Markov Property in discrete time

Definition

Let $(X_n)_{n=0}^{\infty}$ be a discrete time stochastic process with a discrete state space S. $(X_n)_{n=0}^{\infty}$ is a Discrete-time Markov chain (**DTMC**) if for any $j, i, i_{n-1}, \dots, i_0 \in S$,

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j \mid X_n = i) = p(i,j).$$

$p(i,j)$ are called the (one step) transition probabilities

Note that the transition probabilities are not time (n) dependent. We restrict to this so-called *temporally homogeneous* case.

Transition probabilities

If the state space S is finite (say of cardinality k), the transition probabilities can be organized into a $k \times k$ matrix

$$P = \begin{pmatrix} p(1,1) & \cdots & p(1,k) \\ \vdots & & \vdots \\ p(k,1) & \cdots & p(k,k) \end{pmatrix}$$

with non-negative entries such that the row sums are 1:

$$\sum_{j \in S} p(i,j) = 1.$$

Matrices of this kind are called *stochastic matrices*.
If S is countably infinite you can still think at a matrix with infinite dimensions.

Classification of states, recurrence and transience

Let T_y be the time of the first visit to y, without counting X_0.

$$T_y = \min\{n \geq 1 : X_n = y\}$$

T_y is called **hitting time of** y, and if the chain starts in $X_0 = y$ **return time to** y.
T_y is a random variable expressing how many steps are needed to visit y.
Let

$$\rho_{xy} = P_x(T_y < \infty) = P(\text{we will visit } y | X_0 = x),$$

be the probability of visiting y in a finite time if we start at y.

There are two distinct types of states:

- y is recurrent if $\rho_{yy} = 1$;
- y is transient if $\rho_{yy} < 1$.

Classification of states, recurrence and transience

The names *recurrent* and *transient* are better justified by the following properties:

- Recurrent states are visited infinitely many times. IF WE START THERE!
- on the contrary, the number of visits to any transient state is finite. Therefore you will always find a large enough time such that after that time a transient state is never visited any more.

How to prove it?

We basically used that $P(X_{T_y+n} \in A | X_{T_y} = y) = P(X_n \in A | X_0 = y)$ to conclude that the time between two visits to y are all distributed in the same way.

Regeneration

So, we have

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_1 = j \mid X_0 = i) = p(i,j).$$

Is it true that for a random variable τ

$$P(X_{\tau+1} = j \mid X_\tau = i, X_{\tau-1} = i_{\tau-1}, \ldots, X_0 = i_0) = P(X_1 = j \mid X_0 = i) = p(i,j)?$$

NOT IN GENERAL

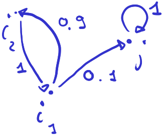
Regeneration

Let τ be defined as the first time n such that $X_{n+1} = j$. In formulas,

$$\tau = \min\{n \geq 0 : X_{n+1} = j\}.$$

Calculate

$$P(X_{\tau+1} = j \mid X_\tau = i) = 1 \neq P(i,j)$$



$$\begin{array}{cccccc} (i_1, & i_2, & i_1, & i_2, & i_1, & j) \\ X_0 & X_1 & X_2 & X_3 & X_4 & X_5 \end{array}$$

$\tau = 4$

(i_1, i_2, i_1)

$(i_1, i_2, i_1, i_2, i_1)$

Stopping times

$\mathcal{F}_i = \sigma(X_u, u \leq i)$

Definition (TEMPI DI ARRESTO)

A *stopping time* for a random process $\{X_i : i \in I\}$ is a random variable τ such that $\{\tau \leq i\} \in \mathcal{F}_i$ for all $i \in I$, where $\{\mathcal{F}_i\}_{i \in I}$ is the natural filtration.

In words, a stopping time is a *good rule for stopping*: by observing the process up to time i (and not the future) I am able to say whether I have stopped ($\tau \leq i$) or not.

