LTCC: Stochastic Processes

Maria De Iorio

Department of Statistical Science

University College London

m.deiorio@ucl.ac.uk

Outline

1. Summary
2. Classification of States
3. Classification of chains
4. Invariant distribution

$$\mathbb{P}(X_n = i, \text{for some } n \geq 1 \mid X_0 = i\} = \begin{cases} = 1 & \text{State } i \text{ is recurrent} \\ < 1 & \text{State } i \text{ is recurrent} \end{cases}$$

summary of previous lecture

- $f_{ij}(n) :=$ prob. 1st visit from j to i occurs at time $n \geq 1$.
- $f_{ij} := \sum_{n \in \mathbb{N}} f_{ij}(n) =$ prob. ever visiting j from i $\quad [\mathbb{N} := \{1, 2, \ldots\}]$
- $f_{jj} =$ prob. ever returning to j
- $T_j :=$ 1st hitting time of state j
- $\mathbb{P}\big(T_j = n \big| X_0 = i\big) = f_{ij}(n), n \geq 1$
- j recurrent $\Leftrightarrow f_{jj} = 1 \Leftrightarrow \sum_n p_{jj}^{(n)} = \infty$
- j transient $\Leftrightarrow f_{jj} < 1 \Leftrightarrow \sum_n p_{jj}^{(n)} < \infty \Rightarrow p_{ij}^{(n)} \to 0, \forall\, i \in \mathcal{S}$

Definition 1 (mean return time)

Define the mean return time to state j as

$$\mu_j := \mathbb{E}\left(T_j \middle| X_0 = j\right) = \sum_{n=1}^{\infty} n\, f_{jj}(n)$$

Remark 2 (a note on mean return time and recurrence/transience)

For j transient, $f_{jj} < 1 \Rightarrow$

$$\mathbb{P}\big(T_j < \infty \big| X_0 = j\big) = \sum_{n \in \mathbb{N}} f_{jj}(n) = f_{jj} < 1$$

i.e. the probability of returning in finite time is strictly less than 1. Therefore it is possible that the return is unbounded:

$$\mathbb{P}\big(T_j = \infty \big| X_0 = j\big) > 0$$

I.e. $\lim_{n\to\infty} f_{jj}(n) > 0$. *The mean return time to a transient state j is*

$$\mu_j = \mathbb{E}\big(T_j \big| X_0 = j\big) = \sum_{n=1}^{\infty} n\, f_{jj}(n) \not< \infty$$

because $f_{jj}(n) \to > 0 \Rightarrow n\, f_{jj}(n) \to \infty$.
However, for j recurrent, μ_j may or may not diverge.

Definition 3 (null and positive recurrence)

j *recurrent. Then*

$$j \text{ is } \begin{cases} \textit{null-recurrent}, & \mu_j = \infty \\ \textit{positive (non-null) recurrent}, & \mu_j < \infty \end{cases}$$

Remark 4

Markov chain returns to a recurrent state w.p. 1 and does so ∞ *many times. However, unlike a positive recurrent state the mean return time of a null-recurrent state diverges.*

Example 5 (random walk,... again)

1. *For* $p \neq q$*, show that all states are transient.*
2. *For* $p = q = 1/2$*, show that all states are null-recurrent.*

Solution We'll prove this for state 0 and then invoke a theorem later on to extend result to all states (recurrence is a class property!).

Recall gambler's ruin. Player A and player B play a series of games.

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

Let $X_n = A$'s lead over B after n games. $X(0) = 0$, $\mathcal{S} = \mathbb{Z}$. Want to first find f_{00} and see if it is $<$ or $=$ to 1. Note $X_n = 0$ [chain enters state 0 at time n] iff A and B win equal number ($n/2$) of games, where n is even. There are

$$\binom{n}{n/2} \qquad [n \text{ choose } n/2]$$

many ways of this occurring, each with probability $(pq)^{n/2}$.
[A wins $n/2$ times $= p^{n/2}$, B wins $n/2$ times $= q^{n/2}$]
I.e. chain returns to state 0 at n w.p.

