LTCC: Stochastic Processes

Maria De Iorio

Department of Statistical Science

University College London

m.deiorio@ucl.ac.uk

Outline

1. Examples:Invariant distributions and limiting behaviour
    - Example 1
    - Example 2
    - Example3
2. Continuous-time Markov chains
    - Definition & Notation
    - Exponential times
    - The Kolmogorov differential equations
    - Invariant distribution and limiting behaviour

Recall: For an irreducible, aperiodic, positive recurrent chain, let π be the (unique) invariant distribution. Then, for all i, j,

$$p_{ij}^{(n)} \to \pi_j \qquad \text{as } n \to \infty$$

Remarks

(a) P^n tends to a matrix with rows all the same and given by π.

(b) We also have $\Pr(X_n = j) = p_j^n \to \pi_j$ as $n \to \infty$ for all j irrespective of the distribution of X_0 . This follows since

$$p_j^{(n)} = \Pr(X_n = j) = \sum_i p_{ij}^{(n)} \Pr(X_0 = i) \to \pi_j \sum_i \Pr(X_0 = i) = \pi_j$$

because $\sum_i \Pr(X_0 = i) = 1$. Thus the distribution of X_n tends to π ; π is the equilibrium distribution.

(c) It can be shown that the proportion of the first n time points spent in state j tends to π_j as $n \to \infty$.

(d) A more general version of the result covering non irreducible chains is: an equilibrium distribution exists if and only if there is an ergodic class into which the chain is certain to be absorbed.

Example 1 (Weather)

Consider a simple weather model, in which there are two types of weather: rainy or sunny. If today is a rainy day, then the probability that tomorrow will also be a rainy day is 0.75, and the probability that tomorrow will be a sunny day is 0.25. On the other hand, if today is a sunny day, then the probability that tomorrow will be rainy day is 0.5, and the probability that tomorrow will also be a sunny day is 0.5.

Assume that our weather model follows a Markov chain. Let $X_n \in \{0, 1\}$ denote the weather on day n, where 0 denotes a rainy day and 1 denotes a sunny day. Then, the process $X_n, n = 0, 1, 2, \ldots$ is a Markov chain on the state space $\mathcal{S} = \{0, 1\}$, with transition matrix given by

$$P = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

We can check by induction that

$$P^n = \frac{1}{4^n}\begin{pmatrix} \frac{2^{2n+1}+1}{3} & \frac{4^n-1}{3} \\ \\ \frac{2(2^{2n}-1)}{3} & \frac{4^n+2}{3} \end{pmatrix}$$

We can verify by induction that $P^\infty = \lim_{n\to\infty} P^n$ exists and is equal to

$$P^\infty = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

Let the initial distribution be

$$p^{(0)} = \big(\Pr(X_0 = 0), \Pr(X_0 = 1)\big)$$

By Chapman-Kolmogorov theorem

$$p^{(n)} = p^{(0)} P^n = \big(\Pr(X_0 = 0), \Pr(X_0 = 1)\big) P^n = \frac{1}{4^n} \begin{pmatrix} \frac{2^{2n+1}+1}{3} & \frac{4^n-1}{3} \\ \\ \frac{2(2^{2n}-1)}{3} & \frac{4^n+2}{3} \end{pmatrix}$$

Then the limiting distribution

$$\begin{aligned} \lim_{n\to\infty} p^{(n)} &= p^{(0)} \lim_{n\to\infty} P^n = p^{(0)} P^{\infty} \\ &= \big(\Pr(X_0 = 0), \Pr(X_0 = 1)\big) \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix} \\ &= \big(\Pr(X_0 = 0) + \Pr(X_0 = 1)\big) \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix} = \pi \end{aligned}$$

Hence

$$\lim_{n\to\infty} \Pr(X_n = 0) = \frac{2}{3}, \qquad \lim_{n\to\infty} \Pr(X_n = 1) = \frac{1}{3}$$

- In particular, this means that regardless of how we start the chain, (i.e., how we set the initial state $p^{(0)}$), in the long run (i.e., as $n \to \infty$), the distribution on the states will always equal π (in this example, we always have $2/3$ chance of finding the chain in state 0 and $1/3$ chance of finding the chain in state 1). This in turn implies that in the long run, the chain will spend $2/3$ of its time in state 0 and $1/3$ of its time in state 1.
- If we set $p^{(0)} = \pi$, then the Chapman-Kolmogorov theorem gives

$$p^{(1)} = \pi P = \pi \text{ and } p^{(n)} = \pi P^n = \pi, \text{ for all } n \geq 1$$

that is, π is an invariant distribution.

