Continious-Time Markov Chains (CTMC)

University of Tehran
Prepared by: ahmad Khonsari

Excerpt from : **CTMC ,** Ward Whitt

Definition of a CTMC

- For the continuous-time Markov chain $\{X(t) : t \geq 0\}$ with N states, the Markov property can be written as
- $P[X(s + t) = j \mid X(s) = i, X(u)=x(u), 0 \leq u < s \ ]= P[X(s + t) = j \mid X(s) = i \ ], i, j \in S, 0 \leq t < \infty,$
- and reflects the fact that the future state at time s+t only depends on the current state at time s .

transition probabilities functions

- We consider the special case of stationary transition probabilities functions (sometimes referred to as homogeneous transition probabilities functions), occurring when

$P[X(s + t) = j \mid X(s) = i] = P[X(t) = j \mid X(0) = i] = P_{ij}(t)$

for all states i and j and for all times s > 0 and t > 0;

- the independence of s characterizes the stationarity.

and

$\mathbf{P}(t) = [P_{ij}(t)]$

is called the ***transition probability matrix function*** (*TPMF*).(a function of time compared to TPM)

Exponential holding time in states of CTMC

- ***Proposition:*** T_i is exponentially distributed
- ***Proof:*** By time homogeneity, we assume that the process starts out in state i. For $s \geq 0$ the event $\{T_i > s\}$ is equivalent to the event $\{X(u) = i$ for $0 \leq u \leq s\}$.
- Similarly, for $s, t \geq 0$ the event $\{T_i > s+t\}$ is equivalent to the event $\{X(u) = i$ for $0 \leq u \leq s + t\}$.

Exponential holding time in states of CTMC

- Therefore,

  $P\,(T_i > s + t \,|\, T_i > s)$

  $= P(X(u) = i \text{ for } 0 \le u \le s + t \,|\, X(u) = i \text{ for } 0 \le u \le s)$

  $= P(X(u) = i \text{ for } s < u \le s + t \,|\, X(u) = i \text{ for } 0 \le u \le s)$

  $= P(X(u) = i \text{ for } s < u \le s + t \,|\, X(s) = i)$

  $= P(X(u) = i \text{ for } 0 < u \le t \,|\, X(0) = i)$

  $= P(T_i > t),$

Exponential holding time in states of CTMC

- where
  - the second equality follows from the simple fact that $P(A \cap B|A) = P(B|A)$, where we let $A = \{X(u) = i \text{ for } 0 \leq u \leq s\}$ and $B = \{X(u) = i \text{ for } s < u \leq s + t\}$.
  - the third equality follows from the Markov property.
  - the fourth equality follows from time homogeneity.

  Therefore, the distribution of T_i has the memoryless property, which implies that it is exponential.

Chapman-Kolmogorov equations

- ***Lemma 1.*** (Chapman-Kolmogorov equations) For all $s \geq 0$ and $t \geq 0$, $P_{i,j}(s + t) = \sum_k P_{i,k}(s)P_{k,j}(t)$
- Or in matrix notation $P(s + t) = P(s)P(t)$
- ***Proof***
- We can compute $P_{i,j}(s + t)$ by considering all possible places the chain could be at time s.
- We then condition and and uncondition, invoking the Markov property to simplify the conditioning; i.e.,

$P_{i,j}(s + t) = P(X(s + t) = j | X(0) = i)$

Chapman-Kolmogorov equations

- ***Proof (cntd.)***

$=\sum_k P(X(s+t)=j, X(s)=k \mid X(0)=i)$

$=$

$\sum_k P(X(s)=k \mid X(0)=i) P(X(s+t)=j \mid X(s)=k, X(0)=i)$

(conditioning on X(s)=k)

$= \sum_k P(X(s)=k \mid X(0)=i) P(X(s+t)=j \mid X(s)=k)$

(Markov property) (uncondition)

$= \sum_k P_{i,k}(s) P_{k,j}(t)$ (stationary transition probabilities)

