Continuous Time Markov Chains (CTMCs)

A continuous time, discrete state Markov process is a process \(\{X(t), t \geq 0\}\) with values in \(\{0, 1, 2, \ldots\}\), such that “given the present, the past and future are independent”, i.e., for all \(s, t \geq 0\) and all states \(i, j, x(u)\),

\[
P(X(t + s) = j \mid X(s) = i, X(u) = x(u), 0 \leq u < s) = P(X(t + s) = j \mid X(s) = i)
\]

stationary transition probability function (independent of \(s\)):

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

Compare:

A discrete time, discrete state Markov process is a process \(\{X_n, n = 0, 1, 2, \ldots\}\) with values in \(\{0, 1, 2, \ldots\}\), such that

\[
P(X_{n+m} = j \mid X_m = i, X_{m-1} = i_{m-1}, \ldots, X_0 = i_0) = P(X_{n+m} = j \mid X_m = i)
\]

\(n\)-step stationary transition probabilities: \(P_{ij}^{(n)}\)

If initial distribution \(P(X(0) = i)\) and \(P_{ij}(\cdot)\) are known then the distribution of \(X(t)\) is known for all \(t\):

\[
P(X(t) = i) = \sum_k P(X(t) = i \mid X(0) = k) \cdot P(X(0) = k)
\]

Compare: \(P(X_n = i) = \sum_k P_{ki}^{(n)} \cdot P(X_0 = k)\).
How does it work?

Suppose process enters state $i$ at time 0. Let

$$T_i = \text{time till next transition, out of state } i$$
$$= \text{sojourn time / holding time in state } i$$

Suppose $X(u) = i$ for $0 \leq u \leq s$, then due to definition:

$$P(T_i > s + t | T_i > s) = P(T_i > t)$$

So $T_i$ is memoryless, hence $T_i \sim \text{exp}(v_i)$ for some $v_i$.

At time $T_i$, the process leaves state $i$ and jumps to a state $j \neq i$ with some probability $P_{ij}$ as in DTMC (independent of $T_i$).

Hence alternative definition:

A CTMC is a process that...

(i) ...stays in state $i$ for a time $T_i \sim \text{exp}(v_i)$, after which it...

(ii) ...moves to some other state $j$ with some probability $P_{ij}$.

Here $P_{ii} = 0$ and $\sum_j P_{ij} = 1$ for all $i$.

(Note that $P_{ij}$ and $P_{ij}(t)$ are not the same!)
Examples

• Poisson process (only $i \to i + 1$)
  State space \{0, 1, 2, \ldots\}
  \[
  T_i \sim \exp(\lambda) \quad \text{so } v_i = \lambda, \quad i = 0, 1, 2, \ldots
  \]
  \[
  P_{ij} = 1, \quad j = i + 1
  \]
  \[
  = 0, \quad j \neq i + 1
  \]

• Pure birth process:
  Same $P_{ij}$ as in Poisson process, but with $v_i = \lambda_i$.
  E.g. Yule process (linear birth process): $\lambda_i = i\lambda$.

• Birth and death process (only $i \to i + 1$ or $i \to i - 1$)
  State space \{0, 1, \ldots\} or \{0, 1, \ldots, N\}
  Time till next ‘birth’ $i \to i + 1$ is $B_i \sim \exp(\lambda_i)$
  Time till next ‘death’ $i \to i - 1$ is $D_i \sim \exp(\mu_i)$
  Time till next transition is $\min(B_i, D_i) \sim \exp(v_i)$
  with $v_i = \lambda_i + \mu_i$

  \[
  P_{i, i+1} = P(B_i < D_i) = \frac{\lambda_i}{\lambda_i + \mu_i}
  \]

  \[
  P_{i, i-1} = P(B_i > D_i) = \frac{\mu_i}{\lambda_i + \mu_i}
  \]

State 0 has $\mu_0 = 0$, hence $v_0 = \lambda_0$ and $P_{0, -1} = 0, P_{0, 1} = 1
First entrance times in birth-death process

Let $T_{ij}$ be the time to reach $j$ from $i$. Then

$$ET_{i,i+1} = \frac{1}{v_i} + P_{i,i+1} \cdot 0 + P_{i,i-1}ET_{i-1,i+1}$$

$$= \frac{1}{\lambda_i + \mu_i} + \frac{\mu_i}{\lambda_i + \mu_i}(ET_{i-1,i} + ET_{i,i+1})$$

$$ET_{i,i+1} = \frac{1}{\lambda_i} + \frac{\mu_i}{\lambda_i}ET_{i-1,i}$$

This can be solved, starting from $ET_{0,1} = 1/\lambda_0$.

