Universitat de Barcelona

Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Passejada a l'atzar: estudi de les cadenes de Markov homogènies

Sana Amazian

Director: Dr. David Márquez Carreras
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 19 de gener de 2020

Abstract

The aim of this project is to be able to identify, understand and describe the main properties of a time-homogeneous Markov chain. A discrete stochastic process which the future only depends on the present and not the past is called Markov chain. If, in addition, it is satisfied that the probabilities that determine the chain are time independent, then we talk about time-homogeneous Markov chains. We will focus on these ones, introducing the most relevant concepts and related results. The theoretical concepts will be completed with some examples to ease the comprehension.

Resum

El propòsit d'aquest projecte és que el lector pugui identificar, entendre i descriure les principals propietats d'una cadena de Markov homogènia. Un procés estocàstic discret que compleix que el futur depèn únicament del present i no del passat s'anomena cadena de Markov. Si, a més, les probabilitats que determinen la cadena són independents del temps, parlem de cadenes de Markov homogènies. Ens centrarem en aquestes, tot veient-ne els conceptes i resultats relacionats més destacables. S'intentarà acompanyar el desenvolupament teòric amb exemples per a una millor comprensió.

2010 Mathematics Subject Classification. 60G05, 60G07, 60J05, 60J10

Agraïments

Vull agrair principalment al meu tutor, David Márquez, pel suport que m'ha brindat des del primer moment; el guiament a l'hora d'escollir el tema; la seva disponibilitat i paciència durant aquests mesos de treball.

No voldria tancar aquesta etapa sense abans agrair a tothom qui n'ha format part; pel seu recolzament i motivació al llarg d'aquests anys.

Índex

1 Introducció

Els conceptes més bàsics de teoria de les cadenes de Markov van ser introduïts per Andrey Markov l'any 1906, de qui van heretar el nom. Des de llavors, diversos matemàtics han treballat en el desenvolupament de la teoria d'aquest tipus de procés estocàstic. Al llarg del projecte, tractarem únicament amb processos de variable temporal discreta. En aquest cas parlarem de *cadenes* de Markov; així mateix, s'anomena *procés* de Markov quan tractem amb el cas continu.

Les cadenes de Markov són considerades el model matemàtic de fenòmens aleatoris temporals més simple. Tot i així, es troba present en diverses disciplines, com ara biologia, economia, física o enginyeria. A més, aquesta simplicitat fa possible un estudi més detallat del seu comportament.

La principal característica d'aquest tipus de procés aleatori és que *no té memòria.* Amb això entenem que el següent pas (*futur*) depèn exclusivament de l'estat en què ens trobem actualment (*present*), sense importar el camí usat per arribar-hi (*passat*). Dins del conjunt de cadenes de Markov, ens centrarem principalment en les *cadenes de Markov homogènies* que són aquelles que la probabilitat de passar d'un estat a un altre és independent del temps. Aquestes propietats, com veurem, ens permetran formalitzar diferents conceptes i resultats, com per exemple, predir com es comportarà una cadena i calcular les probabilitats que quantifiquen aquest comportament.

L'objectiu d'aquest projecte és fer un estudi de la teoria de les cadenes de Markov homogènies discretes. Així, s'ha fraccionat el treball en diversos blocs.

- Inicialment, presentem els conceptes de teoria de probabilitats necessaris pel desenvolupament d'aquesta nova teoria.
- A continuació, es defineix formalment què és una cadena de Markov (i homogènia), tot veient-ne les principals propietats. Això ens permetrà descriure la coneguda equació de Chapman-Kolmogorov.
- Seguidament, veiem una primera classificació dels estats d'una cadena, donant lloc a noves nocions i propietats, necessàries en els blocs posteriors. Estudiarem, també, la probabilitat i el temps mitjà que tarda la cadena en arribar a un conjunt d'estats determinat.
- Durant el següent bloc, el capítol 5, presentem una segona manera de classificar els estats, aquesta vegada tenint en compte l'evolució temporal de la cadena.
- L'últim bloc, que consta dels capítols 6 i 7, ens proposem estudiar el comportament de la cadena a llarg temps, fent necessària, com veurem, la presentació d'un tipus concret de distribucions.

En cada bloc s'intentarà complementar la part teòrica amb exemples, ja siguin de situacions que ens podríem trobar a la vida quotidiana com exemples purament didàctics.

2 Conceptes bàsics

El principal objectiu d'aquesta secció és presentar els conceptes de teoria de probabilitat més rellevants i necessaris pel desenvolupament de la teoria de les cadenes de Markov.

Definició. Un *espai de probabilitat* és una terna $(\Omega, \mathcal{A}, P)$ on

1. Ω s'anomena *espai mostral* i és un conjunt que correspon als diferents resultats que es poden donar en un experiment: $\omega \in \Omega$.
2. $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ té estructura de σ-*àlgebra*, és a dir, es compleix
   (a) $\Omega \in \mathcal{A}$
   (b) $\mathcal{A}$ és estable per pas a complementari: si $A \in \mathcal{A}$, llavors $A^c \in \mathcal{A}$.
   (c) $\mathcal{A}$ és estable per unions numerables: si $\{A_n, n \geq 1\} \subseteq \mathcal{A}$, llavors $\bigcup\limits_{n\geq 1} A_n \in \mathcal{A}$.

   Aquesta estructura ens permet descriure tots els *esdeveniments* possibles.
3. P s'anomena *probabilitat* i és una aplicació $P : \mathcal{A} \longrightarrow [0,1]$ que satisfà
   (a) $P(\Omega) = 1$
   (b) Si $\{A_n : n \geq 1\}$ és una successió de conjunts d'$\mathcal{A}$ disjunts 2 a 2, llavors
   $$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

Seguidament, se citaran les propietats més importants i útils en el càlcul de probabilitats d'esdeveniments:

1. $P(\varnothing) = 0$.
2. Siguin $A_1, ..., A_n \in \mathcal{A}$ disjunts 2 a 2, se satisfà $P\left(\bigcup\limits_{n=1}^{n} A_n\right) = \sum\limits_{n=1}^{n} P(A_n)$.
3. Per a tot $A \in \mathcal{A}$, es compleix $P(A^c) = 1 - P(A)$.
4. Si $A, B \in \mathcal{A}$ tal que $A \subseteq B$, aleshores $P(A) \leq P(B)$.
5. Per a tot $A \in \mathcal{A}$, tenim que $0 \leq P(A) \leq 1$.
6. Si $A, B \in \mathcal{A}$, llavors $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
7. Siguin $A_1, ..., A_n \in \mathcal{A}$, se satisfà $P\left(\bigcup\limits_{n=1}^{n} A_n\right) \leq \sum\limits_{n=1}^{n} P(A_n)$.
8. Sigui $\{A_n : n \geq 1\} \in \mathcal{A}$, tenim
   (a) Si la successió és creixent, $P\left(\bigcup\limits_{n=1}^{n} A_n\right) = \lim\limits_{n\to\infty} P(A_n)$.
   (b) Si la successió és decreixent, $P\left(\bigcap\limits_{n=1}^{n} A_n\right) = \lim\limits_{n\to\infty} P(A_n)$.

   Se la coneix com *Propietat de continuïtat seqüencial de la probabilitat.*

Probabilitat condicionada i independència

En aquesta secció, es definirà el concepte de probabilitat condicionada, tot donant-ne els resultats més rellevants. També, es presenten les nocions d'independència i independència condicional, estudiant la relació de la segona amb el que es coneix per *família de Markov.*

Definició. Siguin $A, B \in \mathcal{A}$ tal que $P(B) > 0$, la *probabilitat condicionada d'A per B* es defineix com

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)}.$$

$P(A \mid B)$ representa la probabilitat que es doni l'esdeveniment A sabent que ha passat B.

Observació. L'aplicació $P(\,\cdot \mid B) : \mathcal{A} \longrightarrow [0,1]$ és una probabilitat sobre $(\Omega, \mathcal{A})$.