▪

Describing a CTMC

- A CTMC is well specified if we specify:
- (1) its initial probability distribution –

p(X(0) = i) for all states i

- (2) its transition probabilities - $P_{i,j}(t)$ for all states i and j and positive times t.
- Thus we use these two elements to compute the distribution of X(t) for each t,

$P(X(t) = j) = \sum_i P(X(0) = i) P_{i,j}(t)$

Describing a CTMC

- Since the CTMC must be at any time in one of the N states, the analogous of DTMC is, for any state *i*

$$\sum_{j=1}^{N} P_{i,j}(t) = 1$$

constructing a CTMC model- four approaches(models)

- <u>for all four models:</u>
- the initial distribution are required and thus we focus on specifying the model beyond the initial distribution.
- The four models are equivalent: you can get to each from any of the others.
- Even though these four approaches are redundant, they are useful because they together give a different more comprehensive view of a CTMC.

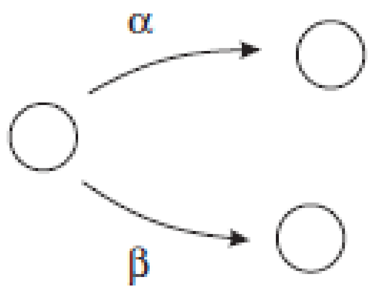
constructing a CTMC (model 1 : DTMC with Exponential Transition Times)

- For the DTMC with transition matrix P (looking at the transition epochs of the CTMC thus p_{ii}=0), the transition probabilities of the embedded chain

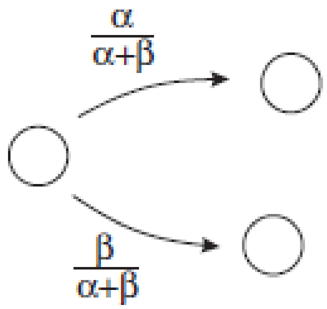
$$
\begin{aligned}
p_{i,j} &= \lim_{\Delta t \to 0} \mathrm{P}\{X_{t+\Delta t} = j \,|\, X_{t+\Delta t} \neq i, X_t = i\} \\
&= \lim_{\Delta t \to 0} \frac{\mathrm{P}\{X_{t+\Delta t} = j, X_{t+\Delta t} \neq i \,|\, X_t = i\}}{\mathrm{P}\{X_{t+\Delta t} \neq i \,|\, X_t = i\}} \\
&= \begin{cases} \dfrac{q_{i,j}}{\Sigma_j \, q_{i,j}} & i \neq j \quad \text{cf. } \mathrm{P}\{\min(X_1, \ldots, X_n) = X_i\} = \frac{\lambda_i}{\lambda_1 + \cdots + \lambda_n}, \text{ when } X_i \sim \mathrm{Exp}(\lambda_i) \\ 0 & i = j \end{cases}
\end{aligned}
$$

constructing a CTMC (model 1 : DTMC with Exponential Transition Times)

- Markov process, transition rates $q_{i,j}$ equilibrium probabilities $\tilde{\pi}_i$



- Embedded Markov chain, transition probabilities $p_{i,j}$
- equilibrium probabilities π_i



constructing a CTMC (model 1 : DTMC with Exponential Transition Times)

- For this DTMC the steady state probability vector is π, the unique probability vector satisfying the equation

$\pi = \pi P$ (1)

- Instead of having each transition take unit time, now we assume that the time required to make a transition from state i has an exponential distribution with rate q_i, and thus mean $1/q_i$, independent of the history before reaching state i.

constructing a CTMC (model 1 : DTMC with Exponential Transition Times)

- Relating the steady-state (stationary) probability vector $\tilde{\pi}$ of the CTMC to steady state probability vector of DTMC π

$$\tilde{\pi}_j = \frac{(\pi_j/q_j)}{\sum_k (\pi_k/q_k)} \qquad (2)$$

- Indeed, this first modelling approach corresponds to treating the CTMC as a special case of a semi-Markov process (SMP)
- We assume that there are no one-step transitions from

any state to itself in the DTMC (no self-loop); i.e., we assume that $P_{i,i}$ = 0 for all i (we look at the chain at transitions)

- this assumption is not critical,(see the third modelling)

constructing a CTMC (model 1 : DTMC with Exponential Transition Times)