‘Similar’ for Variance (see book), but also (not in book):

$$Ee^{-sT_{i,i+1}} = \frac{v_i}{v_i + s} \left( P_{i,i+1} \cdot 1 + P_{i,i-1}Ee^{-sT_{i-1,i+1}} \right)$$

$$= \frac{\lambda_i}{\lambda_i + \mu_i + s} + \frac{\mu_i}{\lambda_i + \mu_i + s}Ee^{-sT_{i-1,i}} Ee^{-sT_{i,i+1}}$$

$$Ee^{-sT_{i,i+1}} = \frac{\lambda_i}{\lambda_i + \mu_i + s - \mu_i Ee^{-sT_{i-1,i}}}$$

This can be solved, starting from $Ee^{-sT_{0,1}} = \lambda_0/(\lambda_0 + s)$.

Finally, if $i < j$ we have $T_{i,j} = T_{i,i+1} + \cdots + T_{j-1,j}$, all terms independent!
Transition probabilities

We did: starting from $i$, when does process reach (fixed) $j$?
Answer: expectation/distribution of first entrance time $T_{ij}$

Now: starting from $i$, where is process after (fixed) time $t$?
Answer: transition probability function

$$P_{ij}(t) = P(X(t) = j|X(0) = i)$$

Known for Poisson process with rate $\lambda$ (let $j \geq i$):

$$P_{ij}(t) = P(j - i \text{ jumps in } (0, t]|X(0) = i)$$

$$= P(j - i \text{ jumps in } (0, t])$$

$$= e^{-\lambda t} \left( \frac{\lambda t}{(j-i)!} \right)^{j-i}$$

More general, pure birth process (let $j \geq i$):

$$P_{ij}(t) = P(X(t) = j|X(0) = i)$$

$$= P(X(t) \leq j|X(0) = i) - P(X(t) \leq j - 1|X(0) = i)$$

$$= P(X_i + \cdots + X_j > t) - P(X_i + \cdots + X_{j-1} > t)$$

with $X_i \sim \exp(\lambda_i)$ independent.

Rest is technical. Suppose $\lambda_i$ are different, then

$$E e^{-s(X_i + \cdots + X_j)} = \prod_{k=i}^{j} \frac{\lambda_k}{\lambda_k + s} = \sum_{k=i}^{j} \alpha_k \frac{\lambda_k}{\lambda_k + s}$$

with

$$\alpha_k = \prod_{r \neq k} \frac{\lambda_r}{\lambda_r - \lambda_k}$$

So

$$P(X_i + \cdots + X_j > t) = \sum_{k=i}^{j} \alpha_k e^{-\lambda_k t}$$
But what if \( P_{i,i+1} \neq 1? \)

In DTMC \( P_{ik}^{(n)} \) were found from difference equations, based on

- Chapman-Kolmogorov and
- ‘short-term behaviour’ \( P_{ij}^{(1)} \)

\[
P_{ij}^{(n)} = \sum_k P_{ik} P_{kj}^{(n-1)} \Rightarrow P^{(n)} = PP^{(n-1)}
\]

In CTMC, \( P_{ik}(t) \) are found similarly from differential equations.
So we need

- Chapman Kolmogorov: \( P_{ij}(t + s) = \sum_k P_{ik}(t) P_{kj}(s) \)
  (in matrix form: \( P(t + s) = P(t)P(s) \))
  Proof: ...

