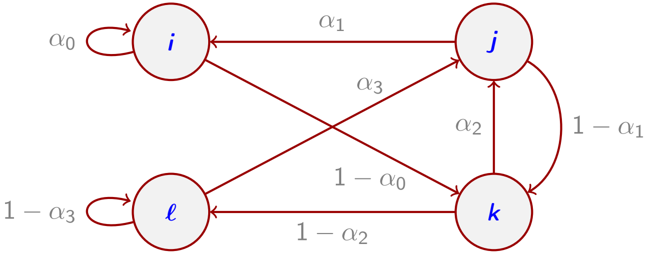
LTCC: Stochastic Processes

Maria De Iorio

Department of Statistical Science
University College London
m.deiorio@ucl.ac.uk



Course structure

- 5×2 hr lectures
- every Monday: Oct. 9 to Nov. 6 (inclusive) usually at 11:15 - 13:00hrs (except next week at 12!!!!!!)
- assignment:
  - handed out (emailed) in 'Spring-time'
  - typically, number of questions $\in \{1, 2, 3\}$
  - you will have *at least one week* to complete and hand in.
  - should only take $\leqslant$ 'a few' hours
  - state space of assessment $= \{0, 1, 2\}$ or $\{A, B, C\}$.

Recommended literature

- Grimmet & Stirzacker (2001) *Probability and Random Processes* [Oxford Uni. Press]
- Ross (1996) *Stochastic Processes* [Wiley]
- Daley & Vere-Jones (2003) *An Introduction to the Theory of Point Processes, Volume I. Elementary Theory and Methods* [Springer]

What is it about?

Introduction to

- discrete time Markov processes
- continuous time Markov processes

Motivation: Markov chains/processes

- physics thermodynamics & statistical mechanics
- chemistry enzyme activity models
- biology epidemic modelling
- sociology population dynamics
- audio music restoration; hidden Markov models used in speech recognition (and bioinformatics and too many other things to mention)
- operational research queueing theory, game theory, (baseball!?) etc
- telecommunications networks etc
- internet search engines, on-line fraud
- computational statistics Markov chain Monte Carlo (Markov random fields)
- economics modelling asset prices, market crashes, etc
- text language can be modelled as a Markov chain; ‘who wrote this text’; spam arms race, etc
- and many, many, more!

Outline

Definition 1

A stochastic process is a family of random variables $\{X(t), t \in \mathcal{T}\}$*, with some indexing set* $\mathcal{T}$ *(informally 'time').*

Remark 2

Indexing set can be

$$\begin{aligned} \mathcal{T} &\subseteq \mathbb{Z} && \textit{discrete} \\ \mathcal{T} = [a, b] &\subseteq \mathbb{R} && \textit{continuous} \end{aligned}$$

Definition 3

The set of all values that X *takes, namely* $\mathcal{S} := \{x \colon X(t) = x, t \in \mathcal{T}\}$ *is called the state space of* X.

Remark 4

For each t, $X(t)$ *is a r.v.:*

$$X(t) \colon \underbrace{\Omega}_{\text{sample space}} \mapsto \overbrace{\mathcal{S}}^{\text{state space}}$$

and we sometimes write $X \colon \Omega \times \mathcal{T} \mapsto \mathcal{S}$.

Note on terminology

If $\mathcal{S} = \{0, 1, 2, \ldots\}$, we refer to the process as integer valued or discrete state process.

If $\mathcal{S} =$ real line, we call $X(t)$ a real-valued stochastic process.

If $\mathcal{S}$ is Euclidean k space, $X(t)$ is called a k-vector process.

If $\mathcal{T} = \{0, 1, 2, \ldots\}$, we refer to $X(t)$ as a discrete time stochastic process.

If $\mathcal{T} = [0, \infty)$, we refer to $X(t)$ as a continuous time stochastic process.