Teorema 2.1 (de les probabilitats totals)**.** *Sigui $\{B_1, \dots, B_n\} \subseteq \mathcal{A}$ una partició de Ω tal que $P(B_i) > 0$, per a tota $i \in \{1, \dots, n\}$. Aleshores, per a qualsevol $A \in \mathcal{A}$,*

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A \mid B_i) P(B_i).$$

Teorema 2.2 (de les probabilitats compostes)**.** *Siguin $\{A_1, \dots, A_n\}$ esdeveniments complint $P(A_1 \cap \dots \cap A_{n-1}) > 0$, llavors*

$$P(A_1 \cap \dots \cap A_n) = P(A_1) P(A_2 \mid A_1) \dots P(A_n \mid A_1 \cap \dots \cap A_{n-1}).$$

Teorema 2.3 (Regla de Bayes)**.** *Siguin $\{A_1, \dots, A_n\}, \{B_1, \dots, B_m\} \subseteq \mathcal{A}$ dues particions de Ω, on cada esdeveniment té probabilitat no nul·la, aleshores*

$$P(A_i \mid B_j) = \frac{P(A_i) P(B_j \mid A_i)}{P(B_j)} = \frac{P(A_i) P(B_j \mid A_i)}{\sum_{k=1}^{n} P(B_j \mid A_k) P(A_k)}.$$

Com s'ha dit, tot seguit s'introduirà el concepte d'independència d'esdeveniments.

Definició. Dos esdeveniments A i B són *independents* si $P(A \cap B) = P(A) \cdot P(B)$.

Observació. Siguin $A, B \in \mathcal{A}$ tal que $P(A), P(B) > 0$.

$$A \text{ i } B \text{ són independents} \iff P(B \mid A) = P(B) \text{ i } P(A \mid B) = P(A).$$

La definició anterior es pot estendre a famílies d'esdeveniments:

Definició. Sigui I un conjunt arbitrari. Un conjunt d'esdeveniments $\{A_i : i \in I\} \subseteq \mathcal{A}$ són *independents* si se satisfà

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \dots P(A_{i_k}).$$

per a tot $\{i_1, \dots, i_k\} \subseteq I$.

Definició. Donats els esdeveniments A, B i C amb $P(C) > 0$, diem que A i B són *condicionalment independents a C* si

$$P(A \cap B \mid C) = P(A \mid C) P(B \mid C).$$

Notem que aquesta noció expressa la independència de A i B respecte $P(\,\cdot \mid C)$.

Definició. Es diu que els esdeveniments A, B i C formen una *família de Markov* si $P(A \cap B) > 0$ i se satisfà

$$P(C \mid A \cap B) = P(C \mid B). \tag{2.1}$$

Observació. Aquests esdeveniments s'han d'entendre com una seqüència cronològica, és a dir, A designa el passat, B el present i C el futur. Així, doncs, la relació descrita a (2.1) es tradueix a dir que només existeix dependència del present i no del passat.

Proposició 2.4. *A, B i C són condicionalment independents si, i només si, formen una família de Markov.*

Variables aleatòries discretes i Esperança

En aquest apartat definim el concepte de variable aleatòria. Més concretament la de tipus discret, que serà de gran importància en el següent capítol. Abans, però, cal introduir la σ-àlgebra de Borel, $\mathcal{B}$, que és aquella generada pels conjunts oberts de $\mathbb{R}$. Considerarem l'espai de probabilitat $(\Omega, \mathcal{A}, P)$.

Definició. Una *variable aleatòria* és una aplicació $X : \Omega \longrightarrow \mathbb{R}$ que satisfà

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$$

per a tot $B \in \mathcal{B}$.

Definició. La *llei* (o *distribució*) d'una variable aleatòria X, P_X, és la probabilitat sobre $(\mathbb{R}, \mathcal{B})$ definida com

$$P_X(B) = P(X^{-1}(B))$$

per a tot $B \in \mathcal{B}$.

Definició. Una variable aleatòria és *discreta* si $X(\Omega)$ és finit o numerable.

Observació. $X(\Omega)$ es pot representar com $\{x_i : i \in I \subseteq \mathbb{N}\}$. Així, doncs, la llei d'una variable aleatòria discreta queda determinada per $P_X(\{x_i\}) = P(X = x_i)$, per a tot $i \in I$.

Definició. Una família de variables aleatòries discretes $X_1, \ldots, X_n$ es diu que és *independent* si per a qualssevol $x_1, \ldots, x_n$, es compleix

$$P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i).$$

Finalment, tanquem el capítol donant la definició d'esperança matemàtica d'una variable aleatòria discreta.

Definició. Diem que una variable aleatòria discreta X té *esperança*, $E(X)$, si, i només si,

$$\sum_{i \in I} |x_i| P(X = x_i) < \infty$$

i en tal cas,

$$E(X) := \sum_{i \in I} |x_i| P(X = x_i).$$

Proposició 2.5. *Per a qualsevol variable discreta X, se satisfà*

$$E(X) = \sum_{k=1}^{\infty} k \cdot P(X = k) = \sum_{k=1}^{\infty} \sum_{r=1}^{k} P(X = k) = \sum_{r=1}^{\infty} \sum_{k=r}^{\infty} P(X = k) = \sum_{r=1}^{\infty} P(X \geq r).$$

3 Cadenes de Markov

El principal objectiu d'aquesta secció és presentar les cadenes de Markov, concretament un tipus específic: les cadenes de Markov homogènies. En veurem els resultats més rellevants, seguit d'una sèrie d'exemples per a una millor comprensió. Finalment, es tancarà la secció introduint el que s'entén per probabilitat de transició en n etapes fins a arribar a descriure l'equació de Chapman-Kolmogorov.

Per tal de poder definir amb rigor la noció de cadena de Markov, necessitem introduir el següent concepte:

Definició. Considerem l'espai de probabilitat $(\Omega, \mathcal{A}, P)$. Un *procés estocàstic* és una família de variables aleatòries indexades

$$\{X_t : t \in T\}$$

on per a cada índex t, $X_t : \Omega \longrightarrow \mathbb{R}$.

A partir d'ara, considerarem només el cas discret, és a dir, quan $T \subseteq \mathbb{Z}^+ = \{0, 1, 2, \ldots\}$. Així, quan parlem de procés estocàstic, ens referirem a $\{X_n : n \geq 0\}$, amb $X_n : \Omega \to I$. On I és un conjunt finit o numerable, que nomenarem *conjunt d'estats* i cada valor $i \in I$ serà un *estat.* Cada variable aleatòria X_n té associada una distribució que, per ara, denotarem per γ^n, és a dir, per a cada $i \in I$,

$$\gamma^n(\{i\}) = \gamma_i^n = P(X_n = i).$$

Definició. Diem que una matriu $\Pi = (p_{ij})_{i,j \in I}$ és una *matriu estocàstica* (o *matriu de transició*) si compleix

(a) $p_{ij} \in [0, 1]$,

(b) Per a tot $i \in I$, $\sum\limits_{j \in I} p_{ij} = 1$.

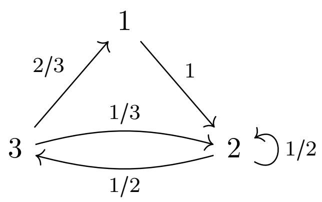
Observacions.

(i) Amb la definició anterior, estem dient que cada fila de la matriu, $(p_{ij})_{j \in I}$, és una probabilitat. A cada valor p_{ij}, se l'anomena *probabilitat de transició.*

(ii) Existeix una correspondència bijectiva entre aquest tipus de matrius i el diagrama que es forma després d'identificar els estats amb vèrtexs i cada probabilitat de transició no nul·la amb arestes.

L'exemple següent servirà per acabar d'entendre la bijecció descrita anteriorment.

$$\begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1/2 & 1/2 \\ 2/3 & 1/3 & 0 \end{pmatrix} \end{array}$$



Si tenim $p_{ij} > 0$, aleshores farem una fletxa de pes p_{ij} que connecti l'estat i amb j. Per exemple, com que $p_{12} = 1$, tindrem una fletxa de pes 1 que connecti l'estat 1 amb l'estat 2. De la mateixa manera, com que $p_{23} = \frac{1}{2} > 0$, tindrem una fletxa sortint de l'estat 2 fins a 3 amb pes $\frac{1}{2}$. I així, amb tots els valors $p_{ij} > 0$.

3.1 Cadenes de Markov homogènies

Ara que ja sabem què és un procés estocàstic, podem procedir a definir cadena de Markov.

Definició. Un procés estocàstic $\{X_n : n \geq 0\}$ que pren valors en un conjunt d'estats I, diem que és una *cadena de Markov* amb distribució inicial $\gamma^0 = \gamma = \{\gamma_i : i \in I\}$ i matriu de transició $\Pi = (p_{ij})_{i,j \in I}$, si per a qualssevol $i_0, \ldots, i_{n-1}, i, j \in I, n \geq 1$ se satisfà

(i) X_0 té llei γ, és a dir, per a $i \in I$, $P(X_0 = i) = \gamma_i$.