- Short-term behaviour for \( h \downarrow 0 \) (as in Poisson Process):

\[
P_{ii}(h) = 1 - v_i h + o(h), \quad P_{ij}(h) = v_i P_{ij} h + o(h), \quad j \neq i
\]

Proof: ...

**Notation:** let

\[
q_{ij} = v_i P_{ij} \quad \text{for } j \neq i
\]

These \( q_{ij} \)’s characterize the process:

\[
v_i = \sum_{j \neq i} q_{ij} \quad \text{and} \quad P_{ij} = \frac{q_{ij}}{\sum_{j \neq i} q_{ij}},
\]
Kolmogorov’s equations for $P_{ij}(t)$

\[
P_{ij}(t + h) = \sum_k P_{ik}(h)P_{kj}(t)
\]

\[
= (1 - v_i h + o(h))P_{ij}(t) + \sum_{k \neq i} (q_{ik} h + o(h))P_{kj}(t)
\]

⇒ Kolmogorov’s backward equation:
(always hold)

\[
P_{ij}'(t) = -v_i P_{ij}(t) + \sum_{k \neq i} q_{ik} P_{kj}(t)
\]

Also:

\[
P_{ij}(t + h) = \sum_k P_{ik}(t)P_{kj}(h)
\]

\[
= P_{ij}(t)(1 - v_j h + o(h)) + \sum_{k \neq j} P_{ik}(t)(q_{kj} h + o(h))
\]

⇒ Kolmogorov’s forward equation:
(almost always hold)

\[
P_{ij}'(t) = -P_{ij}(t)v_j + \sum_{k \neq j} P_{ik}(t)q_{kj}
\]

Initial conditions:

\[
P_{ii}(0) = 1,
\]

\[
P_{ij}(0) = 0, \quad j \neq i
\]
Generator matrix

Let *generator matrix* $Q$ be such that

$Q_{ii} = -v_i$

$Q_{ij} = q_{ij}, \quad i \neq j$

For finite-state Markov chain on \{0, 1, \ldots, N\}:

$$Q = \begin{pmatrix} -v_0 & q_{01} & q_{02} & \cdots & q_{0N} \\ q_{10} & -v_1 & q_{12} & \cdots & q_{1N} \\ q_{20} & q_{21} & -v_2 & \cdots & q_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_{N0} & \cdots & -v_N \end{pmatrix}$$

Also let

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

Then Kolmogorov’s equation in matrix form are

$P'(t) = Q \cdot P(t), \quad $ (backward)

$P'(t) = P(t) \cdot Q, \quad $ (forward)

and initial condition: $P(0) = I$.

Note: row sums of $Q$ are 0

row sums of $P(t)$ are 1 (stochastic matrix)
How to find $Q$ or the (infinitesimal) transition rates?

Example birth-death process:

$$v_i = \lambda_i + \mu_i$$

$$q_{i,i+1} = v_i P_{i,i+1} = (\lambda_i + \mu_i) \frac{\lambda_i}{\lambda_i + \mu_i} = \lambda_i$$

$$q_{i,i-1} = v_i P_{i,i-1} = (\lambda_i + \mu_i) \frac{\mu_i}{\lambda_i + \mu_i} = \mu_i$$

$$q_{i,j} = 0 \text{ for all other } j \neq i \text{ (since then } P_{i,j} = 0)$$

Hence, for a finite-state birth-death process, $Q$ is given by

$$
\begin{pmatrix}
-\lambda_0 & \lambda_0 & 0 \\
\mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 \\
\mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 \\
\vdots & \vdots & \vdots \\
0 & \mu_{N-1} & -(\lambda_{N-1} + \mu_{N-1}) & \lambda_{N-1} \\
& \mu_N & \mu_N & -\mu_N
\end{pmatrix}
$$

In general situation:

First determine state space $S$, then for all $i \in S$:

- $v_i$
- $P_{ij}$
- $q_{ij}$

$\implies Q$
Easy example: the two-state process (on-off source)

Often used, e.g. in telecommunications.
Some source alternates between off- and on-state.
Subsequent off- and on-times are independent, with
Off-time $\sim \exp(\alpha)$, \quad On-time $\sim \exp(\beta)$