Stopping times: examples

Decide which of the following is a stopping time:

1. $T_1 = \min\{n \geq 6 : X_n = 2\}$ ✓
2. $T_2 = \min\{n \geq 1 : X_{n+1} = 2\}$ ✗
3. $T_3 = \min\{n \geq 2 : X_{n-1} = 2\}$ ✓
4. $T_4 = \min\{n > T_1 : X_{n-1} = 2\}$ ✓
5. $T_5 = \min\{n \geq 10 : X_n = X_{n-1}\}$ ✓
6. $T_6 = \min\{n \geq 1 : X_n = X_5\}$ ✗ $\quad \min\{n \geq 7 : X_n = X_5\}$ ✓
7. $T_7 = 10$ ✓

The strong Markov property

If τ is a stopping time, then the strong Markov property holds: for all finite n

$$P(X_{\tau+1} = j \mid X_\tau = i, \tau = n, A) = P(X_1 = j \mid X_0 = i) = p(i,j).$$

for any event A that is $\mathcal{F}_\tau$-measurable ($(\mathcal{F}_n)_{n\in\mathbb{Z}_{\geq 0}}$ being the natural filtration).

$$\mathcal{F}_i = \sigma(X_u : u \leq i)$$

$$P(X_{\tau+1} = j \mid X_\tau = i, \tau = n, A) =$$

$$P(X_{n+1} = j \mid X_n = i, \underbrace{\boxed{\tau = n, A}}_{\in \mathcal{F}_n}) \underset{\text{MARKOV}}{=} P(X_{n+1} = j \mid X_n = i) = p(i,j)$$

Note that $\mathcal{F}_\tau$ is not a random variable! It is a $\sigma-$algebra containing sets of the form

$$A_1 = \{\text{the process visits the state } j \text{ before time } \tau\}$$
$$A_2 = \{\max\{X_i : 1 \leq i \leq \tau\} > K\}$$
$$A_3 = \{\tau \leq 56\}$$
$$A_4 = \{\tau > 5, X_{\tau-3} = j\}$$
$$\ldots$$

$\in \mathcal{F}_\tau$

$\{X_{\tau+1} = 5\} \notin \mathcal{F}_\tau$

Formally,

$$\mathcal{F}_\tau = \{A \in \mathcal{F} \,:\, A \cap \{\tau \leq n\} \in \mathcal{F}_n \quad \forall n\}$$

$\mathcal{F} \supset \mathcal{F}_i$

The strong Markov property

If τ is an almost surely finite stopping time, then

$$P(X_{\tau+1} = j \mid X_\tau = i, A) = P(X_1 = j \mid X_0 = i) = p(i,j).$$

for any event A that is $\mathcal{F}_\tau$-measurable.

$$P(X_{\tau+1} = j \mid X_\tau = i, A) = \sum_{n=0}^{+\infty} P(X_{\tau+1} = j \mid X_\tau = i, \tau = n, A) P(\tau = n \mid X_\tau = i, A)$$

MARKOV FORTE $\rightarrow$ $$= \sum_{n=0}^{+\infty} P(X_1 = j \mid X_0 = i) P(\tau = n \mid X_\tau = i, A)$$

$$= P(X_1 = j \mid X_0 = i) \cdot \underbrace{\boxed{\sum_{n=0}^{+\infty} P(\tau = n \mid X_\tau = i, A)}}_{= 1}$$

Back to recurrence and transience

The time T_y of the first return to y is a stopping time, and the chain starts anew when it hits y. So we can justify $P_y(T^{(1)} < \infty) = \rho_{yy}$ and by induction

$$\boxed{P_y(T_y^{(k)} < \infty) \overset{k}{=} \rho_{yy}}$$

$$P_y(T_y^{(k+1)} < \infty) = \sum_{n=1}^{\infty} P_y(T_y^{(k+1)} < \infty \,|\, T_y^{(k)} = n) P(T_y^{(k)} = n) +$$

$$+ \underbrace{\cancel{P_y(T_y^{(k+1)} < \infty \,|\, T_y^{(k)}}}_{}{}^{0} = \infty) P_y(T_y^{(k)} = \infty)$$

$$= \sum_{n=1}^{\infty} P_y(T_y^{(k+1)} < \infty \,|\, \underline{T_y^{(k)}} = n, \boxed{X_{T_y^{(k)}} = y}) P_y(T_y^{(k)} = n) + 0$$