$$p_{00}^{(n)} = \binom{n}{n/2} (pq)^{n/2}, \qquad n \text{ even}$$

Recall $p_{ij}^{\sim}(s) = \delta_{ij} + f_{ij}^{\sim}(s) p_{jj}^{\sim}(s)$. Then

$$f_{00}^{\sim}(s) = 1 - \frac{1}{p_{00}^{\sim}(s)} \qquad [i, j = 0]$$

$$f_{00}^{\sim}(s) = 1 - \frac{1}{p_{00}^{\sim}(s)}$$

Now

$$p_{00}^{\sim}(s) := \sum_{n=0}^{\infty} s^n p_{00}^{(n)} = \sum_{n=0}^{\infty} \binom{n}{n/2} \left(\sqrt{pq}s\right)^n$$

Recall:

$$\sum_{m=0}^{\infty} \binom{2m}{m} x^m = \frac{1}{\sqrt{1-4x}}$$

Then

$$\begin{aligned} p_{00}^{\sim}(s) &= \sum_{n=0}^{\infty} \binom{n}{n/2} \left(\sqrt{pq}s\right)^n = \sum_{m=0}^{\infty} \binom{2m}{m} \left(\sqrt{pq}s\right)^{2m} \\ &= \sum_{m=0}^{\infty} \binom{2m}{m} \left(pqs^2\right)^m = \frac{1}{\sqrt{1-4pqs^2}} \end{aligned}$$

[exercise?!][1]

[1] a nice alternative here is to use Stirling's approximation to n!

Hence $f_{00}^{\sim}(s) = 1 - \sqrt{1 - 4pqs^2}$, and the probability of returning is

$$f_{00} = \sum_{n=1}^{\infty} f_{00}(n) = \lim_{s \to 1^-} f_{00}^{\sim}(s) = 1 - \sqrt{1 - 4pq} = \begin{cases} = 1, & p = q \\ < 1, & p \neq q \end{cases}$$

Hence, we have state 0 is transient for $p \neq q$ and recurrent for $p = q$. We now need to show null recurrence for $p = q$.

Set $p = q$ and consider mean return time:

$$\begin{aligned}
\mu_0 &:= \mathbb{E}\big(T_0 \big| X_0 = 0\big) \\
&= \sum_{n=1}^{\infty} n\, f_{00}^{(n)} \\
&= \lim_{s \to 1^-} \frac{\mathrm{d}}{\mathrm{d}s} f_{00}^{\sim}(s) \qquad \Big[f_{00}^{\sim}(s) = \sum_{1}^{\infty} s^n f_{00}^n\Big] \\
&= \lim_{s \to 1^-} \frac{\mathrm{d}}{\mathrm{d}s} \big(1 - \sqrt{1 - s^2}\big) \qquad \text{from previous slide} \\
&= \lim_{s \to 1^-} \frac{s}{\sqrt{1 - s^2}} = \infty
\end{aligned}$$

Therefore, state 0 is null-recurrent for $p = q$.

Remark 6 (random walk summary)

$$\begin{array}{ll}
\textit{transient} \left\{ \begin{array}{ll} p < q & X_n \to -\infty \\ p > q & X_n \to +\infty \end{array} \right. & \\
\textit{null-recurrent} \left\{ \begin{array}{ll} p = q & X \textit{ returns to initial state infinitely often} \\ & \textit{although mean return time is infinite} \end{array} \right. &
\end{array}$$

Theorem 7

A recurrent state j *is null iff* $p_{jj}^{(n)} \to 0, n \to \infty$.

A recurrent state j *is null* $\Rightarrow p_{ij}^{(n)} \to 0, \forall i \in \mathcal{S}$.

I.e., a recurrent state j *is positive (non-null) iff* $p_{jj}^{(n)} \not\to 0$

Proof. See Ergodic theorem, later!