Markov processes have no self-loops and their state transitions are characterized by a *generator matrix,* which is analogous to a probability transition matrix. The classification of states have analogous statements for Markov processes where the probability transition matrix is replaced by a generator matrix.
The generator matrix of a Markov process, denoted by *Q,* has entries that are the rates at which the process jumps from state to state. These entries are defined by

$$q_{i,j} = \lim_{\tau \to 0} \frac{P[X(t+\tau)=j \mid X(t)=i]}{\tau} \qquad i \neq j \qquad (3')$$

constructing a CTMC (model 1 : DTMC with Exponential Transition Times)

(We assume that the Markov process is time homogeneous and thus that (3’) is independent of t.)

The total rate out of state i is denoted by q_i and equals

$q_i = \sum_{j \neq i}^{\infty} q_{i,j}$ (4’)

The holding time of state i is exponentially distributed with rate q_i.

By definition, we set the diagonal entries of Q equal to minus the total rate,

$q_{i,i} = -q_i$ (5’)

This implies that the row sums of matrix Q equal 0.

constructing a CTMC (model 1 : DTMC with Exponential Transition Times)

- ***stationary probabilities*** in terms of the generator matrix. Using the results of EMC in SMP (i.e. $\tilde{\pi}_i = \frac{\pi^e{}_i \, E[S_i]}{\sum_{j \in S} \pi^e{}_j \, E[S_j]}$, $i \in S$,) and multiplying (2) by q_{ji} and summing yields [and using (5') $q_{i,i} = -q_i$ and $\pi_i = \sum_{j \neq i} \pi_j p_{j,i} = \sum_{j \neq i} \pi_j \frac{q_{j,i}}{q_j}$ slide 12]

$$\sum_{j=0}^{\infty} \tilde{\pi}_j q_{ji} = \frac{\sum_{j=0}^{\infty} (\pi_j q_{ji} / q_j)}{\sum_k (\pi_j / q_j)} = \frac{\sum_{j \neq i} (\pi_j q_{ji} / q_j) + \pi_i q_{ii} / q_i}{\sum_k (\pi_j / q_j)}$$

$$= \frac{\pi_i - \pi_i}{\sum_k (\pi_j / q_j)} = \mathbf{0}$$ [nel (8.65)]

- Rewriting in matrix form, shows that the stationary probabilities of a Markov process satisfy

πQ = 0,

with the additional normalization requirement that

‖π‖ = 1.

constructing a CTMC (model 2 : Transition Rates and ODE's)

- We look at the chain at any time (so we need to define zero-time transition probabilities, $P_{i,i}(0) = 1$ since there is no instant jump from a state)
- let P(0) = I, where I is the identity matrix; i.e., we set $P_{i,i}(0) = 1$ for all i and we set $P_{i,j}(0) = 0$ whenever $i \neq j$.
- We define $Q \equiv \lim_{h \downarrow 0} \frac{P(h)-I}{h} = \lim_{h \downarrow 0} \frac{P(h)-P(0)}{h}$

= P'(0+) (it is rate)

See Ross prob. Models 9th ed. ch 6 page 378

constructing a CTMC (model 2 : Transition Rates and ODE's)

- Thus the transition rate from state i to state j be defined in terms of the derivatives:

$$Q_{i,j} \equiv \lim_{h \downarrow 0} \frac{P_{i,j}(h) - P_{i,j}(0)}{h} = P'_{i,j}(0+) = \frac{dP_{i,j}(t)}{dt}\Big|_{t=0+} \qquad (3)$$

$$Q_{i,i} \equiv \lim_{h \downarrow 0} \frac{P_{i,i}(h) - P_{i,i}(0)}{h} = \frac{P_{i,i}(h) - 1}{h} = P'_{i,i}(0+) = \frac{dP_{i,i}(t)}{dt}\Big|_{t=0+}$$

- in most treatments of CTMC's instead of above, it is common to assume that

$P_{i,j}(h) = Q_{i,j}h + o(h)$ as $h \downarrow 0$ if $j \neq i$ (4) and

$P_{i,i}(h) - 1 = Q_{i,i}h + o(h)$ as $h \downarrow 0$, (5)

constructing a CTMC (model 2 : Transition Rates and ODE's)