Example 5

Flip a fair coin n *times. Let* $X(n) = \#$ *heads after* n *flips. Then* $X : \Omega \times \mathcal{T} \mapsto \mathcal{S}$ *with sample space* $\Omega = \{\text{'heads'}, \text{'tails'}\}$*, and* $\mathcal{T} = \mathbb{N}$*, with state space* $\mathcal{S} = \mathbb{N}_0$*.*

Definition 6 (Markov process)

A stochastic process is called a Markov process if it satisfies the <u>*Markov property*</u>*, namely*

$$
\begin{aligned}
& \mathbb{P}\Big(\underbrace{X(t) \leq x}_{\text{future}} \,\Big|\, \underbrace{X(t_n) = x_n}_{\text{present}}, \underbrace{X(t_{n-1}) = x_{n-1}, \ldots, X(t_0) = x(t_0)}_{\text{past}}\Big) \\
= \ & \mathbb{P}\Big(\underbrace{X(t) \leq x}_{\text{future}} \,\Big|\, \underbrace{X(t_n) = x_n}_{\text{present}}\Big)
\end{aligned}
$$

$\forall t_0 < t_1 < \ldots < t_n < t \in \mathcal{T}$ *and* $\forall x_0, x_1, \ldots, x_n, x \in \mathcal{S}$*. I.e. the past and future are conditionally independent, given the present.*

Example 7 (random walk)

Define

$$X(n+1) := \begin{cases} X(n)+1\,, & \textit{head} \\ X(n)-1\,, & \textit{tail} \end{cases} \qquad X(0) = 0$$

Then $\mathcal{T} = \mathbb{N}$, $\mathcal{S} \subseteq \mathbb{Z}$, *and* X *is Markov. [Given value of* $X(n)$, *the value of* $X(n+1)$ *is independent of* $X(n-1), \ldots, X(0)$*.]*

We assume that the random direction of each jump is independent of all earlier jumps.

Remark 8 (Markov is one-step away from independence)

Consider joint distribution of a stochastic process X:

$$\begin{aligned} \mathbb{P}\big(X(t_0),\ldots,X(t_n)\big) &= \mathbb{P}\big(X(t_0)\big) \\ &\times \mathbb{P}\big(X(t_1)\big|X(t_0)\big) \\ &\times \mathbb{P}\big(X(t_2)\big|X(t_1),X(t_0)\big) \\ &\times \mathbb{P}\big(X(t_3)\big|X(t_2),X(t_1),X(t_0)\big) \\ &\vdots \qquad \vdots \\ &\times \mathbb{P}\big(X(t_n)\big|X(t_{n-1}),X(t_{n-2}),\ldots,X(t_0)\big) \end{aligned}$$

If X is Markov, then this collapses to

$$\mathbb{P}\big(X(t_0),\ldots,X(t_n)\big) = \mathbb{P}\big(X(t_0)\big)\prod_{i=1}^{n}\mathbb{P}\big(X(t_i)\big|X(t_{i-1})\big)$$

Hence Markov processes are one-step away from independence.

Note on terminology

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

Definition 9 (discrete-time Markov chain)

A discrete-time process $\{X_n, n \in \mathbb{N}_0\}$ with countable discrete state space $\mathcal{S}$ is a Markov chain if

$$\mathbb{P}\Big(\underbrace{X_{n+m} = j}_{\text{future}} \,\Big|\, \underbrace{X_n = i_n}_{\text{present}}, \underbrace{X_{n-1} = i_{n-1}, \ldots, X_0 = i_0}_{\text{past}}\Big)$$

$$= \mathbb{P}\Big(\underbrace{X_{n+m} = j}_{\text{future}} \,\Big|\, \underbrace{X_n = i_n}_{\text{present}}\Big), \forall i_0, i_1, \ldots, i_n, j \in \mathcal{S}; \forall m, n \in \mathbb{N}_0$$

Definition 10 (transition probability)

$\mathbb{P}(X_{n+1} = j | X_n = i)$ *is known as the (one-step)* *transition probability*.