Remark 8

Recall, we now have 3 different classes of states:

- j *transient* $\Leftrightarrow \sum_n p_{jj}^{(n)} < \infty \Rightarrow p_{ij}^{(n)} \to 0, \forall\, i \in \mathcal{S}$
- j *recurrent* $\Leftrightarrow \sum_n p_{jj}^{(n)} = \infty$
  - *null-recurrent:* $p_{ij}^{(n)} \to 0, \forall\, i \in \mathcal{S}$
  - *positive recurrent:* $p_{jj}^{(n)} \not\to 0$

For null-recurrent state j*, the n-step return probs* $p_{jj}^{(n)} \to 0$ *but don't go to zero quick enough for* $\sum_n p_{jj}^{(n)}$ *to converge.*

Sometimes it is only possible to return to a state on an even number of steps (c.f. random walk) or, more generally, on a multiple of d-many steps. This motivates the following state classification

Definition 9

Denote/define the set of all number of steps for which a return to state i is possible by: $D_i := \{n \geq 1 : p_{ii}^{(n)} > 0\}$. Let

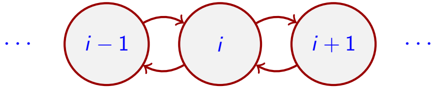
$$d_i := \begin{cases} \gcd D_i\,, & D_i \neq \{\} \\ 0\,, & D_i = \{\} \end{cases} \qquad \textit{greatest common divisor}$$

Then, state i is said to have period d_i and is called

$$\begin{cases} \textit{periodic}, & \text{if } d_i > 1 \\ \textit{aperiodic}, & \text{if } d_i = 1 \end{cases}$$

d_i is is the greatest common divisor of the epochs at which return is possible, i.e. $p_{ii}(n) = 0$ unless n is a multiple of d_i.

Example 10 (random walk: $D_i = \{2, 4, 6, \ldots\}$, $d_i = 2$)



The states all have period 2 and form a single irreducible (later!) class.

- if $p \neq q$ the states are all transients and the walk will eventually drift to $+\infty$ ($p > q$) or $-\infty$ ($p < q$).
- if $p = q = 0.5$ the states are null recurrent so that the walk will return to its initial state infinitely often, although the mean return time between returns will be infinite.

Definition 11

- *State i communicates with state j, written as $i \to j$, if $p_{ij}^{(n)} > 0$ for some $n \geq 0$.*
- *State i and j intercommunicate if $i \to j$ and $j \to i$, and we write $i \leftrightarrow j$.*[a]

[a] caveat: unfortunately, some texts define $i \leftrightarrow j$ if i and j communicate

If two states i and j do not intercommunicate, then either

$$p_{ij}^{(n)} = 0 \qquad \text{for all } n \geq 0$$

or

$$p_{ji}^{(n)} = 0 \qquad \text{for all } n \geq 0$$

or both relations are true.

Lemma 12

The binary relation $\cdot \leftrightarrow \cdot$ *is an* <u>*equivalence relation*</u> *on the state space* $\mathcal{S}$*.*
I.e. $\forall\, i, j, k \in \mathcal{S}$*:*

1. $i \leftrightarrow i$ *[reflexive]*
2. $i \leftrightarrow j \Rightarrow j \leftrightarrow i$ *[symmetry]*
3. $i \leftrightarrow j\,, j \leftrightarrow k \Rightarrow i \leftrightarrow k$ *[transitive]*

Proof

1. $p_{ii}^{(0)} = 1 > 0 \, \forall \, i \in \mathcal{S}$.
2. $i \leftrightarrow j, \Rightarrow \exists m, n \geq 0$ s.t. $p_{ij}^{(m)}, p_{ji}^{(n)} > 0 \Rightarrow j \leftrightarrow i$.
3. $i \leftrightarrow j, j \leftrightarrow k \Rightarrow \exists m, n \geq 0$ s.t. $p_{ij}^{(m)}, p_{jk}^{(n)} > 0$. Chapman-Kolmogorov:

$$\begin{aligned} p_{ik}^{(m+n)} &= \sum_{\ell \in \mathcal{S}} p_{i\ell}^{(m)} \, p_{\ell k}^{(n)} \\ &\geq p_{ij}^{(m)} \, p_{jk}^{(n)} > 0 \end{aligned}$$

I.e., a sum of non-negative terms is greater than or equal to one of its terms. Hence $i \rightarrow k$. By symmetry (repeat with i and k swapped) $k \rightarrow i$ ■

We can now partition the totality of states in equivalence classes:

the states in an equivalence class are those which intercommunicate with each other.

Note: it may be possible, starting in one class, to jump into another class with positive probability, but obviously the chain cannot go back to the original class otherwise the two classes would form just one single class.