- For finite state space, (for infinite state spaces under extra regularity conditions),we have
- $Q_{i,i} = -\sum_{j,j\neq i} Q_{i,j}(t)$ (6)

since $P_{i,j}(t)$ sum over j to 1

$\sum_{j=1}^{N} P_{i,j}(t) = 1$ *so* $P_{i,i}(t) + \sum_{j=1,j\neq i}^{N} P_{i,j}(t) = 1$

$\sum_{j=1,j\neq i}^{N} P_{i,j}(t) = 1 - P_{i,i}(t)$

Dividing by t and let t→0 we obtain (6)

And let

$-Q_{i,i} = q_i$ (7) for all i ,

constructing a CTMC (model 2 : Transition Rates and ODE's)

- Same as DTMC model that is specified via a transition probability matrix P, we can specify a CTMC model via the transition-rate matrix Q.
- In specifying the transition-rate matrix Q, it suffices to specify the off-diagonal elements

$Q_{i,j}$ for $i \neq j$, because the diagonal elements $Q_{i,i}$ are always defined by (6).

- The off-diagonal elements are always nonnegative, whereas the diagonal elements are always negative.
- Each row sum of Q is zero.

constructing a CTMC (model 2 : Transition Rates and ODE's)

- In fact, this approach to CTMC modelling is perhaps best related to modelling with ordinary differential equations,
- We may use Chapman-Kolmogorov equations to find the transition probabilities $P_{i,j}(t)$ from the transition rates $Q_{i,j} \equiv P'_{i,j}(0+)$
- To do this we use the two systems of ordinary differential equations (ODE's) generated by the transition rates namely, Kolmogorov forward and backward ODE's (defined next).

constructing a CTMC (model 2 : Transition Rates and ODE's)

Theorem 1. (Kolmogorov forward and backward ODE's) The transition probabilities satisfy both the Kolmogorov forward differential equations

$P'_{i,j}(s + t) = \sum_k P_{i,k}(s) Q_{k,j}(t) \quad for\ all\ i,j$ (9)

in matrix notation is the matrix ODE

$P'(t) = P(t)Q$ (10)

and the Kolmogorov backward differential equations

$P'_{i,j}(s + t) = \sum_k Q_{i,k}(t) P_{k,j}(t) \quad for\ all\ i,j$ (11)

in matrix notation is the matrix ODE

$P'(t) = QP(t)$ (12)

constructing a CTMC (model 2 : Transition Rates and ODE's)

- ***Proof:*** We apply the Chapman-Kolmogorov equations to write

P(t + h) = P(t)P(h) ,

and then do an asymptotic analysis as h ↓ 0.

- We subtract P(t) from both sides and divide by *h*, to get

$$\frac{P(t+h)-P(t)}{h}=P(t)\frac{P(h)-I}{h}$$

where *I* is the identity matrix

constructing a CTMC (model 2 : Transition Rates and ODE's)

- Recalling that I = P(0), we can let $h \downarrow 0$ to get the desired result (10).
- To get the backward equation (12), we start with

$$P(t + h) = P(h + t) = P(h)P(t)$$

and reason in the same way ▪

constructing a CTMC (model 2 : Transition Rates and ODE's)M/M/1 Queue

- ***Example*** (Transient Probabilities for the *M/M/1* Queue)
- Note that given that the initial state at time 0 was state i,
- Writing the forward equation for the *MIMl1* queue yields
- $\dfrac{dP_{i,0}(t)}{dt} = \mu P_{i,1}(t) - \lambda P_{i,0}(t),$
- $\dfrac{dP_{i,j}(t)}{dt} = \mu P_{i,j+1}(t) + \lambda P_{i,j-l}(t) - (\lambda + \mu)P_{i,j}(t).$

constructing a CTMC (model 2 : Transition Rates and ODE’s) M/M/1 Queue

- ***Example (Cntd.)*** The solution to these equations for this case is then given by

$$P_{i,j}(t) = e^{-(\lambda+\mu)} \left[\rho^{(j-i)/2} I_{j-i}(\alpha t) + \rho^{(j-i-1)/2} I_{j+i+1}(\alpha t) + (1-\rho)\rho^{j} \sum_{k=j+i+2} \rho^{-k/2} I_k(\alpha t) \right]$$