Main questions:

(A) Does an equilibrium measure always exist? If not, under what conditions does it exist? Can there be more than one?
Answer: An invariant measure does not always exist. And yes, there can be more than one.

(B) What does an invariant measure or distribution tell me about the chain?
Answer: See next slides.

(C) How do I calculate π?
Answer: Left hand equations (solve $\pi = \pi P$) or detailed balance equations ($\pi_i p_{ij} = \pi_j p_{ji}$ for all $i, j \in \mathcal{S}$). If these equations have a solution then it is an invariant distribution for P.

Example 2 (Example to (A))

Consider a Markov chain $X_n, n = 0, 1, \ldots$, *with state space* $\mathcal{S} = \{0, 1\}$, *and with transition matrix*

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

By induction, we get for n *odd*

$$P^n = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and for n *even*

$$P^n = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Therefore, $\lim_{n\to\infty} P^n$ *does not exist.*

When does an equilibrium measure exist?
Recall Ergodic Theorem

Theorem 3

If a Markov chain is irreducible and ergodic (positive recurrent and aperiodic), then it has a unique invariant distribution π and for each $j \in \mathcal{S}$ and for $n \to \infty$, we have

$$p_{ij}^{(n)} \to \pi_j, \ \textit{for all } i \in \mathcal{S}$$

Moreover, by the Chapman-Kolmogorov equations, we have

$$p_j^{(n)} = \sum_{i \in \mathcal{S}} p_i^{(0)} p_{ij}^{(n)} \to \pi_j \sum_{i \in \mathcal{S}} p_i^{(0)} = \pi_j$$

so an equilibrium distribution exists (and is the same as the invariant distribution).

Consider the Markov Chain with state space $\mathcal{S} = \{0, 1, 2, \dots\}$ and with transition probabilities

$$p_{j,j+1} = p, p_{j0} = 1 - p, \text{ where } p \in (0, 1)$$

So

$$P = \begin{pmatrix} 1-p & p & 0 & \cdots & 0 \\ 1-p & 0 & p & \cdots & 0 \\ 1-p & 0 & 0 & p & \cdots \\ \vdots & & \ddots & & \vdots \end{pmatrix}$$

One can show that this chain is irreducible and aperiodic. Now, to compute the stationary distribution of this chain, we need to solve the following system of linear equations $\pi = \pi P$:

$$\pi_0 = (1-p)\sum_i \pi_i; \pi_1 = \pi_0 p; \pi_i = p\pi_{i-1}\dots; \sum_i \pi_i = 1$$

This gives

$$\pi_0 = 1 - p; \pi_1 = p(1-p); \pi_i = p^i \pi_0 = p^i(1-p), \forall i \in \mathcal{S}$$

discrete and continuous time Markov

		$\mathcal{T}$: discrete	$\mathcal{T}$: continuous
$\mathcal{S}$	discrete, countable	discrete-time Markov chain	continuous-time Markov chain
	continuous	✗	✗

In previous lectures, we considered discrete time Markov chains, typically:

$$\{X_n, n \in \mathcal{T}\}, \quad \mathcal{T} \subseteq \mathbb{N}$$

We will now look at continuous time Markov chains:

$$\{X(t), t \in \mathcal{T}\}, \quad \mathcal{T} \subseteq \mathbb{R}^+ := [0, \infty)$$

Again, we'll restrict the development to countable state space $\mathcal{S} \subseteq \mathbb{Z}$. I.e.

$$X = \{X_n\} : \Omega \times \mathbb{N} \mapsto \mathcal{S} \subseteq \mathbb{Z} \qquad \text{[discrete-time]}$$

$$X = \{X(t)\} : \Omega \times \mathbb{R}^+ \mapsto \mathcal{S} \subseteq \mathbb{Z} \qquad \text{[continuous-time]}$$

Definition 4 (continuous-time Markov chains)