Generally, the one-step transition probability depends on three indexes: i, j, and n. We will consider the case where this is constant w.r.t. n.

Definition 11 (time-homogeneity)

Markov chain X *is* *time-homogenous* *if*

$$\mathbb{P}\left(X_{n+1} = j \middle| X_n = i\right) = \mathbb{P}\left(X_1 = j \middle| X_0 = i\right), \forall n \in \mathbb{N}_0 \text{ and } i, j \in \mathcal{S}$$

P

For $\mathcal{S} = \{0, 1, \ldots, N-1\}$, the transition matrix $\mathbf{P} \in [0,1]^{N \times N}$ is

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0,N-1} \\ p_{10} & p_{11} & \cdots & p_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N-1,0} & p_{N-1,1} & \cdots & p_{N-1,N-1} \end{bmatrix}$$

Definition 13

A stochastic matrix is a matrix $\mathbf{P} = (p_{ij})_{i,j \in \mathcal{S}}$ *which satisfies*

1. $p_{ij} \geq 0, \forall i, j \in \mathcal{S}$ *[$\mathbf{P}$ has non-neg. entries]*
2. $\sum_{j \in \mathcal{S}} p_{ij} = 1, \forall i \in \mathcal{S}$ *[rows of $\mathbf{P}$ sum to 1]*

Theorem 14

The transition matrix is a stochastic matrix.

Proof p_{ij} is a probability $\Rightarrow$ 1 and $\sum_{j \in \mathcal{S}} \mathbb{P}(X_1 = j | X_0 = i) = 1 \Rightarrow$ 2.

bayer A and player B play a series of games. Now

$$\begin{aligned} \mathbb{P}(A \text{ wins}) &= p \\ \mathbb{P}(B \text{ wins}) &= 1 - p =: q \end{aligned}$$

where the outcome of each game is independent. Find the transition probability matrix.

b $\ddots$ [

$$\begin{matrix} p_{-1,-1} & p_{-1,0} & p_{-1,1} \\ p_{0,-1} & p_{0,0} & p_{0,1} \\ p_{1,-1} & p_{1,0} & p_{1,1} \end{matrix}$$

$$\ddots$$

$$\begin{matrix} \ddots & & & \\ q & 0 & p & \\ & q & 0 & p \\ & & & \ddots \end{matrix} \qquad \square(0)$$

Remark 15 (marginal distribution notation)

Denote marginal probability that the chain is in state j at time n as:

$$p_j^{(n)} := \mathbb{P}(X_n = j)$$

Then the row vector

$$\mathbf{p}^{(n)} := \left(p_j^{(n)}\right)_{j\in\mathcal{S}}$$

is the distribution (pmf) of X_n with initial distribution $\mathbf{p}^{(0)}$. E.g., for $\mathcal{S} = \mathbb{Z}$, the marginal distribution (at time n) is $\mathbf{p}^{(n)} = \left[\ldots, p_{-1}^{(n)}, p_0^{(n)}, p_1^{(n)}, \ldots\right]$.

Example 16

Remark 17

Next week or so, we will consider the following:

- $\mathbf{p}^{(n)}$ *distn. of* A*'s lead over* B *after* n *games*
- *In particular* $\lim_{n\to\infty} p_0^{(n)}$*, prob. that* A *and* B *win equal #games for large* n *(in the long run)*
- *distn. of #games played until* $X_n = 0$*:* $\left[\mathbb{P}(T_0 = n)\right]_{n\in\mathcal{T}}$ *where* $T_j := \min\{n > 0 \colon X_n = j\}$
- $\mathbb{E}(T_0)$ *mean return time to state* 0*.*
- $X_n \to ?$

Before that, it will be instructive to introduce a couple more examples.