A Markov chain is *irreducibile* if the equivalence relation induces only one class, i.e. if all the states intercommunicate.

Example: consider the transition probability matrix

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \vdots & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \vdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \vdots & 0 & 1 & 0 \\ 0 & 0 & \vdots & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \vdots & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{P_1} & 0 \\ 0 & \mathbf{P_2} \end{bmatrix}$$

This Markov chain divides into the 2 classes composed of states $\{1, 2\}$ and $\{3.4, 5\}$.

If X_0 lies in the first class, then the state of the system remain in the first class with transition matrix $\mathbf{P_1}$. Similarly if X_0 lies in the second class.

Basically we have two completely unrelated processes labelled together.

Example: random walk with 2 absorbing states.

$$\mathbf{P} = \begin{array}{|cccccccc|c} & & & & & & & & \text{states} \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & q & 0 & p & \cdots & 0 & 0 & 0 & 2 \\ \vdots & & & & & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & q & 0 & p & a-1 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 1 & a \end{array}$$

We have 2 classes: $\{0\}, \{1, 2, \ldots, a-1\}, \{a\}$. It is possible to reach the first and third from the second class, but it is not possible to return to the second from either the first or the third class.

We now bring together the ideas of classifying the states and that of intercommunicating equivalence classes.

Theorem 13

$i \leftrightarrow j \Rightarrow$

1. $d_i = d_j$
2. - i *transient* $\Leftrightarrow j$ *transient*
   - i *recurrent* $\Leftrightarrow j$ *recurrent*
3. i *null-recurrent* $\Leftrightarrow j$ *null-recurrent*

Sketch Proof

1. Recall $D_i := \{n \geq 1 \colon p_{ii}^{(n)} > 0\}$, $d_i := \gcd D_i$, $D_i \neq \{\}$. Then

$$i \leftrightarrow j \Rightarrow \exists m, n \geq 0 \text{ s.t. } p_{ij}^{(m)}, p_{ji}^{(n)} > 0 \qquad \text{as } i \leftrightarrow j$$

Chapman-Kolmogorov:

$$p_{ii}^{(m+n)} = \sum_{\ell \in \mathcal{S}} p_{i\ell}^{(m)}\, p_{\ell i}^{(n)} \geq p_{ij}^{(m)}\, p_{ji}^{(n)} > 0$$

Hence $\exists \ell \geq 1$ s.t. $p_{ii}^{(\ell)} > 0 \Rightarrow \ell \in D_i$. I.e. $D_i \neq \{\}$.

Let $\ell \in D_i$. Now,

$$\begin{aligned} p_{jj}^{(m+n+\ell)} &= \sum_{r\in\mathcal{S}} \sum_{u\in\mathcal{S}} p_{jr}^{(n)}\, p_{ru}^{(\ell)}\, p_{uj}^{(m)} \\ &\geq p_{ji}^{(n)}\, p_{ii}^{(\ell)}\, p_{ij}^{(m)} > 0, \forall\, \ell \in D_i \end{aligned}$$

Hence $m + n + \ell \in D_j$. Similarly $m + n + 2\ell \in D_j$. This means that $\exists k_1, k_2 \in \mathbb{N}$ s.t.

$$\begin{aligned} m+n+\ell &= k_1 d_j && \text{since } d_j \text{ is the gcd} \\ m+n+2\ell &= k_2 d_j \end{aligned}$$

$\Rightarrow \ell = (k_2 - k_1)\, d_j \Rightarrow \forall\, \ell \in D_i$. we have d_j is a common divisor of D_i. But d_i is greatest common divisor of D_i. Therefore $d_j \leq d_i$. By symmetry (swap i and j and repeat all previous arguments), we also have that $d_i \leq d_j \Rightarrow$ 1.

