_{12}P_{11}-q_{21}P_{12}\end{aligned}\right\}\quad \begin{aligned}&\text{conditioned on } x(0)=1\\ &X(t)\text{ varies}\end{aligned}$$

$$\left.\begin{aligned}\dot{P}_{21}(t)&=q_{21}P_{22}-q_{12}P_{21}\\ \dot{P}_{22}(t)&=q_{12}P_{21}-q_{21}P_{22}\end{aligned}\right\}\quad \begin{aligned}&\text{conditioned on } x(0)=2\\ &X(t)\text{ varies}\end{aligned}$$

# Why There Must be Two Sets of Differential Equations with One Solution

Backward Equations: $\frac{d\left[P(t)\right]}{dt} = [Q]\left[P(t)\right];\ [P(0)] = I$

Forward Equations: $\frac{d\left[P(t)\right]}{dt} = \left[P(t)\right][Q];\ \ [P(0)] = I$

The forward equation acts on each row of P(t) separately (corresponding to a single initial state), while the backward equation acts on each column separately (corresponding to a single final state). By differentiating term-by-term, one can check that the solution to the backward equation is the following generalization of the scalar solution to $\frac{dp(t)}{dt} = qp(t)$ , with init. cond. p(0):

$$\left[P(t)\right] = e^{[Q]t}\left[P(0)\right],$$

$$e^{[Q]t} = \sum_{k=0}^{\infty} Q^k t^k \Big/ k!$$

Begin with solution to the backward equation. Since

$$\left[P(t)\right] = e^{[Q]t}\left[P(0)\right] = \sum_{k=0}^{\infty}\frac{\left[Q\right]^k t^k}{k!}\left[P(0)\right],$$

if [P(0)] commutes with [Q] (and therefore with $[Q]^k$ )

$$\left[P(t)\right] = \sum_{k=0}^{\infty}\frac{\left[Q\right]^k t^k}{k!}\left[P(0)\right] = \sum_{k=0}^{\infty}\left[P(0)\right]\frac{\left[Q\right]^k t^k}{k!} = \left[P(0)\right]e^{[Q]t}$$

and therefore

$$\frac{d}{dt}\left[P(t)\right] = \frac{d}{dt}\sum_{k=0}^{\infty}\left[P(0)\right]\frac{\left[Q\right]^k t^k}{k!} = \sum_{k=1}^{\infty}\left[P(0)\right]\frac{\left[Q\right]^k t^{(k-1)}}{(k-1)!} =$$

$$\sum_{k=0}^{\infty}\left[P(0)\right]\frac{\left[Q\right]^k t^{(k)}}{(k)!}Q = [P(0)]e^{[Q]t}Q = \left[P(t)\right]Q$$

i.e., [P(t)] is also a solution to the forward equation.

# Conclusion

Though solutions to

$$\frac{d\left[P(t)\right]}{dt} = [Q]\left[P(t)\right]$$

and

$$\frac{d\left[P(t)\right]}{dt} = \left[P(t)\right][Q]$$

differ in general, they coincide whenever the initial conditions commute with [Q]. This is always the case when [P(0)] = I.

## REVERSIBILITY AND BACKWARD PROCESSES

The embedded chain of an ergodic Markov process has a backward chain with transition probabilities $\pi_i P_{ij}^* = \pi_j P_{ji}$.

Also, if $X(t) = j$, the time since the last transition is an exponential rv with rate $\nu_j$.

$$\begin{array}{c}
\text{State } i \quad | \qquad \text{State } j, \qquad | \quad \text{State } k \\
\hline
t_1 \qquad \textit{rate } \nu_j \qquad t_2
\end{array}$$

Thus, for ergodic Markov process in steady state, the backward process is also a Markov process. (Alternatively, look at sampled time chain.)

Backward rates $q_{ij}^*$ given by $\nu_i P_{ij}^*$.

Using $\pi_j = p_j \nu_j a$, ($a$ is normalizing factor, $\left(\sum_i p_j \nu_j\right)^{-1}$), $p_i \nu_i P_{ij}^* = p_j \nu_j P_{ji}$, which gives the rates for the reverse process

$$q_{ij}^* = v_i P_{ij}^* = \frac{v_i \pi_j P_{ji}}{\pi_i} = \frac{v_i \left(p_j v_j a\right) P_{ji}}{\left(p_i v_i a\right)} = \frac{p_j q_{ji}}{p_i}$$

**Reverse Process:** $p_i q_{ij}^* = p_j q_{ji}$

Definition: A Markov process is **reversible** if $q_{ij}^* = q_{ij}$ for all *i,j*.

**An ergodic Markov process is reversible iff its embedded chain is reversible.**

If the embedded chain is reversible $\left(P_{ij} = P_{ij}^*\right)$, then

$$q_{ij}^* = v_i P_{ij}^* = v_i P_{ij} = q_{ij}$$

so the process is reversible also. Conversely, if the process is reversible, so is the embedded chain.