A continuous-time stochastic process

$$X = \{X(t)\} : \Omega \times \mathbb{R}^+ \mapsto \mathcal{S} \subseteq \mathbb{Z}$$

is a <u>*continuous-time Markov chain*</u> *(CTMC) if it satisfies the* <u>*Markov property*</u>*, namely:*

$$\mathbb{P}\Big(\underbrace{X(\tau+t)=j}_{\textit{future}} \,\Big|\, \underbrace{X(\tau)=i}_{\textit{present}}, \underbrace{X(u)=x(u), u \in [o,\tau)}_{\textit{past}}\Big)$$

$$= \mathbb{P}\Big(\underbrace{X(\tau+t)=j}_{\textit{future}} \,\Big|\, \underbrace{X(\tau)=i}_{\textit{present}}\Big) =: \underbrace{p_{ij}(t)}_{\textit{transition prob.}}$$

Definition 5 (time homogeneity)

A CTMC X is <u>*time homogenous*</u> *if*

$$\mathbb{P}(X(\tau+t)=j|X(\tau)=i) = \mathbb{P}(X(t)=j|X(0)=i)$$

Definition 6 (transition matrix)

The matrix $\mathbf{P}(t) = (p_{ij}(t))_{ij \in \mathcal{S}}$ is called the transition probability matrix of X.

There is a transition matrix for each value of $t \in [0, \infty)$.

Theorem 7 (transition matrix is stoch. semigroup)

The family $\{\mathbf{P}(t), t \in \mathbb{R}^+\}$ is a stochastic semigroup, that is:

1. $\mathbf{P}(0) = \mathbf{I}$
2. $\mathbf{P}(t)$ *is a stochastic matrix, i.e.* $p_{ij}(t) \geq 0$ *and* $\sum_{j \in \mathcal{S}} p_{ij}(t) = 1$
3. *The Chapman Kolmogorov equations hold:* $\mathbf{P}(t + \tau) = \mathbf{P}(t)\mathbf{P}(\tau)$, *if* $t, \tau \geq 0$.

Proof See the discrete case. This is proved by conditioning on position at s.

Remark 8 (notation)

Write distribution of $X(t)$ *as the row vector* $\mathbf{p}(t) := (p_j(t))_{j\in\mathcal{S}}$ *with* $p_j(t) := \mathbb{P}\left(X(t) = j\right), j \in \mathcal{S}$. *I.e.*

$$\mathbf{p}(t) = \sum_{i\in\mathcal{S}} \mathbb{P}\big(X(0) = i\big)\, \mathbb{P}\big(X(t) = j \big| X(0) = i\big) = \mathbf{p}(0)\, \mathbf{P}(t)$$

Remark 9

I.e. $p_{ij}(h)$ *tells us the probability of going from* i *to* j *in the interval* $(\tau, \tau + h]$. *But as* $h \to 0$, *it is 'useful' to introduce some constraints...*

Question: How long will this process remain in a given state, say $j \in \mathcal{S}$?

Explicitly, suppose $X(0) = j$ and let $T_j =$ the time when we transition away from state j. To find the distribution of T_j, we let $t, \tau \geq 0$ and we show, buy means of the Markov property and the time-homogeneity, that

$$\Pr(T_j > \tau + t | T_j > \tau) = \Pr(T_j > t)$$

Therefore, T_j satisfies the loss of memory property, and is therefore exponentially distributed (since the exponential random variable is the only continuous random variable with this property).

Lemma 10 (memoryless property and Expon)

1. *$X \sim Expon(\lambda) \Rightarrow \mathbb{P}(X > \tau + t)|X > \tau) = \mathbb{P}(X > t)$, for $\tau, t > 0$.*
2. *X continuous random variable with $\mathbb{P}(X > \tau + t)|X > \tau) = \mathbb{P}(X > t)$, for $\tau, t > 0 \Rightarrow X \sim Expon(\lambda)$.*

The Chapman-Kolmogorov equations are

$$\mathbf{P}(\tau + t) = \mathbf{P}(\tau)\mathbf{P}(t)$$

- In continuous time, it is possible for chains to behave strangely, e.g. to run through an infinite number of states in finite time.
- We exclude such possibilities and consider only continuous-time Markov chains for which the transition probabilities, $p_{ij}(t)$, are differentiable at $t = 0$, with $p_{ii} \to 1$ and $p_{ij}(t) \to 0, i \neq 0$ for $t \downarrow 0$.