For ❷ we follow a similar initial step as before:

$$i \leftrightarrow j \Rightarrow \exists m, n \geq 1, \text{ s.t. } p_{ij}^{(m)}, p_{ji}^{(n)} > 0$$

As before, use Chapman-Kolmogorov to show

$$p_{ii}^{(m+n+r)} \geq p_{ij}^{(m)}\, p_{jj}^{(r)}\, p_{ji}^{(n)} \qquad (1)$$

Similarly

$$p_{jj}^{(r)} \geq p_{ji}^{(m)}\, p_{ii}^{(r-n-m)}\, p_{ij}^{(n)} \qquad (2)$$

$$\begin{aligned} (1) \Rightarrow p_{jj}^{(r)} &\leq c_1 p_{ii}^{(r+k)} \qquad \text{set } k = m+n && (3) \\ (2) \Rightarrow p_{jj}^{(r)} &\geq c_2 p_{ii}^{(r-k)} && (4) \end{aligned}$$

But this means that $\sum_r p_{jj}^{(r)}$ and $\sum_r p_{ii}^{(r)}$ converge or diverge together.

For ❸, recall recurrent state i is null iff $p_{ii}^{(n)} \to 0$. In this case, (3) $\Rightarrow p_{jj}^{(n)} \to 0$. Similarly, if $\Rightarrow p_{jj}^{(n)} \to 0$ then (4) $\Rightarrow p_{ii}^{(n)} \to 0$. ■

Remark 14

Intercommunication decomposes the state space into equivalence classes of aperiodic/periodic and transient/(null-or positive)recurrent states, i.e. states in the same class behave similarly.

Example 15

Recall in random walk example we showed that, for state 0*:*

$$\begin{array}{rl} \textit{transient} & p \neq q \\ \textit{null-recurrent} & p = q \end{array}$$

Also recall, for random walk, that $i \leftrightarrow j \, \forall \, i, j \in \mathcal{S}$*. Hence all states are transient for* $p \neq q$*; all states are null-recurrent for* $p = q$*; and we have solved problem Example 5!*

Definition 16

A set $\mathcal{C} \subseteq \mathcal{S}$ *of states is:*

1. *closed* if $p_{ij} = 0, \ \forall\, i \in \mathcal{C}, j \notin \mathcal{C}$
2. *irreducible* if $i \leftrightarrow j \ \forall\, i, j \in \mathcal{C}$

Remark 17

Once a chain takes a value in a closed set of states it never leaves $\mathcal{C}$ *subsequently. A closed set containing one state is called absorbing. By definition, the equivalence classes induced by the* $\cdot \leftrightarrow \cdot$ *operator are irreducible. We can call a closed set of classes* $\mathcal{C}$ *aperiodic (or null etc) if all states in* $\mathcal{C}$ *have this property. Furthermore, if the entire state space is, say, recurrent, then it is irreducible and we can talk of a recurrent, irreducible chain, etc.*

Theorem 18 (decomposition theorem)

State space $\mathcal{S}$ can be uniquely partitioned as

$$\mathcal{S} = \mathcal{T} \bigcup_{\ell} \mathcal{C}_\ell$$

where $\mathcal{T}$ is a set of transient states and $\mathcal{C}_\ell$ are irreducible closed sets of recurrent states.

'Proof' Omitted. Assume $\mathcal{C}_r$ not closed for some r and see p224 G&S. ■

- If $X_0 \in \mathcal{C}_\ell$, then the chain never leaves $\mathcal{C}_\ell$ and we might take $\mathcal{C}_\ell$ to be the whole state space
- If $X_0 \in \mathcal{T}$, then the chain either stays in $\mathcal{T}$ forever or moves eventually to one of the $\mathcal{C}_\ell$ where it subsequently remains.
- When $\mathcal{S}$ is finite, the chain cannot stay in the transient states forever
- For a finite Markov chain, not all states can be transient and there can be no null recurrent states. Thus, if $\mathcal{S}$ is finite and irreducible, it must be positive recurrent.

Example 19

Let $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$ *and*

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$\{1, 2\}$ and $\{5, 6\}$ are irreducible closed sets and therefore contain recurrent non-null states. States 3 and 4 are transient because $3 \to 4 \to 6$ but return from 6 is impossible.

All states have period 1 because $p_{ii}(1) > 0$ for all i.

Lemma 20

If state space is finite then at least one state is recurrent and all recurrent states are positive (non-null).