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

La segona condició de la definició anterior s'anomena *propietat de Markov* i ens indica que el futur (X_{n+1}) només depèn del present (X_n), i no del passat.

Definició. Una cadena de Markov es diu que és una *cadena de Markov homogènia* si es compleix

$$P(X_{n+1} = j \mid X_n = i) = p_{ij},$$

per a tot $i, j \in I$ i $n \geq 0$.

Aquesta igualtat és coneguda com *propietat d'homogeneïtat* i ens diu que l'evolució del sistema és independent del temps: la probabilitat condicionada de l'estat j en l'instant $n+1$ per l'estat i en l'instant n és independent d'n.

A partir d'ara, ens centrarem únicament en les cadenes de Markov homogènies.

Teorema 3.1.1. *Un procés estocàstic $\{X_n : n \geq 0\}$ que pren valors al conjunt d'estats I és una cadena de Markov homogènia amb distribució inicial γ i matriu de transició Π si, i només si, per a tot $i_0, \ldots, i_n \in I$ i per a tot $n \geq 0$ se satisfà*

$$P(X_0 = i_0, \ldots, X_n = i_n) = \gamma_{i_0} p_{i_0 i_1} \ldots p_{i_{n-1} i_n}. \tag{3.1.1}$$

Demostració.

Mirem, primer, la implicació directa. Suposem que $\{X_n : n \geq 0\}$ és una cadena de Markov homogènia i volem veure que es compleix la igualtat (3.1.1). Tenim

$$\begin{aligned} P(X_0 = i_0, ..., X_n = i_n) &= P(X_0 = i_0)P(X_1 = i_1 \mid X_0 = i_0) \times \ldots \\ &\qquad \ldots \times \; P(X_n = i_n \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}) \\ &= P(X_0 = i_0)P(X_1 = i_1 \mid X_0 = i_0) \times \ldots \times P(X_n = i_n \mid X_{n-1} = i_{n-1}) \\ &= \gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}. \end{aligned}$$

On en la primera igualtat s'ha utilitzat el teorema de les probabilitats compostes, en la segona, la propietat de Markov i finalment, en la tercera, la propietat d'homogeneïtat d'una cadena de Markov homogènia.

Recíprocament, si considerem $n = 0$, tenim que per a tot $i \in I$, $P(X_0 = i) = \gamma_i$, per tant X_0 té llei γ. Així, es compleix la primera condició de la definició de cadena de Markov. Falta per veure, doncs, que se satisfan les propietats de Markov i homogeneïtat.
D'una banda, per la definició de probabilitat condicionada i la igualtat (3.1.1), tenim

$$\begin{aligned} P(X_{n+1} = i_{n+1} \mid X_0 = i_0, \ldots, X_n = i_n) &= \frac{P(X_0 = i_0, ..., X_{n+1} = i_{n+1})}{P(X_0 = i_0, ..., X_n = i_n)} \\ &= \frac{\gamma_{i_0} p_{i_0 i_1} \cdots p_{i_n i_{n+1}}}{\gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}} = p_{i_n i_{n+1}}. \end{aligned}$$

D'altra banda, fent servir el teorema de les probabilitats totals i la igualtat (3.1.1), resulta

$$\begin{aligned}
P(X_{n+1} = i_{n+1} \mid X_n = i_n) &= \frac{P(X_n = i_n, X_{n+1} = i_{n+1})}{P(X_n = i_n)} \\
&= \frac{\displaystyle\sum_{i_0,\ldots,i_{n-1} \in I^{n-1}} P(X_0 = i_0, \ldots, X_n = i_n, X_{n+1} = i_{n+1})}{\displaystyle\sum_{i_0,\ldots,i_{n-1} \in I^{n-1}} P(X_0 = i_0, \ldots, X_n = i_n)} \\
&= \frac{\displaystyle\sum_{i_0,\ldots,i_{n-1} \in I^{n-1}} \gamma_{i_0} p_{i_0 i_1} \cdots p_{i_n i_{n+1}}}{\displaystyle\sum_{i_0,\ldots,i_{n-1} \in I^{n-1}} \gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}} = p_{i_n i_{n+1}}.
\end{aligned}$$

Per tant, hem obtingut

$$P(X_{n+1} = i_{n+1} \mid X_0 = i_0, \ldots, X_n = i_n) = P(X_{n+1} = i_{n+1} \mid X_n = i_n) = p_{i_n i_{n+1}}.$$

□

Notació. Com s'ha pogut observar amb el resultat anterior, una cadena de Markov homogènia queda determinada per la seva distribució inicial i la matriu de transició, per això, aquest tipus de cadenes de Markov es denoten per $CMH(\gamma, \Pi)$.

Proposició 3.1.2. *Sigui $\{X_n : n \geq 0\}$ una $CMH(\gamma, \Pi)$, aleshores $\{X_{n+m} : m \geq 0\}$ és una $CMH(\mathcal{L}(X_m), \Pi)$, on $\mathcal{L}(X_m)$ és la llei de la variable aleatòria discreta X_m.*

Demostració. Fent servir el teorema de les probabilitats totals, i el teorema anterior a $CMH(\gamma, \Pi)$, obtenim:

$$\begin{aligned}
P(X_{0+m} = j_0, \ldots, X_{n+m} = j_n) &= \\
&= \sum_{i_0,\ldots,i_{m-1} \in I^{m-1}} P(X_0 = i_0, \ldots, X_{m-1} = i_{m-1}, X_m = j_0, \ldots, X_{n+m} = j_n) \\
&= \sum_{i_0,\ldots,i_{m-1} \in I^{m-1}} \gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{m-1} j_0} p_{j_0 j_1} \cdots p_{j_{n-1} j_n} \\
&= p_{j_0 j_1} \cdots p_{j_{n-1} j_n} \sum_{i_0,\ldots,i_{m-1} \in I^{m-1}} \gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{m-1} j_0} \\
&= P(X_m = j_0)\; p_{j_0 j_1} \cdots p_{j_{n-1} j_n}.
\end{aligned}$$

□

Per tancar la secció, donarem alguns exemples de cadenes de Markov homogènies.

Exemples.

(i) *Passejada aleatòria sobre $\mathbb{Z}$*

Sigui una partícula que es mou en línia recta fent passos d'una unitat. Un pas és una unitat a la dreta amb probabilitat $p \in (0,1)$, o bé, una unitat a l'esquerra amb probabilitat $q = 1 - p$.

Considerem $X_0 = 0$ la distribució inicial, és a dir, la posició inicial de la partícula. Així, definim la posició a l'etapa n com: $X_n = X_0 + \xi_1 + \ldots + \xi_n$, on $\{\xi_n : n \geq 1\}$ són variables aleatòries independents i idènticament distribuïdes tals que $P(\xi_n = 1) = p$ i $P(\xi_n = -1) = q$, per a qualsevol n. Així, ξ_n designa el desplaçament a l'etapa n.

Passem a veure que $\{X_n : n \geq 0\}$ és una cadena de Markov homogènia. Utilitzant que $\{\xi_n : n \geq 1\}$ són variables aleatòries independents i idènticament distribuïdes, per a tot n i qualssevol $i_0, \ldots, i_{n-1}, i, j \in I$, tenim

$$\begin{aligned}
&P(X_{n+1} = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) = \\
&\qquad = \frac{P(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i, X_{n+1} = j)}{P(X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i)} \\
&\qquad = \frac{P(X_0 = i_0, \xi_1 = i_1 - i_0, \ldots, \xi_n = i - i_{n-1}, \xi_{n+1} = j - i)}{P(X_0 = i_0, \xi_1 = i_1 - i_0, \ldots, \xi_n = i - i_{n-1})} \\
&\qquad = \frac{P(X_0 = i_0, \xi_1 = i_1 - i_0, \ldots, \xi_n = i - i_{n-1}) P(\xi_{n+1} = j - i)}{P(X_0 = i_0, \xi_1 = i_1 - i_0, \ldots, \xi_n = i - i_{n-1})} \\
&\qquad = P(\xi_{n+1} = j - i).
\end{aligned}$$

D'altra banda,

$$\begin{aligned}
P(X_{n+1} = j \mid X_n = i) &= \frac{P(X_n = i, X_{n+1} = j)}{P(X_n = i)} \\
&= \frac{P(X_n = i, \xi_{n+1} = j - i)}{P(X_n = i)} \\
&= \frac{P(X_n = i) P(\xi_{n+1} = j - i)}{P(X_n = i)} = P(\xi_{n+1} = j - i).
\end{aligned}$$

Així, es compleix la propietat de Markov i com que $P(\xi_{n+1} = j - i) := p_{ij}$ no depèn d'n, tenim que $\{X_n : n \geq 0\}$ és una cadena de Markov homogènia.