Then $\{X(t)\}$ is a CTMC (birth-death process) with
state space $\{0, 1\}$ (0=off, 1=on), and

$$Q = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}$$

Solution of $P'(t) = P(t) \cdot Q$, $P_{ij}(0) = \delta_{ij}$:

$$P_{00}'(t) = -\alpha P_{00}(t) + \beta P_{01}(t)$$
$$= -\alpha P_{00}(t) + \beta (1 - P_{00}(t))$$

$$\Rightarrow P_{00}(t) = \frac{\beta}{\alpha + \beta} + \frac{\alpha}{\alpha + \beta} e^{-(\alpha+\beta)t}$$

$$\Rightarrow P_{01}(t), P_{10}(t), P_{11}(t) \text{ also known.}$$

Limiting probabilities:

$$\lim_{t \to \infty} P_{00}(t) = \lim_{t \to \infty} P_{10}(t) = \frac{\beta}{\alpha + \beta}$$

$$\lim_{t \to \infty} P_{01}(t) = \lim_{t \to \infty} P_{11}(t) = \frac{\alpha}{\alpha + \beta}$$
Less easy example: \( n \) identical on-off sources

Let \( X(t) = \# \) active sources at time \( t \).

How to find \( Q \)?

Suppose \( X(t) = i \)

- each source that is on turns off after time \( \sim \exp(\beta) \)
  \( \Rightarrow \) time until the first turns off \( \sim \exp(i\beta) \)

- each source that is off turns on after \( \sim \exp(\alpha) \)
  \( \Rightarrow \) time until the first turns on \( \sim \exp((n - i)\alpha) \)

So: \( \{X(t)\} \) is a birth-death process.

Next jump after time \( \sim exp(i\beta + (n - i)\alpha) \),
to state \( i + 1 \) w.p. \( \frac{(n-i)\alpha}{(n-i)\alpha + i\beta} \) or
to state \( i - 1 \) w.p. \( \frac{i\beta}{(n-i)\alpha + i\beta} \)

\[ Q = \cdots \]

Finding \( P_{ij}(t) \) is difficult, but for \( t \rightarrow \infty \)?
Next example:

Two independent machines are on (1) or off (0).
Let $X_i(t) =$ state of machine $i$ at time $t$. $i = 1, 2$
and $X(t) = (X_1(t), X_2(t))$

If $\{X_1(t)\}$ and $\{X_2(t)\}$ are CTMC’s as before, with parameters $(\alpha, \beta)$ and $(\lambda, \mu)$ respectively,
then also $\{X(t)\}$ is a continuous time Markov chain.
E.g. in state $(0, 0)$,
• machine 1 turns on after time $\sim \exp(\alpha)$, or
• machine 2 turns on after time $\sim \exp(\lambda)$,
whichever happens first. Hence:

$$v_{00} = \alpha + \lambda$$
$$q_{00,01} = v_{00}P_{00,01} = (\alpha + \lambda) \frac{\lambda}{\alpha + \lambda} = \lambda$$

and

$$Q = \begin{pmatrix}
00 & -(\alpha + \lambda) & \lambda & \alpha & 0 \\
01 & \mu & -(\alpha + \mu) & 0 & \alpha \\
10 & \beta & 0 & -(\beta + \lambda) & \lambda \\
11 & 0 & \beta & \mu & -(\beta + \mu)
\end{pmatrix}$$
Limiting behaviour

Accessibility, communicating states, communicating class, irreducibility, transience, recurrence, ...

... are all defined similarly as in DTMCs, with transition probabilities replaced by transition rates $q_{ij}$.

or:

... are defined via the embedded Markov chain with transition probabilities $P_{ij} = q_{ij}/v_i$.

In particular:

$\{X(t)\}$ is irreducible $\iff P_{ij}(t) > 0$ for all $i, j \in S$

(note: positive/null recurrence cannot be defined as in DTMCs)

Theorem (no proof):

If CTMC is irreducible and positive recurrent, then

$$P_j = \lim_{t \to \infty} P_{ij}(t) > 0$$

exists, independent of $i$.