Proof Let state space $\mathcal{S}$ be finite ($|\mathcal{S}| < \infty$). Assume all states are transient, then $p_{ij}^{(n)} \to 0,\ \forall\, i, j \in \mathcal{S}$. But this means that the finite sum:

$$\lim_{n\to\infty} \sum_{j\in\mathcal{S}} p_{ij}^{(n)} = 0$$

which is a contradiction, since $\sum_{j\in\mathcal{S}} \mathbb{P}\big(X_n = j \big| X_0 = i\big) = 1$. Therefore $\exists$ recurrent class.

Now, assume $\exists$ null recurrent class $\mathcal{C}_0 \neq \{\}$. Then, for $i \in \mathcal{C}_0$:

$$p_{ij}^{(n)} \to 0, \quad j \in \mathcal{C}_0$$

and, since recurrent classes are closed:

$$p_{ij}^{(n)} = 0, \quad j \in \mathcal{S} \setminus \mathcal{C}_0$$

and again we have that the following finite sum leads to a contradiction:

$$\lim_{n\to\infty} \sum_{j \in \mathcal{S}} p_{ij}^{(n)} = 0, \qquad \mathcal{S} = \mathcal{C}_0 \cup \mathcal{S} \setminus \mathcal{C}_0 \qquad \blacksquare$$

Remark 21

A good reason for classifying states and classes of states (as well as entire chains) in this way is that it allows us to talk about what happens to the Markov chain, or at least the probability distribution of the Markov chain, in the long run.

Remark 22 (brief recap on notation)

The probability distribution of the chain at time n *is written as the row vector*

$$\mathbf{p}^{(n)} = \left(p_j^{(n)}\right)_{j\in\mathcal{S}}, \quad p_j^{(n)} = \mathbb{P}(X_n = j)$$

and we had, from Chapman-Kolmogorov that

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

i.e., the distribution at time n *is the initial distribution (vector-matrix) multiplied by the* n*-step transition matrix.*

Definition 23 (invariant distribution)

The row vector $\boldsymbol{\pi} = \left(\pi_j\right)_{j\in\mathcal{S}}$ *is called an* <u>*invariant distribution*</u> *of the chain if*

- $\boldsymbol{\pi}$ *is a probability vector:* $\pi_j \geq 0,\ \forall j \in \mathcal{S},\ \sum_{j\in\mathcal{S}} \pi_j = 1.$
- $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$, *i.e.* $\pi_j = \sum_{i\in\mathcal{S}} \pi_i\, p_{ij}\,,\ \forall j \in \mathcal{S}$

(Unfortunately, an invariant distribution is sometimes called a <u>stationary</u> distribution.)

Remark 24

As the name invariant suggests, $\boldsymbol{\pi}$ *remains unchanged by* $\mathbf{P}$*:*

$$\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}\mathbf{P}^2 = \cdots = \boldsymbol{\pi}\mathbf{P}^n$$

Note that if the initial distribution is invariant: $\mathbf{p}^{(0)} = \boldsymbol{\pi}$, *then*

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

i.e. the distribution of the chain remains the same over all time.

Definition 25

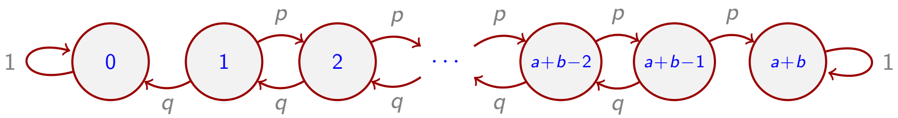
If, for every initial distribution $\mathbf{p}^{(0)}$ *we have that* $\mathbf{p}^{(n)} \to \boldsymbol{\pi}$, *then* $\boldsymbol{\pi}$ *is an equilibrium distribution.*

An invariant distribution need not exist, and if it does exist, it need not be unique. However, for an irreducible chain,

unique invariant distribution exists $\leftrightarrow$ chain is positive recurrent

Before formal result, note following example where $\boldsymbol{\pi}$ is not unique.