- Suppose the chain is in state $X(t) = i$ at time t. What happens in the interval $(t, t+h)$?
  - (a) nothing happens with prob. $p_{ii}(h) + \mathrm{o}(h)$, the error term accounts also for the possibility that the chain moves out of i and back in the interval
  - (b) the chain moves to a new state j with prob.

$$p_{ij}(h) + \mathrm{o}(h)$$

  So the probability of two or more transition in the interval $(t, t+h)$ is o(h.) (Proof Omitted.)

- We are interested in the behaviour of $p_{ij}(t)$ for small t. It can be shown that there exists constants q_{ij} such that

$$p_{ij}(t) = \delta_{ij} + q_{ij}t + o(t)$$

  where δ_{ij} is the Kronecker delta ($\delta_{ii} = 1$, $\delta_{ij} = 0$ for $j \neq i$). Since $p_{ij}(t)$ is a probability, $q_{ij} \geq 0, q_{ii} \leq 0$ and $\sum_j q_{ij} = 0$, so that $q_{ii} = -\sum_{j \neq i} q_{ij}$.

<u>Note:</u> An expression $A(t)$ is $o(t)$ as $t \to 0$ if $A(t)/t \to 0$ as $t \to 0$.

The matrix $\mathbf{Q} = (q_{ij})$ is called the **generator** of the chain and takes the role of the transition matrix $\mathbf{P}$ for discrete time chains.

Summary: for a chain starting from $X(t) = i$, we have:

- nothing happens in $(t, t+h)$ with probability $1 + q_{ii}h + o(h)$
- the chain jumps to a new state $j \neq i$ with probability $q_{ij}h + o(h)$
- since $\sum_j p_{ij}(t) = 1$ and

$$1 = \sum_j p_{ij}(t) \approx 1 + h \sum_j q_{ij}$$

so that

$$\sum_j q_{ij} = 0, \quad \forall i \qquad \text{or } \mathbf{Q1'} = \mathbf{0'}$$

- We can now say

$$\mathbf{P}(t) = I + \mathbf{Q}t + \mathbf{o}(t)$$

where $\mathbf{o}(t)$ is a matrix of $o(t)$ terms. By tome homogeneity, this implies that $\Pr(X(t+h) = j \mid X(t) = i) \approx \delta_{ij} + q_{ij}h$ for small h.

- **Forward Equations:**

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}$$

To show this, condition on the position at time t, and let h be small, so that

$$\mathbf{P}(t+h) = \mathbf{P}(t)\mathbf{P}(h) = \mathbf{P}(t)(\mathbf{I} + \mathbf{Q}h + \mathbf{o}(h))$$

and therefore

$$\mathbf{P}(t+h) - \mathbf{P}(t))/h \to \mathbf{P}(t)\mathbf{Q} \text{ as } h \to 0.$$

- **Backward Equations:**

$$\mathbf{P}'(t) = \mathbf{Q}\mathbf{P}(t),$$

which is proved similarly by conditioning on the position at time h.

- Subject to the boundary condition $\mathbf{P}(0) = \mathbf{I}$, these equations have formal solution (care is needed when $\mathcal{S}$ is not finite)

$$\mathbf{P}(t) = \exp(t\mathbf{Q}) := \sum_{n=0}^{\infty} \frac{t^n \mathbf{Q}^n}{n!}$$

where $\mathbf{Q}^0 = \mathbf{I}$. Thus the transition probabilities are specified by $\mathbf{Q}$, and so the chain is specified by $\mathbf{Q}$ and the initial distribution $\mathbf{p}(0)$.

We can also write

- **Forward Equations:**

$$p'(t) = \sum_k p_{ik}(t) q_{kj}$$

Backward Equations:

$$p'(t) = \sum_k q_{ik} p_{kj}(t)$$

Some results:

Theorem 11

When CTMC leaves state i, it jumps to state $j \neq i$ w.p. $-q_{ij}/q_{ii}$

Remark 12

Thus far, we know transition probabilities (conditioned on the event that a jump occurs) but we don't yet know anything about when these jumps occur

Definition 13 (holding time)

The holding time T_j of a continuous-time, homogenous stochastic process X is defined by

$$T_j := \inf\{t > 0 \colon X(\tau + t) \neq j | X(\tau) = j\}\,,\ \forall \tau\,.$$

Theorem 14

X homogenous CTMC with generator $\mathbf{Q}$. Then $T_j \sim Expon(-q_{jj})$.