STOPPING TIME

STRONG MARKOV $\to$

$$= \sum_{n=1}^{\infty} \boxed{P_y(T_y < \infty)} P_y(T_y^{(k)} = n)$$

$$= P_y(T_y < \infty) \boxed{\sum_{n=1}^{\infty} P_y(T_y^{(k)} = n)}$$

$$= \underline{P_y(T_y < \infty)} \boxed{P_y(T_y^{(k)} < \infty)}$$

$$= \underline{\rho_{yy}} \boxed{\rho_{yy}^{k}} = \rho_{yy}^{k+1}.$$

Classification of states, communication classes

EQUIVALENTE: $i \Rightarrow j$ SE o $i=j$ o $P_i(T_j < \infty) > 0$

We say that state j is reachable from i (and write $i \Rightarrow j$) if there exist an $m \geq 0$ such that

$$p^{(m)}(i,j) > 0.$$

$$= \begin{cases} 0 & \text{SE } i \neq j \\ 1 & \text{SE } i=j \end{cases}$$

By definition, $p^{(0)}(i,j) = \delta_{ij}$, where δ_{ij} is the Kronecker symbol, i.e. the (i,j) element of the identity matrix. QUINDI $i \Rightarrow i$ PERCHÉ $p^{(0)}(i,i) > 0$

If both $i \Rightarrow j$ and $j \Rightarrow i$ we say that state i communicates with state j , and write $i \Leftrightarrow j$.

Communication classes, irreducibility

It is easy to prove that communication is an equivalence relation. Both symmetry and reflexivity are obvious. Transitivity is proved in this way.
Since $i \Rightarrow j$, there exist an m such that $p^{(m)}(i,j) > 0$.
Since $j \Rightarrow k$, there exist an n such that $p^{(n)}(j,k) > 0$, therefore

$$p^{(m+n)}(i,k) \overset{\text{c.k.}}{=} \sum_{s \in S} p^{(m)}(i,s)p^{(n)}(s,k) \geq \underbrace{p^{(m)}(i,j)}_{>0}\underbrace{p^{(n)}(j,k)}_{>0} > 0.$$

As any equivalence relation, communication partitions the state space into equivalence classes, also called communication classes.

A Markov chain is is called irreducible, if all the states communicate with each other.

A subset $A \subset S$ of the state space is called absorbing (or closed), if when the chain is started in A, there is no way to get out of A.

$i \in A$, $j \notin A \Rightarrow i \not\Rightarrow j$

Example

Consider the following transition graph (where transition probabilities are not explicitly written, but are understood to be numbers strictly between 0 and 1):

$$\begin{tikzcd}
1 \arrow[d, bend left] & 2 \arrow[l] \arrow[r] \arrow[d] & 4 \arrow[r] & 6 \arrow[dl] \\
5 \arrow[u, bend left] & 3 \arrow[l] \arrow[ur] & 7 \arrow[u] &
\end{tikzcd} \tag{1}$$

Describe the communication classes, and find out which one of them are absorbing.

$\{1,5\}$ $\{4,6,7\}$ RICORRENTI

$\{2,3\}$ TRANSIENTI

Finite state spaces and recurrence

When S is finite, things are nice.

Theorem

Consider a Markov chain with state space S.

1. *If $A \subset S$ is a finite absorbing communication class, all its states are recurrent*
2. *if S is finite, it can be partitioned as $S = T \cup R_1 \cup \cdots \cup R_k$, where T contains all the transient states and the R_i are absorbing communication classes (and therefore they only contain recurrent states)*
3. *if a DTMC on a finite state space is irreducible, it is recurrent (that means that all states are recurrent)*

We will only prove the first statement. The third and second follow immediately as Corollaries.

Finite state spaces and recurrence

In order to prove the first statement we will split the prof into two separate parts

1. $A \subset S$ is a finite absorbing communication class, it has to contain at least a recurrent state;
2. Recurrence is a class property, i.e., if $x \Leftrightarrow y$ and x is recurrent, so is y.

Finite state spaces and recurrence

- $A \subset S$ is a finite absorbing communication class, it has to contain at least a recurrent state.

Indeed if all the states where transient, for each $i \in A$ there would be a time N_i such that after N_i, state i is not visited anymore. Then if I take a time n that is larger than $\boxed{\max_i\{N_i\}}$ at time n the process would not be in A anymore. But that would contradict that A is absorbing.

- Recurrence is a class property, i.e., if $x \Leftrightarrow y$ and x is recurrent, so is y.

For this we first define a useful object.