Ara, podem donar-ne la matriu de transició:

$$\Pi = \begin{pmatrix}
\ddots & & & & & & \iddots \\
 & 0 & p & 0 & 0 & 0 & \\
 & q & 0 & p & 0 & 0 & \\
\cdots & 0 & q & 0 & p & 0 & \cdots \\
 & 0 & 0 & q & 0 & p & \\
 & 0 & 0 & 0 & q & 0 & \\
\iddots & & & & & & \ddots
\end{pmatrix}$$

Notem que el diagrama de la cadena és el següent:

$$\ldots \; -2 \underset{q}{\overset{p}{\rightleftarrows}} -1 \underset{q}{\overset{p}{\rightleftarrows}} \mathbf{0} \underset{q}{\overset{p}{\rightleftarrows}} 1 \underset{q}{\overset{p}{\rightleftarrows}} 2 \; \ldots$$

Cal destacar que en cap moment durant la justificació anterior s'ha utilitzat la llei de ξ_n amb $n \geq 1$. Això dóna lloc al següent resultat:

Proposició 3.1.3. *Sigui $\{Y_n : n \geq 1\}$ una successió de variables aleatòries independents i idènticament distribuïdes tal que $Y_n : \Omega \longrightarrow I$. Definint $X_n = \sum_{i=1}^{n} Y_i$, $\{X_n : n \geq 0\}$ és una CMH.*

I encara podem donar una generalització de la proposició anterior:

Proposició 3.1.4. *Sigui $\{Z_n : n \geq 1\}$ una successió de variables aleatòries independents i idènticament distribuïdes que prenen valors a I. Sigui $f : I \times I \longrightarrow I$ una aplicació i $X_0 : \Omega \longrightarrow I$ una variable aleatòria independent de $\{Z_n : n \geq 1\}$. Aleshores, si definim $X_{n+1} = f(X_n, Z_{n+1})$, amb $n \geq 0$, tenim que $\{X_n : n \geq 0\}$ és una cadena de Markov homogènia.*

Demostració. De manera similar que s'ha fet en l'exemple (i), utilitzarem el fet que $\{Z_n : n \geq 1\}$ són variables aleatòries independents i idènticament distribuïdes i independents de X_0:

$$\begin{aligned}
&P(X_{n+1} = j \mid X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i) = \\
&\qquad = \frac{P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i, X_{n+1} = j)}{P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i)} \\
&\qquad = \frac{P(X_0 = i_0, f(i_0, Z_1) = i_1, \dots, f(i_{n-1}, Z_n) = i, f(i, Z_{n+1}) = j)}{P(X_0 = i_0, f(i_0, z_1) = i_1, \dots, f(i_{n-1}, z_n) = i)} \\
&\qquad = \frac{P(X_0 = i_0, f(i_0, Z_1) = i_1, \dots, f(i_{n-1}, Z_n) = i) P(f(i, Z_{n+1}) = j)}{P(X_0 = i_0, f(i_0, Z_1) = i_1, \dots, f(i_{n-1}, Z_n) = i)} \\
&\qquad = P(f(i, Z_{n+1}) = j).
\end{aligned}$$

D'altra banda,

$$\begin{aligned}
P(X_{n+1} = j \mid X_n = i) &= \frac{P(X_n = i, X_{n+1} = j)}{P(X_n = i)} \\
&= \frac{P(X_n = i, f(i, Z_{n+1}) = j)}{P(X_n = i)} \\
&= \frac{P(X_n = i) P(f(i, Z_{n+1}) = j)}{P(X_n = i)} = P(f(i, Z_{n+1}) = j).
\end{aligned}$$

Així, doncs, se satisfà la propietat de Markov i com que $P(f(X_n, Z_{n+1}) = j) = p_{ij}$ no depèn d'n, tenim que, a més, es compleix la propietat d'homogeneïtat i, per tant, $\{X_n : n \geq 0\}$ és una cadena de Markov homogènia. $\square$

(ii) *Passejada aleatòria sobre $\mathbb{Z}$ amb barreres absorbents*

Siguin dos jugadors A i B jugant a cara i creu amb capital a i b, respectivament. Es diu que el joc ha finalitzat quan un dels dos jugadors es queda sense capital. En cada partida, cada individu aposta 1 unitat del seu capital, on el perdedor ha de cedir la seva moneda al guanyador. El jugador A llença una moneda, de manera que és el guanyador si toca cara i perd si toca creu, amb probabilitats $p \in (0,1)$ i $q = 1 - p$, respectivament.

Entenem $\{X_n : n \geq 0\}$ com el procés que descriu l'evolució del capital del jugador A. Tenim $X_0 = a$ i $X_{n+1} = X_n + \xi_{n+1}$, amb $X_n : \Omega \longrightarrow \{0, 1, \dots, a+b\}$ i ξ_{n+1} com s'ha definit en l'exemple anterior. Aleshores, $\{X_n : n \geq 0\}$ és una $CMH(\gamma, \Pi)$, on:

(a) $\gamma = \delta_{\{a\}}$ es coneix com la *Delta de Dirac* i satisfà $P(X_0 = a) = 1$ i $P(X_0 \neq a) = 0$.

(b) La matriu de transició $(c+1)\times(c+1)$, amb $c=a+b$, és:

$$\Pi = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ q & 0 & p & \dots & 0 & 0 & 0 \\ 0 & q & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & p & 0 \\ 0 & 0 & 0 & \dots & q & 0 & p \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}$$

El joc comença a l'estat a. Si després d'm partides, el jugador A té un capital de $k \notin \{0,c\}$, aleshores, en la pròxima partida, el capital d'A pot passar a $k+1$ amb probabilitat p, o bé, $k-1$ amb probabilitat q. Quan $k \in \{0,c\}$ (el jugador A o el B s'ha arruïnat, respectivament) el joc s'acaba i cap jugador pot canviar de capital, és a dir, $p_{00} = p_{cc} = 1$.

Donem el diagrama de la cadena:

$$1 \circlearrowright 0 \xleftarrow[q]{} 1 \underset{q}{\overset{p}{\rightleftarrows}} 2 \ \dots \ a-1 \underset{q}{\overset{p}{\rightleftarrows}} \boldsymbol{a} \underset{q}{\overset{p}{\rightleftarrows}} a+1 \ \dots \ c-2 \underset{q}{\overset{p}{\rightleftarrows}} c-1 \xrightarrow{p} c \circlearrowleft 1$$

(iii) *Passejada aleatòria sobre $\mathbb{Z}$ amb barreres reflectants*

Ens trobem en la mateixa situació que abans, però ara quan un dels jugadors es queda sense capital, rep 1 unitat per part de l'altre jugador per tal de mantenir el joc. Ara, la matriu de transició és:

$$\Pi = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ q & 0 & p & \dots & 0 & 0 & 0 \\ 0 & q & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & p & 0 \\ 0 & 0 & 0 & \dots & q & 0 & p \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

I el diagrama de la cadena:

$$0 \underset{q}{\overset{1}{\rightleftarrows}} 1 \underset{q}{\overset{p}{\rightleftarrows}} 2 \ \dots \ a-1 \underset{q}{\overset{p}{\rightleftarrows}} \boldsymbol{a} \underset{q}{\overset{p}{\rightleftarrows}} a+1 \ \dots \ c-2 \underset{q}{\overset{p}{\rightleftarrows}} c-1 \underset{1}{\overset{p}{\rightleftarrows}} c$$

3.2 Probabilitat de transició en m etapes: Eq. de Chapman-Kolmogorov

En l'anterior secció hem estudiat les probabilitats de transició p_{ij}, que indiquen el pas d'un estat al següent, ara ens preguntem: quina és la probabilitat que després d'm etapes, la cadena estigui en un estat determinat?