- where $\rho = \frac{\lambda}{\mu}$ and α=2 $\mu\sqrt{\rho}$ and

- $I_k(x) = \sum_{m=0}^{\infty} \frac{\left(\frac{x}{2}\right)^{k+2m}}{(k+m)!m!} \qquad k \geq -1$

is the series expansion for the modified Bessel function of the first kind.

constructing a CTMC (model 2 : Transition Rates and ODE's) M/M/1 Queue

- ***Example (Cntd.)*** It is difficult to have any intuition regarding the solution except for its limiting, and thus stationary, values.
- (no need for normalization eq. since initial condition P(0) =(0,…0, 1,0,,…) being in state i at t=0 ($p_{ii}(0)=1$) is an extra equation)
- In the third term (i.e. coefficient $(1-\rho)\rho^j$) we see factors corresponding to the stationary distribution.
- it must be $\lim_{t\to\infty} P_{i,j}(t) = (1-\rho)\rho^j$ independent of i.
- The solution of transient probabilities suggests that :
- $lim_{t\to\infty} e^{-(\lambda+\mu)}\rho^{(j-i)/2} I_{j-i}(\alpha t) = 0$
- $lim_{t\to\infty} e^{-(\lambda+\mu)}\rho^{(j-i-1)/2} I_{j+i+1}(\alpha t) = 0$
- $lim_{t\to\infty} e^{-(\lambda+\mu)} \sum_{k=j+i+2} \rho^{-k/2} I_k(\alpha t) = 1$

constructing a CTMC (model 2 : Transition Rates and ODE's)

- Equations (10 &12) are matrix ODE's in *t* that can be similarly solved as the scalar ODE f'(t)=qf(t) and have matrix exponential solution.
- (P(0) = I, the initial condition plays no role) In particular, as a consequence of Theorem 1, and If all entries of ***Q*** are bounded,(***Q*** is said to be **uniform:** the name comes from uniformization of CTMC in model 4) we have the following corollary.
- $Q_{i,j}=\infty$ means instantaneous jump from state I upon entering this state

constructing a CTMC (model 2 : Transition Rates and ODE's)

- ***Theorem 2***. (matrix exponential representation) The transition function can be expressed as a matrix-exponential function of the rate matrix Q, i.e.,

$$P(t) = e^{Qt} = \sum_{n=0}^{\infty} \frac{Q^n t^n}{n!} \quad (13)$$

This matrix exponential is the unique solution to the two ODE's with initial condition $P(0) = I$.

constructing a CTMC (model 2 : Transition Rates and ODE's)

- Proof: If we verify or assume that we can interchange summation and differentiation in (13), we can check that the displayed matrix exponential satisfies the two ODE's

$$P'(t) = \frac{d}{dt}\sum_{n=0}^{\infty}\frac{Q^n t^n}{n!} = \sum_{n=0}^{\infty}\frac{d}{dt}\frac{Q^n t^n}{n!}$$

$$= \sum_{n=0}^{\infty}\frac{nQ^n t^{n-1}}{n!} = \mathrm{Q}\sum_{n=0}^{\infty}\frac{Q^n t^n}{n!} = Qe^{Qt} \qquad \blacksquare$$

Summary of some Models of Markov Processes

Type of Process		Self-Loops	Holding Time
Semi-Markov Processes		No	Arbitrary
Markov chains	Model 1	Yes	$H_i = 1$
	Model 2	No	Geometric, $E[H_i] = (1 - p_{i,i})^{-1}$
Markov processes	Continuous time	No	Exponential, $E[H_i] = q_i^{-1}$
	Uniformized–Model 1	Yes	$H_i = 1$
	Uniformized–Model 2	No	Geometric, $E[H_i] = q_{\max}/q_i$