Example 18 (gambler's ruin)

A and B play for chips; loser pays winner 1 chip.

$$\begin{aligned} \mathbb{P}(A \text{ wins}) &= p \\ \mathbb{P}(B \text{ wins}) &= 1 - p =: q \end{aligned}$$

A starts with a chips, B starts with b chips. Let X_n = #chips A has after n games. Game ends when A or B is bankrupted $(p_{00} = p_{a+b,a+b} = 1)$.

X is Markov with state space

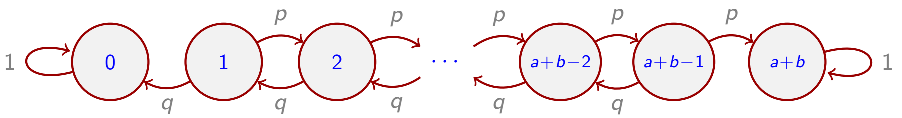
$$\mathcal{S} = \{ \underbrace{0,}_{A \text{ is bankrupted}} 1, \ldots, \underbrace{a+b}_{B \text{ is bankrupted}} \}$$

Transition matrix:

$$\mathbf{P} = \begin{bmatrix} 1 & & & & & \\ q & 0 & p & & & \\ & q & 0 & p & & \\ & & \ddots & \ddots & \ddots & \\ & & & q & 0 & p \\ & & & & & 1 \end{bmatrix}$$

If/when X_n reaches state 0 or $a + b$, it stays there. The states 0 and $a + b$ are called absorbing states.

gambler's ruin state transition diagram



Remark 19

Finite state space does not necessarily imply existence of absorbing states.

Example 20 (urn model)

Consider 2 urns. Urn A contains N white balls. Urn B contains N black balls. At each turn (time index $n = 1, 2, \ldots$) a ball is chosen at random from each urn and the two balls are interchanged. Denote the # of black balls in urn A, after nth interchange, by $\{X_n, n \in \mathbb{N}_0\}$.

X is Markov. $X_0 = 0$ (urn A starts out with 0 black balls.) State space: $\mathcal{S} = \{0, \ldots, N\}$. Transition probabilities:

- one more black ball in A: $p_{i,i+1} = \mathbb{P}\big(X_{n+1} = i + 1 \big| X_n = i\big)$
- one less black ball in A: $p_{i,i-1} = \mathbb{P}\big(X_{n+1} = i - 1 \big| X_n = i\big)$
- no change in # black balls in A: $p_{i,i} = \mathbb{P}\big(X_{n+1} = i \big| X_n = i\big)$

- one more black ball in A: $p_{i,i+1} = \mathbb{P}\big(X_{n+1} = i+1 \big| X_n = i\big)$
$$= \mathbb{P}\big(\text{white from } A, \text{black from } B \big| N-i \text{ white in } A,\ N-i \text{ blacks in } B\big)$$
$$= \left(\frac{N-i}{N}\right)^2 = \left(1 - \frac{i}{N}\right)^2 =: q_i^*$$

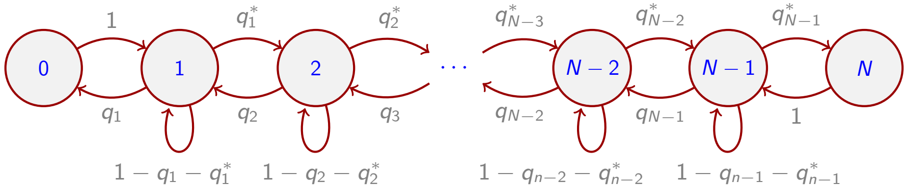
- one less black ball in A: $p_{i,i-1} = \mathbb{P}\big(X_{n+1} = i-1 \big| X_n = i\big)$
$$= \mathbb{P}\big(\text{black from } A, \text{white from } B \big| i \text{ black in } A,\ i \text{ white in } B\big)$$
$$= \left(\frac{i}{N}\right)^2 =: q_i$$