Example 26 (gambler's ruin)



Any $\boldsymbol{\pi}$ *of the form* $[\pi_0, 0, 0, \ldots, 0, 0, 1-\pi_0]$, *where* $0 \leq \pi_0 \leq 1$ *is an invariant distribution (exercise: check). I.e.* $\boldsymbol{\pi}$ *is not unique.*
The states $\{1, \ldots, a+b-1\}$ *are transient; the states* 0 *and* $a+b$ *are separate irreducible, positive recurrent, aperiodic classes.*
Note also that, since the chain cannot stay in the finitely many states $\{1, \ldots, a+b\}$ *forever, the limiting distribution* $\lim_{n\to\infty} \mathbf{p}^{(n)}$ *is either*

$$[1, 0, 0, \ldots, 0] \text{ or } [0, 0, 0, \ldots, 1]$$

Remark 27

On the other hand, the urn model is irreducible, positive recurrent (and aperiodic), so a unique invariant distribution exists. (We'll return to that example later on.)

Theorem 28 (p 227, G&S)

An irreducible chain has invariant distribution $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$ *iff all states are positive recurrent. Furthermore,* $\boldsymbol{\pi}$ *is the unique invariant distribution and is given by* $\boldsymbol{\pi} = (\pi_j)_{j \in \mathcal{S}}$ *with* $\pi_j = \mu_j^{-1}, \ \forall\, j \in \mathcal{S}$ *where* μ_j *is the mean return time to state* j*. [Intuitively, if on average, a chain visits state* j *once every* μ_j *time steps, then the prob. that the chain is in state* j *(in the long run) should* $= 1/\mu_j$*]*

proof of Theorem 28 plan

1. show all j transient $\Rightarrow \boldsymbol{\pi} = 0$
2. existence of $\boldsymbol{\pi} = \left(\mu_j^{-1}\right)_j$
3. $\boldsymbol{\pi} = \left(\mu_j^{-1}\right)_j \Rightarrow$ all states are positive recurrent

1: An irreducible chain implies all states are transient or all states are recurrent. Assume all states are transient and that an invariant distribution $\boldsymbol{\pi}$ exists, i.e.

$$\pi_j = \sum_{i \in \mathcal{S}} \pi_i \, p_{ij}^{(n)}$$

But transience implies $p_{ij}^{(n)} \to 0 \; \forall \, i, j \Rightarrow \boldsymbol{\pi} = \mathbf{0}$ which contradicts property that $\boldsymbol{\pi}$ is a probability vector. Hence, all states must be recurrent.

2: [Want $\exists ! \boldsymbol{\pi}$ s.t. $\pi_j = \mu_j^{-1}$] Suppose we define the initial distribution as:

$$\pi_j := \mathbb{P}\big(X_0 = j\big)\ ,\ \forall j \in \mathcal{S}$$

Then, by definition ($T_j =$ first hitting time of j)

$$\begin{aligned} \pi_j\, \mu_j &= \mathbb{P}\big(X_0 = j\big)\, \mathbb{E}\big(T_j \big| X_0 = j\big) \\ &= \sum_{n=1}^{\infty} \mathbb{P}\big(X_0 = j\big)\, \mathbb{P}\big(T_j \geq n \big| X_0 = j\big) \qquad \left[\mathbb{E}Y = \sum_{n=1}^{\infty} \mathbb{P}(Y \geq n)\right] \\ &= \sum_{n=1}^{\infty} \mathbb{P}\big(T_j \geq n, X_0 = j\big) \end{aligned}$$

Since $T_j := \min\{n \geq 1 \colon X_n = j\}$, we have $\mathbb{P}\big(T_j \geq 1\big) = 1$.
We now look at each term in the sum.
The first term on the RHS (i.e. for $n = 1$) is

$$\mathbb{P}\big(T_j \geq 1, X_0 = j\big) = \mathbb{P}\big(X_0 = j\big)$$

For $n = 1 : \mathbb{P}\big(T_j \geq 1, X_0 = j\big) = \mathbb{P}\big(X_0 = j\big)$.
Now, for $n \geq 2$:

$$\begin{aligned}
& \mathbb{P}\big(T_j \geq n, X_0 = j\big) \\
= \;& \mathbb{P}\big(X_0 = j, X_{1:n-1} \neq j\big) \\
= \;& \mathbb{P}\big(X_{1:n-1} \neq j\big) - \mathbb{P}\big(X_0 \neq j, X_{1:n-1} \neq j\big) && [\mathbb{P}(A, B) = \mathbb{P}(B) - \mathbb{P}(\overline{A}, B)] \\
= \;& \mathbb{P}\big(X_{0:n-2} \neq j\big) - \mathbb{P}\big(X_0 \neq j, X_{1:n-1} \neq j\big) && [\text{time-homog.}] \\
=: \;& \alpha_{n-2} - \alpha_{n-1} && [\alpha_n := \mathbb{P}\big(X_{0:n} \neq j\big)]
\end{aligned}$$