Interpretations of $\{P_j\}$:

- Limiting distribution
- Long run fraction of time spent in $j$
- Stationary distribution

Last bullet: If $P(X(0) = i) = P_i$ for all $i \in S$, then $P(X(t) = j) = P_j$ for all $j \in S$ and all $t > 0$.

Proof: on board

Note 1: no periodicity issue.

Note 2: $P_i, P_{ij}$ and $P_{ij}(t)$ are all different.

Other notation for $P_j = \lim_{t \to \infty} P_{ij}(t)$ is $\pi_j$ (as in DTMC).
Balance equations

To find the limiting distribution, take limit \( t \to \infty \) in forward equations:

\[
\lim_{t \to \infty} P'_{ij}(t) = \lim_{t \to \infty} \left( \sum_{k \neq j} q_{kj} P_{ik}(t) - v_j P_{ij}(t) \right)
\]

becomes

\[
0 = \sum_{k \neq j} q_{kj} P_k - v_j P_j
\]

so limiting distribution satisfies

\[
\begin{align*}
  v_j P_j & = \sum_{k \neq j} q_{kj} P_k \\
  \sum_j P_j & = 1
\end{align*}
\]

(balance equations) (normalization)

Interpretation 'rate out = rate in':

Balance equation for set \( A \) (summing above over \( i \in A \)):

\[
\sum_{j \in A} P_j \sum_{i \notin A} q_{ji} = \sum_{i \notin A} P_i \sum_{j \in A} q_{ij}
\]
Balance equations in matrix form

Balance equations can also be found in matrix form from forward equation:
Let $t \to$ in $P'(t) = P(t)Q$:

$$0 = \begin{bmatrix}
P_0 & P_1 & P_2 & \ldots \\
P_0 & P_1 & P_2 & \ldots \\
P_0 & P_1 & P_2 & \ldots \\
\vdots & \vdots & \vdots & \ddots
\end{bmatrix} Q$$

where $P_i = \lim_{t \to \infty} P_{ki}(t)$ for all $k$

Hence, with $P = (P_0, P_1, \ldots)$ solve

$$\begin{cases} 
PQ = 0 \quad \text{(balance equations)} \\
\sum_j P_j = 1 \quad \text{(normalization)}
\end{cases}$$

Compare with DTMC:

$$\begin{cases} 
\pi P = \pi \quad \text{(balance equations)} \\
\sum_j \pi_j = 1 \quad \text{(normalization)}
\end{cases}$$

Birth death process:

$$-\lambda_0 P_0 + \mu_1 P_1 = 0$$
$$\lambda_0 P_0 - (\lambda_1 + \mu_1) P_1 + \mu_2 P_2 = 0$$
$$\lambda_1 P_1 - (\lambda_2 + \mu_2) P_2 + \mu_3 P_3 = 0$$
$$\vdots$$

or

$$\begin{align*}
\lambda_0 P_0 &= \mu_1 P_1 \\
(\lambda_1 + \mu_1) P_1 &= \lambda_0 P_0 + \mu_2 P_2 \\
(\lambda_2 + \mu_2) P_2 &= \lambda_1 P_1 + \mu_3 P_3 \\
&\vdots
\end{align*}$$
Summing first $i$ equations (or applying balance equations to set $A = \{0, 1, \ldots, i - 1\}$):

$$\lambda_{i-1}P_{i-1} = \mu_i P_i$$

Recursively:

$$P_i = \frac{\lambda_{i-1}}{\mu_i} P_{i-1} = \ldots = \frac{\lambda_{i-1}\lambda_{i-2} \cdots \lambda_0}{\mu_i \mu_{i-1} \cdots \mu_1} P_0$$

Two cases:

- Normalization $\sum_{i=0}^{\infty} P_i = 1$ is possible:

$$P_0 = \left[ \sum_{i=0}^{\infty} \frac{\lambda_{i-1}\lambda_{i-2} \cdots \lambda_0}{\mu_i \mu_{i-1} \cdots \mu_1} \right]^{-1}$$

Positive recurrent process

- Normalization $\sum_{i=0}^{\infty} P_i = 1$ is not possible:

$$\sum_{i=0}^{\infty} \frac{\lambda_{i-1}\lambda_{i-2} \cdots \lambda_0}{\mu_i \mu_{i-1} \cdots \mu_1} = \infty, \quad P_i = 0$$

Transient or null-recurrent process; no limiting distribution.
Embedded Markov chain

Let \( \{X(t), t \geq 0\} \) be some irreducible and positive recurrent CTMC with parameters \( v_i \) and \( P_{ij} = q_{ij}/v_i \). Let \( Y_n \) be the state of \( X(t) \) after \( n \) transitions.