Sigui $\{X_n : n \geq 1\}$ una $CMH(\gamma, \Pi)$, definim

$$p_{ij}^{(m)} := P(X_{n+m} = j \mid X_n = i)$$

per a tot $n, m \geq 0$ i $i, j \in I$. És coneguda com la *probabilitat de transició en m etapes* i indica la probabilitat que partint de l'estat i arribem a l'estat j en exactament m passos. Així mateix, tenim que la *matriu de transició en m etapes* és

$$\Pi_m := (p_{ij}^{(m)})_{i,j\in I}$$

Observacions.

(i) Si $m = 0$, tenim $p_{ij}^{(0)} = P(X_n = j \mid X_n = i)$, que només té sentit quan $i = j$, és a dir, $p_{ij}^{(0)} = \begin{cases} 1 & si \quad i = j \\ 0 & si \quad i \neq j \end{cases}$.
D'aquesta manera, $\Pi_0 = Id$.

(ii) Si $m = 1$, tenim $p_{ij}^{(1)} = P(X_{n+1} = j \mid X_n = i) = p_{ij}$, que és el cas que hem vist en la secció **3.1**.

Proposició 3.2.1. *Sigui $\{X_n : n \geq 1\}$ una $CMH(\gamma, \Pi)$ i Π_m la matriu de transició en m etapes, amb $m \geq 0$. Aleshores, $\Pi_m = \Pi^m = \Pi \cdot \overset{m)}{\cdots} \cdot \Pi$.*

Demostració. És suficient veure que es compleix: $p_{ij}^{(m)} = \sum_{k \in I} p_{ik}^{(m-1)} p_{kj}$ per $m \geq 2$, que es tradueix a dir que $\Pi_m = \Pi_{m-1}\Pi$. D'aquesta manera, iterant successivament, obtindrem $\Pi_m = \Pi_{m-2}\Pi \cdot \Pi = \ldots = \Pi \cdot \overset{m)}{\cdots} \cdot \Pi$. Efectivament, fent servir la definició de probabilitat condicionada, el teorema de les probabilitats totals i que $\{X_n : n \geq 1\}$ és una $CMH(\gamma, \Pi)$, tenim

$$\begin{aligned}
p_{ij}^{(m)} = P(X_{n+m} = j \mid X_n = i) &= \frac{P(X_{n+m} = j, X_n = i)}{P(X_n = i)} \\
&= \frac{\sum_{k \in I} P(X_{n+m} = j, X_{n+m-1} = k, X_n = i)}{P(X_n = i)} \\
&= \sum_{k \in I} \frac{P(X_{n+m} = j, X_{n+m-1} = k, X_n = i)}{P(X_{n+m-1} = k, X_n = i)} \cdot \frac{P(X_{n+m-1} = k, X_n = i)}{P(X_n = i)} \\
&= \sum_{k \in I} P(X_{n+m} = j \mid X_{n+m-1} = k, X_n = i) P(X_{n+m-1} = k \mid X_n = i) \\
&= \sum_{k \in I} p_{ik}^{(m-1)} p_{kj}.
\end{aligned}$$

□

A partir d'aquest resultat, obtenim

1. Π_m és una matriu estocàstica.
2. Sabent que $\Pi_{l+k} = \Pi_l \cdot \Pi_k$, amb $k, l \geq 0$, podem donar el que es coneix per *Equació de Chapman-Kolmogorov*:
$$p_{ij}^{(l+k)} = \sum_{h \in I} p_{ih}^{(l)} p_{hj}^{(k)}.$$
3. Repetint iterativament l'equació de Chapman-Kolmogorov, resulta
$$p_{ij}^{(m)} = \sum_{i_1 \in I} p_{ii_1}^{(m-1)} p_{i_1 j} = \sum_{i_1, \ldots, i_{m-1} \in I} p_{ii_1} p_{i_1 i_2} \cdot \ldots \cdot p_{i_{m-1} j}.$$

Amb el que sabem fins ara, donada una cadena de Markov homogènia $\{X_n : n \geq 0\}$, podem donar-ne la llei de X_n:

Proposició 3.2.2. *Sigui $\{X_n : n \geq 0\}$ una $CMH(\gamma, \Pi)$. Denotem $\gamma^{(n)} = \gamma \Pi_n$. Aleshores, per a tot $k \in I$,*
$$P(X_n = k) = \gamma_k^{(n)}.$$

Demostració.

$$P(X_n = k) = \sum_{h \in I} P(X_n = k \mid X_0 = h) P(X_0 = h) = \sum_{h \in I} P(X_0 = h) p_{hk}^{(n)} = \sum_{h \in I} \gamma_h p_{hk}^{(n)}.$$

□

Com a conseqüència de la proposició anterior, tenim:

Corol·lari 3.2.3. *Sigui $\{X_n : n \geq 0\}$ una $CMH(\gamma, \Pi)$. Es compleix*

(i) $P(X_{l+k} = i) = \sum_{j \in I} P(X_{l+k} = i \mid X_l = j) P(X_l = j) = \sum_{j \in I} \gamma_j^{(l)} p_{ji}^{(k)}$, *és a dir,* $\gamma^{(l+k)} = \gamma^{(l)} \Pi_k$.

(ii) $\gamma_j^{(n)} = P(X_n = j) = \sum_{h \in I} \gamma_h p_{hj}^{(n)} = \sum_{h \in I} \sum_{i_1, \ldots, i_{n-1} \in I} \gamma_h p_{h i_1} \ldots p_{i_{n-1} j}$.

Amb el resultat següent, donem les distribucions del procés en dimensió finita:

Proposició 3.2.4. *La distribució inicial γ i la matriu de probabilitats Π d'una cadena de Markov homogènia $\{X_n : n \geq 0\}$, determinen la llei del vector aleatori $(X_{n_1}, \ldots, X_{n_k})$ amb $0 \leq n_1 < \ldots < n_k$.*

Demostració. Estudiem, primer, el cas $(X_0, \ldots, X_k)$, és a dir, quan hi ha una única etapa entre un estat i el següent. Usant la fórmula de les probabilitats totals, obtenim

$$\begin{aligned} P(X_0 = i_0, ..., X_k = i_k) &= P(X_0 = i_0) P(X_1 = i_1 \mid X_0 = i_0) \times \ldots \\ &\qquad \ldots \times \ P(X_k = i_k \mid X_0 = i_0, \ldots, X_{k-1} = i_{k-1}) \\ &= \gamma_{i_0} p_{i_0 i_1} \ldots p_{i_{k-1} i_k}. \end{aligned}$$

Ara, de manera més general, estudiem el cas $(X_{n_1}, \ldots, X_{n_k})$ amb $n_1 < \ldots < n_k$, és a dir quan es pot donar que hi hagi més d'una etapa entre estats.

$$\begin{aligned} &P(X_{n_1} = i_1, ..., X_{n_k} = i_k) = \\ &= P(X_{n_1} = i_1) P(X_{n_2} = i_2 \mid X_{n_1} = i_1) \times \ldots \times P(X_{n_k} = i_k \mid X_{n_1} = i_1, \ldots, X_{n_{k-1}} = i_{k-1}) \\ &= \sum_{h \in I} \gamma_h p_{h i_1}^{(n_1)} p_{i_1 i_2}^{(n_2 - n_1)} \ldots p_{i_{k-1} i_k}^{(n_k - n_{k-1})}. \end{aligned}$$

□

Per concloure el capítol, donem dos exemples senzills per veure'n el càlcul de $p_{ij}^{(n)}$.