- no change in # black balls in A: $p_{i,i} = 1 - q_i - q_i^*$

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & & & & \\ q_1 & 1 - q_1 - q_1^* & q_1^* & & & \\ & q_2 & 1 - q_2 - q_2^* & q_2^* & & \\ & & \ddots & \ddots & \ddots & \\ & & & q_{N-1} & 1 - q_{N-1} - q_{N-1}^* & q_{N-1}^* \\ & & & & & 0 \end{bmatrix}$$

See also state transition diagram. I.e., if/when you are in state 0 (resp. N), you never stay there, you always go back to state 1 (resp. $N-1$). States 0 and N are called reflecting barriers.

urn state transition diagram



Example 21 (weather 'forecasting')

1. *2 possible weather conditions on any single day: {rainy, sunny}.*
2. *Tomorrow's weather depends* **only** *on today's weather.*
3. $\mathbb{P}\big(\textit{rain tomorrow}\big|\textit{rain today}\big) = \alpha$
4. $\mathbb{P}\big(\textit{sunny tomorrow}\big|\textit{sunny today}\big) = \beta$

Find transition probabilities.

Solution

$$① \Rightarrow Y_n := \begin{cases} 1, & \text{sunny on } n\text{th day} \\ 0, & \text{rains on } n\text{th day} \end{cases} \qquad \mathcal{S} = \{0, 1\}$$

② $\Rightarrow Y_n$ Markov, and

$$③ \Rightarrow p_{00} = \alpha, \quad p_{01} = 1 - \alpha, \qquad ④ \Rightarrow p_{10} = 1 - \beta, \quad p_{11} = \beta$$

I.e.

$$\mathbf{P} = \begin{bmatrix} \alpha & 1 - \alpha \\ 1 - \beta & \beta \end{bmatrix}$$

Example 22 (weather 'forecasting' II)

- *Now assume tomorrow's weather depends only on weather of today* **and** *yesterday*

Find transition probabilities.

Solution

$$\mathbb{P}\left(Y_{n+1} \middle| Y_n, Y_{n-1}, \dots\right) = \mathbb{P}\left(Y_{n+1} \middle| Y_n, Y_{n-1}\right) \neq \mathbb{P}\left(Y_{n+1} \middle| Y_n\right)$$

Hence, Y_n not Markov (at least the way we have defined it). However...

Remark 23

Define $\{X_n\}$ *s.t.*

$$X_n := Y_{n-1} + 2Y_n, \qquad X_0 = 0$$

i.e. for $Y_{n-1} = i$ *and* $Y_n = j$, *we have* $X_n := i + 2j$. *then* X *is Markov.*

Proof ...

$$
\begin{aligned}
\mathbb{P}\big(X_{n+1}\big|X_n, X_{n-1}, \dots\big) &= \mathbb{P}\big(Y_{n+1}, Y_n\big|Y_n, Y_{n-1}, \dots\big) \\
&= \mathbb{P}\big(Y_{n+1}\big|Y_n, Y_{n-1}, \dots\big) \\
&= \mathbb{P}\big(Y_{n+1}\big|Y_n, Y_{n-1}\big) \\
&= \mathbb{P}\big(Y_{n+1}, Y_n\big|Y_n, Y_{n-1}\big) \\
&= \mathbb{P}\big(X_{n+1}\big|X_n\big) \qquad \blacksquare
\end{aligned}
$$

Homework: compute transition matrix of X and draw state transition diagram

Recall notation from Remark 15:

$$\begin{aligned} p_j^{(n)} &= \mathbb{P}(X_n = j) \\ p_{ij} &= \mathbb{P}(X_{n+1} = j | X_n = i) \end{aligned}$$

Definition 24 (n-step transition probabilities)

Define the <u>n-step transition probabilities</u> by

$$p_{ij}^{(n)} := \mathbb{P}(X_n = j | X_0 = i)$$

and the <u>n-step transition matrix</u> by

$$\mathbf{P}^{(n)} := \left(p_{ij}^{(n)}\right)_{i,j\in\mathcal{S}}$$

Remark 25

$\mathbf{P}^{(0)} = I$, *(the identity matrix):* $p_{ij}^{(0)} = \mathbb{P}(X_0 = j | X_0 = i) = \delta_{ij}$