I.e.

$$\begin{aligned}
\pi_j \mu_j &= \underbrace{\mathbb{P}\big(X_0 = j\big)}_{n = 1 \text{ term}} + \underbrace{\sum_{n=2}^{\infty} \alpha_{n-2} - \alpha_{n-1}}_{\text{telescopes: } \alpha_0 - \cancel{\alpha_1} + \cancel{\alpha_1} - \cancel{\alpha_2} + \cancel{\alpha_2} - \dots} \\
&= \mathbb{P}\big(X_0 = j\big) + \alpha_0 - \lim_{n \to \infty} \alpha_n
\end{aligned}$$

But, $\lim_{n\to\infty} \alpha_n = \mathbb{P}(X_m \neq j, \forall m) =$ prob. chain never visits j. But X is recurrent. Therefore this prob. is 0 and we have:

$$\pi_j \mu_j = \mathbb{P}\big(X_0 = j\big) + \mathbb{P}\big(X_0 \neq j\big) = 1 \qquad [\alpha_0 := \mathbb{P}\big(X_0 \neq j\big)]$$

❸: [want: $\boldsymbol{\pi} = \left(\mu_j^{-1}\right)_j \Rightarrow$ all states are positive recurrent]

We have shown $\exists! \boldsymbol{\pi}$ s.t. $\pi_j = \mu_j^{-1}$. From Defn. 3, we want to show $\mu_j < \infty$, i.e. that $\pi_j > 0$. Suppose that $\pi_j = 0$ for some j. Then

$$0 = \pi_j = \sum_{k \in \mathcal{S}} \pi_k \, p_{kj}^{(n)} \geq \pi_i p_{ij}^{(n)}, \ \forall \, i \in \mathcal{S}, \ n \geq 0$$

This implies $\pi_i = 0$, $\forall \, p_{ij}^{(n)} > 0$, i.e. $\forall \, i \to j$. But, chain is irreducible, i.e. $i \leftrightarrow j, \forall i, j \in \mathcal{S}$. Therefore

$$\pi_j = 0 \text{ for some } j \Rightarrow \pi_i = 0 \, \forall i \in \mathcal{S} \Rightarrow \sum_{j \in \mathcal{S}} \pi_j = 0 \neq 1$$

which contradicts Defn. 23. Hence $\mu_j < \infty$ and therefore all states are positive recurrent. ■

Theorem 29

For an irreducible, aperiodic Markov chain

$$p_{ij}^{(n)} \to \frac{1}{\mu_j}, \text{ as } n \to \infty \ \forall i, j \in \mathcal{S}$$

Proof omitted. *see, e.g. G&S pp232-235*

Remark 30

If a Markov chain is transient, or null recurrent, then

$$\mu_j = \infty \text{ and } p_{ij}^{(n)} \to 0$$

If a Markov chain is positive recurrent then

$$p_{ij}^{(n)} \to \pi_j = \frac{1}{\mu_j}$$

I.e. (informally) if, on average, a chain visits state j once every μ_j steps then the probability that the chain is in state j (in the long run) should be $1/\mu_j$.

Theorem 31

For an irreducible Markov chain with period d:

$$p_{jj}^{(nd)} \to \frac{d}{\mu_j}\text{, as } n \to \infty \ \forall j \in \mathcal{S}$$

and $p_{jj}^{(m)} = 0$ if d is not a divisor of m.

Theorem 32 (Ergodic theorem [general)

For any aperiodic state j of a Markov chain,

$$p_{jj}^{(n)} \to \frac{1}{\mu_j}\text{, as } n \to \infty$$

and (recall, f_{ij} is prob. of ever going to j from i) for any ther state i

$$p_{ij}^{(n)} \to \frac{f_{ij}}{\mu_j}, i \neq j$$

C.f. p235 G&S