(i) Considerem la cadena que donats $\alpha, \beta > 0$, té per matriu de transició i diagrama:

$$\Pi = \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix}$$

$$\begin{tikzcd} 1 \arrow[loop left, "1-\alpha"] \arrow[r, bend left, "\alpha"] & 2 \arrow[loop right, "1-\beta"] \arrow[l, bend left, "\beta"] \end{tikzcd}$$

Com que per $n \geq 1$ es compleix: $\Pi_n = \Pi_{n-1} \Pi$, tenim

$$p_{11}^{(n)} = p_{11}^{(n-1)} p_{11} + p_{12}^{(n-1)} p_{21} = (1 - \alpha) p_{11}^{(n-1)} + \beta p_{12}^{(n-1)}.$$

A més, com que Π_{n-1} és una matriu estocàstica, es compleix $p_{11}^{(n-1)} + p_{12}^{(n-1)} = 1$. Així, queda

$$p_{11}^{(n)} = (1-\alpha-\beta)p_{11}^{(n-1)} + \beta.$$

Per tant, obtenim la relació de recurrència següent:

$$\begin{cases} p_{11}^{(n)} = (1-\alpha-\beta)p_{11}^{(n-1)} + \beta, \\ p_{11}^{(0)} = 1. \end{cases}$$

Iterant, obtenim la següent expressió:

$$p_{11}^{(n)} = p_{11}^{(0)}(1-\alpha-\beta)^n + \beta\sum_{k=0}^{n}(1-\alpha-\beta)^k.$$

Finalment, resolent la sèrie geomètrica, resulta

$$p_{11}^{(n)} = \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n.$$

Observem que si considerem valors d'n prou petits, com que Π té dimensió 2, podem trobar l'expressió de $p_{ij}^{(n)}$ calculant directament $\Pi_n = \Pi^n$. Considerant $n = 2$, tenim

$$\Pi_2 = \begin{pmatrix} \alpha^2 - 2\alpha + \alpha\beta + 1 & -\alpha^2 + 2\alpha - \alpha\beta \\ -\beta^2 - \alpha\beta + 2\beta & \beta^2 + \alpha\beta - 2\beta + 1 \end{pmatrix}$$

I efectivament,

$$p_{11}^{(2)} = \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^2 = \alpha^2 - 2\alpha + \alpha\beta + 1.$$

(ii) *Mutació d'un virus*

Suposem que un virus pot existir en N varietats diferents. En cada generació, pot mantenir-se en la mateixa varietat, o bé, pot mutar-ne a qualsevol altra d'aquestes amb probabilitat $\alpha \in (0,1)$. Ens preguntem, quina és la probabilitat que la varietat de la n-èssima generació sigui la mateixa que l'original?

Es tracta d'una cadena de Markov homogènia amb estats $\{1,\dots,N\}$, així Π és una matriu $N \times N$, complint

$$p_{ij} = \begin{cases} 1-\alpha, & i = j \\ \frac{\alpha}{N-1}, & i \neq j \end{cases}$$

Aleshores, el problema es tradueix en saber quant val $p_{11}^{(n)}$. Cal tenir en compte que en algun moment, el virus passarà de l'estat inicial a un dels restants amb probabilitat α, i que també, en un altre instant, es donarà la transició d'un dels altres estats a l'inicial. Per tant, com que els $N-1$ estats restants tenen un comportament simètric, podem entendre'ls com un únic i estudiar la cadena de només 2 estats següent:

$$\begin{tikzcd} \textit{inicial} \arrow[loop left, "1-\alpha"] \arrow[r, bend left, "\alpha"] & \textit{restants} \arrow[l, bend left, "\frac{\alpha}{N-1}"] \arrow[loop right, "1-\frac{\alpha}{N-1}"] \end{tikzcd}$$

Observem que és la mateixa cadena estudiada en l'exemple (i) amb $\beta = \frac{\alpha}{N-1}$, per tant, obtenim

$$p_{11}^{(n)} = \frac{\frac{\alpha}{N-1}}{\alpha + \frac{\alpha}{N-1}} + \frac{\alpha}{\alpha + \frac{\alpha}{N-1}}\left(1 - \alpha - \frac{\alpha}{N-1}\right)^n = \frac{1}{N} + \frac{N-1}{N}\left(1 - \frac{\alpha N}{N-1}\right)^n.$$

4 Estructura de classes

En aquest capítol, veurem una manera de classificar els diferents estats d'una cadena de Markov, podent així determinar la irreductibilitat de la cadena. Seguidament, analitzarem la periodicitat d'aquestes classes, que ens permetrà descriure el que es coneixen per subclasses cícliques. Finalment, es presenten les nocions de probabilitat i temps mig d'absorció, estudiant, així, l'evolució temporal de la cadena. En cada cas, es donaran diferents exemples per acabar d'entendre els conceptes donats.

Com fins ara, considerem que $\{X_n : n \geq 0\}$ és una $CMH(\gamma, \Pi)$, on cada variable aleatòria discreta X_n pren valors a I finit o numerable i $\Pi = (p_{ij})_{i,j \in I}$.

4.1 Comunicació entre estats

De vegades, es possible *fraccionar* una cadena de Markov en peces més petites, de manera que es pot estudiar el comportament de cada una d'aquestes per a una millor comprensió del total. En aquesta secció farem aquest estudi.

Definició. Un estat $j \in I$ és *accessible* des d'un l'estat $i \in I$ si existeix $n \geq 0$ tal que

$$p_{ij}^{(n)} = P(X_n = j \mid X_0 = i) > 0.$$

En aquest cas, escrivim $i \longrightarrow j$.

Definició. Diem que dos estats $i, j \in I$ es *comuniquen* i ho denotarem per $i \longleftrightarrow j$, si se satisfà $i \longrightarrow j$ i $j \longrightarrow i$.

Observacions

1. Per a qualsevol $n \geq 1$, es compleix $p_{ij}^{(n)} = \sum_{i_1,\ldots,i_{n-1} \in I} p_{ii_1} p_{i_1 i_2} \ldots p_{i_{n-1} j}$. Així, doncs, $p_{ij}^{(n)} > 0$ si, i només si, existeix almenys un camí $i = i_0, i_1, \ldots, i_{n-1}, i_n = j$ des de i fins a j, de manera que $p_{ii_1} p_{i_1 i_2} \ldots p_{i_{n-1} j} > 0$.

2. La relació $\longleftrightarrow$ és d'equivalència al conjunt d'estats I:

   - *Reflexiva*: Tenim $p_{ii}^{(0)} = 1 > 0$, per tant $i \longleftrightarrow i$, per a qualsevol $i \in I$.
   - *Simètrica*: Si $i \longleftrightarrow j$, existeixen $n, m \geq 0$ tal que $p_{ij}^{(n)} > 0$ i $p_{ji}^{(m)} > 0$. Per tant, $j \longleftrightarrow i$.
   - *Transitiva*: Suposem que $i \longleftrightarrow j$ i $j \longleftrightarrow k$. Sabem, doncs, que existeixen $n, m \geq 0$ tal que $p_{ij}^{(n)} > 0$ i $p_{jk}^{(m)} > 0$. A més, fent ús de l'equació de Chapman-Kolmogorov obtenim

     $$p_{ik}^{(n+m)} = \sum_{h \in I} p_{ih}^{(n)} p_{hk}^{(m)} \geq p_{ij}^{(n)} p_{jk}^{(m)} > 0.$$

     Així, $i \longrightarrow k$. Anàlogament, obtindrem $k \longrightarrow i$. Per tant, $i \longleftrightarrow k$.

3. La relació anterior genera una partició disjunta d'I en classes d'equivalència: els elements d'una classe seran els estats que es comuniquen entre ells. Així, és possible començar en una classe i entrar-ne a una altra amb probabilitat positiva, però no es pot tornar a la inicial, perquè implicaria que dos estats de classes diferents es comuniquen.

Definició. Una classe C és *tancada*, si donat $i \in C$ tal que $i \longrightarrow j$, aleshores $j \in C$. En altres paraules, C és una classe de la qual no se'n pot sortir. A més, se satisfà $\sum_{j \in C} p_{ij} = 1$.

Definició. Un estat $i \in I$ es diu que és *absorbent* (o *tancat*) si $\{i\}$ és una classe tancada. Tenim, doncs, que es compleix $p_{ii} = 1$.

Definició. Una cadena és *irreductible* si hi ha una única classe d'equivalència, és a dir, que tots els estats es comuniquen.

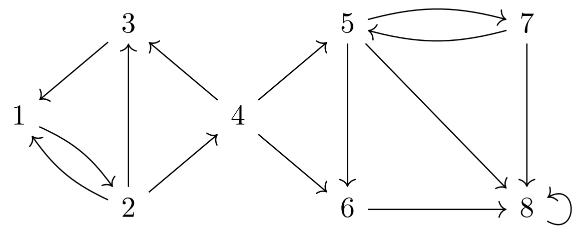
Definició. Un estat $i \in I$ es diu que és *de pas* (o *no essencial*), si existeixen $m \geq 1$ i $j \in I$ tal que $p_{ij}^{(m)} > 0$, però $p_{ji}^{(n)} = 0$, per a tot $n \geq 1$.

Amb això estem dient que, a partir d'un cert instant, si ens trobem a l'estat i, és possible sortir-ne, però no tornar-hi mai més. Per contra, entenem per estats *essencials* als restants, és a dir, aquells dels quals se'n pot sortir, però eventualment tornar-hi, incloent en aquest conjunt els estats absorbents.