Then \( \{Y_n, n = 0, 1, \ldots\} \) is DTMC with transition probs. \( P_{ij} \): Embedded chain. (Note: \( P_{jj} = 0 \)).

Now let \( P_i \) satisfy
\[
\sum_j P_j = 1 \quad \text{and} \quad P_j v_j = \sum_{i \neq j} P_i q_{ij}
\]
and let \( \pi_i \) satisfy
\[
\sum_j \pi_j = 1 \quad \text{and} \quad \pi_j = \sum_{i \neq j} \pi_i P_{ij}
\]
Is \( \pi_j = P_j \)?

No, not in general
Yes, if \( v_i \equiv v \) for all \( i \).

Example: on board
Uniformization

Idea:

• if CTMC has different \( v_i \)'s, choose a uniform \( v \geq \sup_{i \in S} v_i \)
• Compensate by adding 'self-transitions' \( i \rightarrow i \)
• Then define 'embedded' Markov chain

Suppose process is in state \( i \). Then:

• Stay in \( i \) for \( X \sim \exp(v) \)
• Then either leave \( i \) w.p. \( \frac{v_i}{v} \), or 'return' to \( i \) w.p. \( 1 - \frac{v_i}{v} \).

Consequences:

• Jumps occur according to Poisson Process, rate \( v \)
• Sojourn time in \( i \) is \( S = \sum_{k=1}^{N} X_i \) with \( N \sim \text{geom}(v_i/v) \) and \( X_i \) i.i.d. \( \sim \exp(v) \)
  So (e.g. by considering LST \( \phi_S(s) = g_N(\phi_X(s)) \)), we have \( S \sim \exp(v_i) \! \! ! \)

New DTMC has

\[
P_{ii} = 1 - \frac{v_i}{v} \\
P_{ij} = \frac{v_i}{v} P_{ij} \quad (j \neq i)
\]

Again an embedded process, but at event times of a Poisson process rate \( v \), with self transitions.

In this way almost (!) any CTMC (with parameters \( v_i, P_{ij} \)) can be viewed as a combination of a PP (with rate \( v \)) and a DTMC (with transition probabs. \( P_{ij}^* \))
Properties of uniformized Markov chain

Uniformizing DTMC has same stationary distribution as the original CTMC:

\[ \pi_j = \sum_{i \in S} \pi_i P_{ij}^* \]

\[ \pi_j = \pi_j (1 - \frac{v_j}{v}) + \sum_{i \neq j} \pi_i P_{ij} \frac{v_i}{v} \]

\[ v_j \pi_j = \sum_{i \neq j} v_i P_{ij} \pi_i \]

\[ v_j \pi_j = \sum_{i \neq j} q_{ij} \pi_i \]

Same set of equations as for \( P_i \); result follows by uniqueness.

Relation between transition probabilities:

Let \( N(t) \) be the Poisson process, at rate \( v \), then:

\[ P_{ij}(t) = P(X(t) = j | X(0) = i) \]

\[ = \sum_{n=0}^{\infty} P(X(t) = j | X(0) = i, N(t) = n) P(N(t) = n | X(0) = i) \]

\[ = \sum_{n=0}^{\infty} P_{ij}^*(n) e^{-vt} \frac{(vt)^n}{n!} \]

See example 6.21 for uniformization of two-state CTMC (parameters \( \lambda \) and \( \mu \)) at rate \( \lambda + \mu \) to find \( P_{00}(t) \), etcetera.

Other numerical applications: see Section 6.8 (but not for the exam!)