Remark 26

X *time-homogeneous*

$\mathbb{P}(X_{n+m} = j | X_m = i) = \mathbb{P}(X_n = j | X_0 = i) = p_{ij}^{(n)} \quad \forall n, m \in \mathcal{T};\ i, j \in \mathcal{S}$

Theorem 27 (Chapman-Kolmogorov (cf p215, G&S))

$\mathbf{P}^{(m+n)} = \mathbf{P}^{(m)}\mathbf{P}^{(n)}$, *for all* $m, n \in \mathbb{N}_0$

Proof I.e., we want to show that:

$$p_{ij}^{(m+n)} = \sum_{k \in \mathcal{S}} p_{ik}^{(m)} p_{kj}^{(n)}$$

for all $m, n \in \mathbb{N}_0$ and $i, j \in \mathcal{S}$. LHS is

$$\begin{aligned}
p_{ij}^{(m+n)} &= \mathbb{P}\big(X_{n+m} = j \big| X_0 = i\big) \\
&= \sum_{k \in \mathcal{S}} \mathbb{P}\big(X_{n+m} = j, X_m = k \big| X_0 = i\big) \\
&= \sum_{k \in \mathcal{S}} \mathbb{P}\big(X_{n+m} = j \big| X_m = k, \cancel{X_0 = i}\big)\mathbb{P}\big(X_m = k \big| X_0 = i\big) \quad [Markov] \\
&= \sum_{k \in \mathcal{S}} \mathbb{P}\big(X_n = j \big| X_0 = k\big)\mathbb{P}\big(X_m = k \big| X_0 = i\big) \quad [t-homog.] \\
&= \sum_{k \in \mathcal{S}} p_{ik}^{(m)} p_{kj}^{(n)} \qquad \blacksquare
\end{aligned}$$

Corollary 28

$\mathbf{P}^{(n)} = \mathbf{P}^n$

Proof (I.e., this says the nth step transition matrix is the nth power of the one-step transition matrix.) From Chapman-Kolmogorov

$$\mathbf{P}^{(m+n)} = \mathbf{P}^{(m)}\mathbf{P}^{(n)}$$

But

$$\begin{aligned}
\mathbf{P}^{(n)} = \mathbf{P}^{(n-1+1)} &= \mathbf{P}^{(n-1)}\mathbf{P}^{(1)} \\
&= \mathbf{P}^{(n-1)}\mathbf{P} \\
&= \mathbf{P}^{(n-2)}\mathbf{P}^{(1)}\mathbf{P} \\
&= \mathbf{P}^{(n-2)}\mathbf{P}\mathbf{P} \\
&= \mathbf{P}^{(n-2)}\mathbf{P}^2 \\
&\vdots \quad \vdots \\
&= \mathbf{P}^n \qquad \blacksquare
\end{aligned}$$

Corollary 29

$\mathbf{p}^{(n)} = \mathbf{p}^{(0)}\mathbf{P}^{n}$

Proof I.e. this says distn. at time n is row-matrix product of initial distn. with n-step transition matrix. We have

$$\begin{aligned} p_j^{(n)} = \mathbb{P}(X_n = j) &= \sum_{i\in\mathcal{S}} \mathbb{P}\big(X_n = j, X_0 = i\big) \\ &= \sum_{i\in\mathcal{S}} \mathbb{P}\big(X_0 = i\big)\mathbb{P}\big(X_n = j \big| X_0 = i\big) \\ &= \sum_{i\in\mathcal{S}} p_i^{(0)} p_{ij}^{(n)} \end{aligned}$$

i.e. $\mathbf{p}^{(n)} = \mathbf{p}^{(0)}\mathbf{P}^{(n)}$. Chapman-Kolmogorov (Corollary 28) completes the proof. ■