A continuació, donarem uns quants exemples per acabar d'entendre aquestes definicions:

Exemples. Quan s'estudien les diferents classes d'una cadena, el valor de les probabilitats de transició deixen de tenir importància: només interessa identificar les que no són nul·les. Així, als diagrames només s'il·lustraran les fletxes corresponents a les $p_{ij} > 0$.

(i) Considerem la cadena que té associat el següent diagrama:



- Els estats 2, 3 i 4 són accessibles des de 1, i alhora, podem accedir a 1 des de qualsevol d'aquests tres estats. De la mateixa manera, tenim que 2 comunica amb 3 i 4, i que 3 comunica amb 4. Així, obtenim que aquests estats pertanyen a la mateixa classe.
- Partint de l'estat 5 (resp. 7), l'únic estat del qual podem tornar-hi és el 7 (resp. 5). Per tant, l'estat 5 només es comunica amb el 7, i viceversa.
- L'estat 6 no es comunica amb cap altre estat a part de si mateix, $p_{66}^{(0)} = 1$, per tant $\{6\}$ és una classe. Notem que no és tancada, ja que $p_{66} = 0$.
- Una vegada arribem a l'estat 8 ja no podem sortir-ne, és a dir, 8 és un estat absorbent i, per tant, $\{8\}$ una classe tancada.

Per tant la cadena té 4 classes: $\{1,2,3,4\}$, $\{5,7\}$, $\{6\}$ i $\{8\}$, i no és irreductible.

Observem que discutir quins estats són essencials o de pas, pot reduir-se a fer aquest estudi per classes, ja que el raonament usat per a un estat de la classe serà aplicable a un altre pel fet de comunicar-se.

- Tenim que $p_{45} > 0$, però des de 5 no podrem tornar mai a 4, per tant 4 és un estat de pas, i conseqüentment, els estats 1, 2 i 3, també.

- L'estat 8 és accessible des de 5, però com que 8 és absorbent no es pot retornar a 5. Per tant, 5 i 7 també són estats de pas.
- Com que $p_{68} = 1 > 0$ i 8 és absorbent, tenim que 6 és un estat de pas.
- L'únic estat essencial de la cadena és el 8 per ser absorbent.

(ii) *Passejada aleatòria sobre* $\mathbb{Z}$

$$\ldots\ -2 \rightleftarrows -1 \rightleftarrows \mathbf{0} \rightleftarrows 1 \rightleftarrows 2\ \ldots$$

Tots els estats es comuniquen entre ells, per tant tots ells són essencials. La cadena és irreductible, ja que tenim una única classe formada per a tot $i \in \mathbb{Z}$.

(iii) *Passejada aleatòria sobre* $\mathbb{Z}$ *amb barreres absorbents*

$$\circlearrowleft 0 \longleftarrow 1 \rightleftarrows 2\ \ldots\ldots\ N-2 \rightleftarrows N-1 \longrightarrow N \circlearrowright$$

Clarament, els estats 0 i N són absorbents, així que són estats essencials i cada un forma una classe tancada. Pel que fa a la resta d'estats, tenim que donat un estat $i \in \{1,\ldots,N-1\}$ podem accedir a qualsevol altre $j \in \{1,\ldots,N-1\}$, i viceversa. Tenim, doncs, que $C = \{1,\ldots,N-1\}, \{0\}$ i $\{N\}$ són les classes de la cadena i, per tant, no és irreductible. Notem, a més, que $p_{10} > 0$, però $p_{01}^{(n)} = 0$ per a qualsevol $n \geq 1$, per tant els estats de la classe C són de pas.

4.2 Periodicitat i classes cícliques

En algunes ocasions, ens interessarà estudiar el comportament periòdic d'una cadena de Markov. Per exemple, considerem la passejada aleatòria sobre $\mathbb{Z}$ amb $p \in (0,1)$. Sigui C_0 el conjunt dels estats parells i C_1 els senars, tenim que $I = C_0 \cup C_1$ i necessàriament en cada etapa passem d'un conjunt a l'altre. Així, després de dues etapes tornarem a estar al conjunt inicial. En aquesta secció estudiarem aquest fenomen de manera més rigorosa.

Definició. Sigui $i \in I$ un estat essencial. El *període* d'aquest estat es defineix com

$$d_i = mcd\{n \geq 1 : p_{ii}^{(n)} > 0\}.$$

Diem que un estat $i \in I$ és *aperiòdic* quan $d_i = 1$.

Proposició 4.2.1. *Si dos estats* $i, j \in I$ *es comuniquen, aleshores* $d_i = d_j$.

Demostració.

Si $i = j$ és evident. Suposem que $i \neq j$, així el fet que es comuniquin implica que existeixen $m, n \geq 1$ tal que $p_{ij}^{(m)} > 0$ i $p_{ji}^{(n)} > 0$. Si considerem $k \geq 1$ tal que $p_{jj}^{(k)} > 0$, de l'equació de Chapman-Kolmogorov, en traiem

$$p_{ii}^{(m+k+n)} \geq p_{ij}^{(m)} p_{jj}^{(k)} p_{ji}^{(n)} > 0,$$
$$p_{ii}^{(m+n)} \geq p_{ij}^{(m)} p_{ji}^{(n)} > 0.$$

Per definició de d_i, obtenim que $d_i \mid m+k+n$ i $d_i \mid m+n$, per tant d_i divideix la diferència, és a dir: $d_i \mid (m+k+n) - (m+n) = k$. Com que havíem considerat qualsevol k tal que $p_{jj}^{(k)} > 0$, tenim que $d_i \mid d_j$. Anàlogament, intercanviant i per j, obtenim $d_j \mid d_i$. Per tant, $d_i = d_j$. □

Aquest resultat ens permet parlar de període d d'una classe (respecte la relació $\longleftrightarrow$), així direm que una classe és *periòdica* si $d > 1$, o bé, *aperiòdica* si $d = 1$.

Sigui d el període d'una classe C i siguin $i, j \in C$. Siguin r, s i t nombres naturals tal que $p_{ij}^{(r)} > 0$, $p_{ij}^{(s)} > 0$ i $p_{ji}^{(t)} > 0$, amb $r \neq s$. Tenim,

$$p_{ii}^{(r+t)} \geq p_{ij}^{(r)} p_{ji}^{(t)} > 0,$$
$$p_{ii}^{(s+t)} \geq p_{ij}^{(s)} p_{ji}^{(t)} > 0.$$

D'aquesta manera, com que $d \mid r + t$ i $d \mid s + t$, tenim $d \mid r - s$, és a dir, $r - s = \dot{d}$. Amb això estem dient que si $r = ad + b$, amb a i b enters tal que $0 \leq b \leq d - 1$, aleshores existeix un enter c tal que $s = cd + b$. Conseqüentment, si j és accessible des de i en n passos, se satisfà $n \equiv b \pmod d$ on $0 \leq b \leq d - 1$ depèn de i i j, però és independent d'n. En altres paraules, qualsevol camí que uneixi i amb j amb probabilitat no nul·la, tindrà longitud b mòdul d. D'aquesta manera, cada parella $i, j \in C$ determina un element $\overline{b} \in \mathbb{Z}/d$.

Dit això, fixat $i \in C$, podem definir

$$\begin{aligned} C_0 &= \{j \in C : p_{ij}^{(n)} > 0 \text{ implica } n \equiv 0 \pmod d\} \\ C_1 &= \{j \in C : p_{ij}^{(n)} > 0 \text{ implica } n \equiv 1 \pmod d\} \\ &\vdots \\ C_{d-1} &= \{j \in C : p_{ij}^{(n)} > 0 \text{ implica } n \equiv d - 1 \pmod d\} \end{aligned}$$

Els conjunts $C_0, \ldots, C_{d-1}$ s'anomenen *subclasses cícliques* i compleixen $C = \bigcup_{k=0}^{d-1} C_k$. Considerem $C_d = C_0$, per tant $C_u = C_v$ si $u \equiv v \pmod d$.

Proposició 4.2.2. *Sigui C una classe de període d. Si $j \in C_b$ i $p_{jk} > 0$, aleshores $k \in C_{b+1}$.*

Demostració. Sigui $n \geq 1$ tal que $p_{ij}^{(n)} > 0$. Tenim que es compleix

$$p_{ik}^{(n+1)} \geq p_{ij}^{(n)} p_{jk} > 0.$$

Per tant, com que $n \equiv b \pmod d$, tenim que $n + 1 \equiv b + 1 \pmod d$, i així $k \in C_{b+1}$. □

Amb el resultat anterior, diem que si sortim de $i \in C_0$, la cadena es trasllada d'una subclasse C_b a C_{b+1} en una etapa.