Recurrence and expected number of visits

The random variable

$$N_n(y) = \sum_{i=1}^{n} \mathbb{1}_{\{X_i = y\}}$$

VISITE A y ENTRO TEMPO n SENZA CONTARE X_0

is the number of visits to y before time n. We define $N(y) = \lim_{n\to\infty} N_n(y) = \boxed{\sum_{i=1}^{\infty} \mathbb{1}_{\{X_i=y\}}}$. We can easily prove that

VISITE A y SENZA CONTARE X_0

$$\mathbb{E}_x[N(y)] = \sum_{n=1}^{\infty} p^{(n)}(x, y)$$

indeed the sequence $N_n(y)$ is a.s. increasing and monotone convergence holds

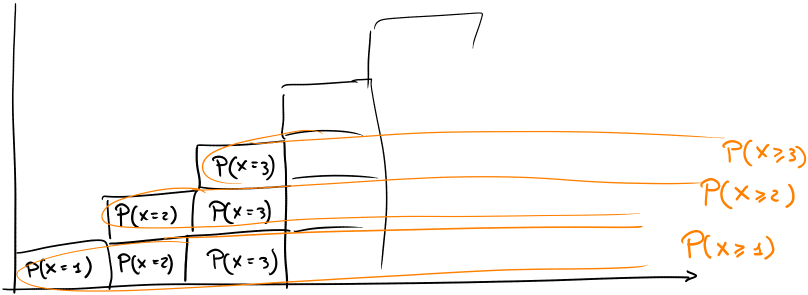
$$\mathbb{E}_x[N(y)] = \sum_{n=1}^{\infty} \mathbb{E}_x[\mathbb{1}_{\{X_n=y\}}] = \sum_{n=1}^{\infty} \overbrace{P_x(X_n = y)}^{n\text{-step t.p.}} = \sum_{n=1}^{\infty} p^{(n)}(x, y)$$

$$= \lim_{n\to+\infty} E[N_n(y)] = \lim_{n\to+\infty} \sum_{i=1}^{n} E[\mathbb{1}\{X_i=y\}]$$

Cute trick

If X is a random variable with values in the non-negative integers, then

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} iP(X = i) = \boxed{\sum_{i=1}^{\infty} P(X \geq i).}$$



Theorem

y is recurrent if and only if $\mathbb{E}_y[N(y)] = \infty$

$\Rightarrow$) OVVIA SE $X_0 = y$ ALLORA $N(y) = \infty$

$\Leftarrow$) NON È OVVIA

$P_x(T_y^{(1)} < \infty) = \rho_{xy}$

In general

$$E_x[N(y)] = \sum_{k=1}^{\infty} P_x(N(y) \geq k) = \sum_{k=1}^{\infty} P_x(T_y^{(k)} < \infty) \overset{\text{SMP}}{=} \sum_{k=1}^{\infty} \rho_{xy}\rho_{yy}^{k-1}$$

$$= \begin{cases} 0 & \text{if } \rho_{xy} = 0 \\ \infty & \text{if } \rho_{xy} > 0 \text{ and } \rho_{yy} = 1 \\ \dfrac{\rho_{xy}}{1-\rho_{yy}} & \text{if } \rho_{xy} > 0 \text{ and } \rho_{yy} < 1 \end{cases}$$

Communication and recurrence

- Recurrence is a class property, i.e., if $x \Leftrightarrow y$ and x is recurrent, so is y.

Since x communicates with y, there exists k and m such that $p^{(\underline{k})}(x, y)$ and $p^{(\underline{m})}(y, x)$ are both positive. Now for any integer n we have

$$p^{(m+n+k)}(y, y) \geq \underline{p^{(m)}(y, x)p^{(n)}(x, x)p^{(k)}(x, y)} \qquad \forall n$$

hence

$$\begin{aligned}
\mathbb{E}_y[N(y)] &= \sum_{r=1}^{\infty} p^{(r)}(y, y) \geq \sum_{r=m+k+1}^{\infty} p^{(r)}(y, y) \\
&\geq \sum_{n=1}^{\infty} p^{(m+n+k)}(y, y) \geq p^{(m)}(y, x)p^{(k)}(x, y) \overbrace{\boxed{\sum_{n=1}^{\infty} p^{(n)}(x, x)}}^{= \infty \text{ x RIC.}} = \infty
\end{aligned}$$

since x is recurrent and $p^{(k)}(x, y)$ and $p^{(m)}(y, x)$ are both positive.

Communication and transience

Transience is also class property, i.e., if $x \Leftrightarrow y$ and x is transient, so is y.

If x is transient and $x \Leftrightarrow y$, if we assumed that y is recurrent we would need to conclude that x is recurrent by what seen before, which is a contradiction.