Remark 15

A fuller picture now emerges. Given Q, the chain is in state i for exponentially distributed holding time, with mean $-1/q_{ii}$. when it leaves state i it jumps to state j with probability $-q_{ij}/q_{ii}$. It then stays in state j for exponentially distributed time, with mean $-1/q_{jj}$. When it leaves state j it jumps to, say, state k with probability $-q_{kj}/q_{jj}$, etc. Successive holding times are independent

A brief aside...

Definition 16 (embedded Markov chain)

Let Y_n be a sequence of states visited by $\{X(t)\}$. Then Y_n forms a discrete-time M.C. called the embedded Markov chain, aka jump chain (G&S), (of $\{X(t)\}$) with transition probabilities

$$p_{ij} := \begin{cases} -q_{ij}/q_{ii}, & j \neq i \\ 0, & j = i \end{cases}$$

This representation of the Markov process is particularly useful in simulations.

Remark 17 (summary so far: generator)

The generator determines:

1. *embedded (jump chain) probabilities* $-q_{ij}/q_{ii}$
2. *holding times* T_j
3. *transition matrix (previous result)*

In fact, it can be argued that all you need to specify a CTMC is

1. $\mathbf{p}(0)$ *[initial distribution]*
2. $\mathbf{Q}$ *[generator]*

Concept of invariant and equilibrium distributions also holds in continuous time. These, too, are determined by the generator.

Definition 18

A row vector $\boldsymbol{\pi}$ *s.t.* $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}(t), \forall t \geq 0$ *is called an invariant distribution*

Theorem 19

$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}(t) \Leftrightarrow \boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$

(e.g., use $\boldsymbol{\pi}\mathbf{P}(t) = \boldsymbol{\pi}\exp(t\mathbf{Q})$).

Remarks

(a) this implies that

$$\sum_j \pi_j q_{jk} = 0 \qquad \text{for all } k$$

(b) If $\mathbf{p}^{(0)} = \boldsymbol{\pi}$, then $\mathbf{p}^{(t)} = \boldsymbol{\pi}\mathbf{P}(t) = \boldsymbol{\pi}$ for all t.

Definition 20

An equilibrium distribution $\boldsymbol{\pi}$ *exists (as in discrete time case) if, for each* $j \in \mathcal{S}$,

$$p_{ij}(t) \to \pi_j,\ t \to \infty\, \forall\, i \in \mathcal{S}$$

If this is the case, $p'_{ij}(t) \to 0$ as $t \to \infty$. For each i, j pair, the forward equation is

$$p'_{ij}(t) = \sum_{k \in \mathcal{S}} p_{ik}(t) q_{kj}$$

so that, as $t \to \infty$, we obtain

$$0 = \sum_{k \in \mathcal{S}} \pi_k q_{kj} \ \text{ or } \ \mathbf{0} = \boldsymbol{\pi}\mathbf{Q}.$$

Thus $\boldsymbol{\pi}$ satisfies $\boldsymbol{\pi}\mathbf{Q} = \mathbf{0}$ and is an invariant distribution. (Note that taking the limit in the backward equations just gives

$$0 = \sum_{k \in \mathcal{S}} q_{ik} \pi_j = \pi_j \sum_{k \in S} q_{ik} = 0).$$

Just as in the discrete-time case, equilibrium and invariant distributions are closely linked.

Proposition 21

An equilibrium distribution is an invariant distribution.

Definition 22

A continuous-time Markov chain is defined to be irreducible if, for every i *and* j, $p_{ij}(t) > 0$ *for some* t.

Theorem 23 (ergodic theorem in continuous time)

For an irreducible continuous-time Markov chain:

(a) *if there exists an invariant distribution* $\boldsymbol{\pi}$ *then it is unique and* $p_{ij}(t) \to \pi_j$ *as* $t \to \infty$;

(b) *if there is no invariant distribution then* $p_{ij}(t) \to 0$ *as* $t \to \infty$.

Sketch proof G&S, p261.

Next Week: important examples.