**All Birth-Death processes are reversible because their embedded chains are reversible.**

**THEOREM** (*Guessing theorem*): Given $\{q_{ij}\}$, if prob. vector $\{p_i\}$ with $\sum_i p_i \nu_i < \infty$ and a set of non-neg. numbers $\{q_{ij}^*\}$ exist satisfying

$$\sum_j q_{ij} = \sum_j q_{ij}^* \text{ for all } i$$
$$p_i q_{ij} = p_j q_{ji}^* \text{ for all } i,j$$

then $\{p_i\}$ is set of steady state probabilities and $\{q_{ij}^*\}$ is set of rates for the backward transition process.

**PROOF:**

$$\sum_i p_i q_{ij} = p_j \sum_i q_{ji}^*$$

$$= p_j \sum_i q_{ji} = p_j \nu_j \quad \text{for all } j$$

Thus $\{p_i\}$ is steady state probability vector. QED

This theorem can be used to guess steady state probabilities and, when process is reversible, to show it. All B/D processes are reversible since embedded chains are reversible.

## Recall Burke's Theorem for Sampled Time M/ M/ m Chain

(We actually only proved it for M/ M/ 1 chains, but the proof for M/ M/ m chains is essentially identical.)

If the chain is in steady state with arrival rate $\lambda < \mu$,

a) the departure process is Bernoulli with rate $\lambda$

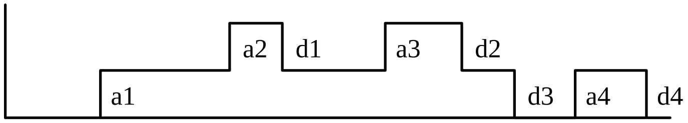
b) the state $X_n$ at any time $n\delta$ is independent of departures prior to $n\delta$ .

## Burke's Theorem For M/M/m with $\lambda < \mu$,

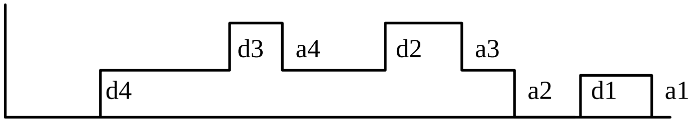
a) Departures are Poisson with rate $\lambda$

b) $X(t)$ is independent of departures before $t$

c) For FCFS, given departure at $t$, arrival time of that customer is independent of departures prior to $t$.

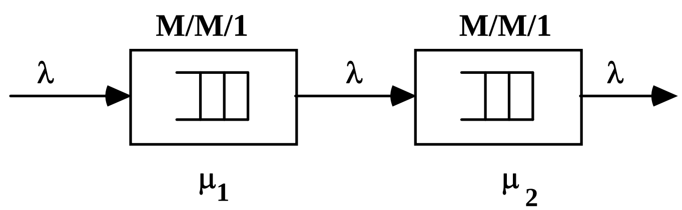



right moving (forward) M/M/1 process




left moving (backward) M/M/1 process

# INTERCONNECTIONS OF M/M/1 QUEUES IN STEADY STATE




Consider two steady-state queues in tandem; arrivals to first are Poisson at rate $\lambda$. Both have independent exponential service times, independent of all arrivals. Output of first queue is Poisson at rate $\lambda$, so **both queues are M/M/1**.

State $Y(t)$ of second queue is a function of its arrival epochs before $t$ and its service durations before $t$.

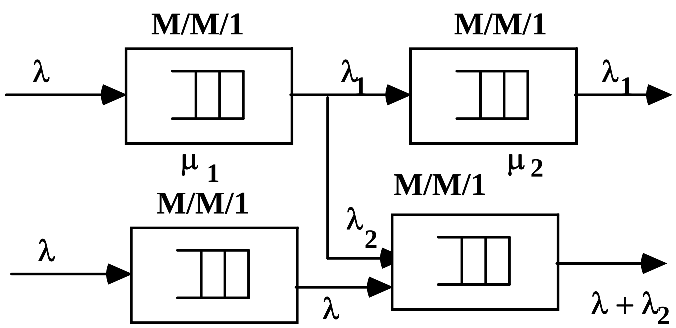
Those arrival epochs are the departures of first queue before $t$; they are independent of state $X(t)$ of first queue.

Thus $X(t)$ and $Y(t)$ are independent, and

$$P\big(X(t)=i,\, Y(t)=j\big)=(1-\rho_1)\rho_1^i(1-\rho_2)\rho_2^j;\ \rho_1=\lambda/\mu_1, \rho_2=\lambda/\mu_2$$

Note that $Y(t+\tau),\ \tau>0$, is not independent of $X(t)$.

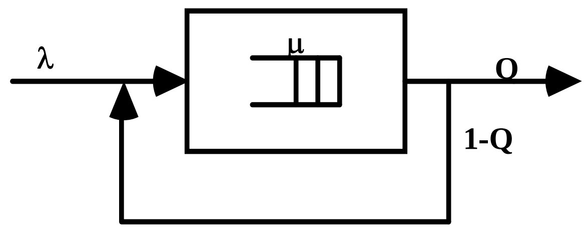
## Same argument applies to any network without feedback




Assume service intervals are all independent of each other and of arrivals; top and bottom arrivals Poisson; splitting is independent of processes.

Then queues on right have Poisson inputs, independent states at $t$.

## Feedback makes things more complex.




Assume $\lambda \ll \mu, Q \ll 1-Q$. Then inputs and outputs look like


Note that a feedback departure does not change the state of the system. Exogenous departures Poisson, at rate $\mu Q$.