Exemples.

(i) *Passejada aleatòria sobre $\mathbb{Z}$ amb $X_0 = 0$*

La cadena té una única classe essencial $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$ i té període 2. Efectivament, partint de l'estat 0, la cadena hi podrà tornar al cap de $2, 4, 6, \ldots$ etapes. Així, $d_0 = 2$. A més, com que hem provat que el període és una propietat de classe, es té $d = 2$. Tenim, doncs, dues subclasses cícliques:

$$C_0 = \{j \in C : p_{0j}^{(n)} > 0 \text{ implica } n \equiv 0 \pmod 2\} = \{0, \pm 2, \pm 4, \ldots\},$$
$$C_1 = \{j \in C : p_{0j}^{(n)} > 0 \text{ implica } n \equiv 1 \pmod 2\} = \{\pm 1, \pm 3, \ldots\}.$$

Notem que la descripció de les subclasses cícliques depèn de l'estat de referència que agafem. En aquest cas, hem considerat $i = 0$, que coincidiran per a qualsevol altre estat parell. Si considerem, però, $i = 1$, tindrem

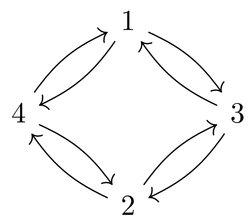
$$C_0 = \{j \in C : p_{1j}^{(n)} > 0 \text{ implica } n \equiv 0 \pmod 2\} = \{\pm 1, \pm 3, \ldots\},$$
$$C_1 = \{j \in C : p_{1j}^{(n)} > 0 \text{ implica } n \equiv 1 \pmod 2\} = \{0, \pm 2, \pm 4, \ldots\}.$$

(ii) *Passejada aleatòria sobre $\mathbb{Z}$ amb barreres absorbents*

Les diferents classes d'aquesta cadena són: $\{0\}, \{N\}$ i $\{1, \ldots, N-1\}$. La definició de període només té sentit per estats essencials i ja hem vist que només els estats 0 i N ho són, per tant només cal estudiar la periodicitat d'aquests. Prenent l'estat 0, com que $p_{00} = 1$, tenim que per a qualsevol $n \geq 1$, $p_{00}^{(n)} = 1 > 0$. Així, $d_0 = 1$. El mateix raonament és vàlid per la classe tancada $\{N\}$, per tant és una classe aperiòdica.

(iii) Considerem la cadena de Markov homogènia que té per matriu de transició i diagrama associat següents:

$$\begin{pmatrix} 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{pmatrix}$$



Tots els estats es comuniquen, per tant la cadena és irreductible, tenint com a única classe $\{1, 2, 3, 4\}$. Estudiant el període de l'estat 1, obtenim que podem tornar-hi després d'un nombre parell d'etapes, per tant $d_1 = 2$, i conseqüentment, $d = 2$. Les subclasses cícliques per a aquest estat són:

$$C_0 = \{j \in \{1, 2, 3, 4\} : p_{1j}^{(n)} > 0 \text{ implica } n = \dot{2}\} = \{1, 2\},$$
$$C_1 = \{j \in \{1, 2, 3, 4\} : p_{1j}^{(n)} > 0 \text{ implica } n = \dot{2} + 1\} = \{3, 4\}.$$

4.3 Temps d'entrada i probabilitats d'absorció

En aquesta secció estudiarem quan una cadena de Markov entra a un cert subconjunt del conjunt d'estats I i el temps mig que tarda en arribar-hi.

Definició. Sigui $\{X_n : n \geq 0\}$ una cadena de Markov homogènia amb conjunt d'estats I finit o numerable i matriu de transició Π. El *temps d'entrada* a un conjunt $A \subset I$ és una variable aleatòria $H_A : \Omega \longrightarrow \mathbb{N} \cup \{\infty\}$ definida com

$$H_A(\omega) = \inf\{n \geq 0 : X_n(\omega) \in A\},$$

de manera que assumirem $\inf\{\varnothing\} = \infty$. La probabilitat que partint de l'estat i, la cadena entri a A és

$$\lambda_i^A = P(H_A < \infty \mid X_0 = i).$$

Quan A és una classe tancada, diem que λ_i^A és la *probabilitat d'absorció*.

Observacions.

- Si A és una classe tancada i $i \in A$, aleshores $\lambda_i^A = 1$.
- Si B és una classe tancada tal que $A \cap B = \varnothing$ i $i \in B$, aleshores $\lambda_i^A = 0$.

Definició. El temps mitjà que necessita $\{X_n : n \geq 0\}$ per arribar a A s'anomena *temps mig d'absorció* i es defineix de la manera següent:

$$m_i^A = E(H_A \mid X_0 = i) = \sum_{n=1}^{\infty} nP(H_A = n \mid X_0 = i) + \infty P(H_A = \infty \mid X_0 = i).$$

Cal destacar que l'expressió $\infty P(H_A = \infty \mid X_0 = i)$ és purament notacional. S'utilitza per denotar que val 0 quan $P(H_A = \infty \mid X_0 = i) = 0$, i ∞ altrament.

A partir d'ara considerarem $A \subset I$ una classe tancada, $\lambda^A = \{\lambda_i^A : i \in I\}$ les probabilitats d'absorció i $m^A = \{m_i^A : i \in I\}$ els temps mig d'absorció. Estudiarem els resultats generals de com calcular-los.

Teorema 4.3.1. *La probabilitat d'absorció λ^A és la solució minimal no negativa del sistema d'equacions lineal*

$$\begin{cases} \lambda_i^A = 1, & i \in A \\ \lambda_i^A = \sum_{j \in I} p_{ij} \lambda_j^A, & i \notin A \end{cases}$$

Per minimal *entenem que si $x = \{x_i \geq 0 : i \in I\}$ és una altra solució del sistema, aleshores $x_i \geq \lambda_i^A$, per a tota $i \in I$.*

Demostració. Primer, cal provar que λ_A satisfà el sistema d'equacions anterior. Suposem $X_0 = i \in A$, aleshores $H_A = 0$, que implica $\lambda_i^A = 1$. Suposem, ara, que $X_0 = i \notin A$, aleshores $H_A \geq 1$. Aplicant la propietat de Markov, obtenim

$$P(H_A < \infty \mid X_0 = i, X_1 = j) = P(H_A < \infty \mid X_1 = j) = \lambda_j^A.$$

Així,

$$\begin{aligned} \lambda_i^A = P(H_A < \infty \mid X_0 = i) &= \frac{P(H_A < \infty, X_0 = i)}{P(X_0 = i)} = \sum_{j \in I} \frac{P(H_A < \infty, X_0 = i, X_1 = j)}{P(X_0 = i)} \\ &= \sum_{j \in I} \frac{P(H_A < \infty, X_0 = i, X_1 = j)}{P(X_0 = i, X_1 = j)} \cdot \frac{P(X_0 = i, X_1 = j)}{P(X_0 = i)} \\ &= \sum_{j \in I} P(H_A < \infty \mid X_0 = i, X_1 = j) P(X_1 = j \mid X_0 = i) = \sum_{j \in I} p_{ij} \lambda_j^A. \end{aligned}$$

Ara que hem vist que efectivament λ^A és solució, passem a veure que és la minimal. Sigui $x = \{x_i \geq 0 : i \in I\}$ una altra solució del sistema. Se satisfà $\lambda_i^A = x_i = 1$, per a qualsevol $i \in A$. Suposem, ara, $i \notin A$, aleshores

$$x_i = \sum_{j \in I} p_{ij} x_j = \sum_{j \in A} p_{ij} + \sum_{j \notin A} p_{ij} x_j.$$

Observem que $\sum_{j\in A} p_{ij}$ denota la probabilitat d'arribar a A des de i en una única etapa. Substituint x_j obtenim

$$\begin{aligned}
x_i &= \sum_{j\in A} p_{ij} + \sum_{j\notin A} p_{ij} \left(\sum_{k\in A} p_{jk} + \sum_{k\notin A} p_{jk} x_k \right) \\
&= P(X_1 \in A \mid X_0 = i) + P(X_1 \notin A, X_2 \in A \mid X_0 = i) + \sum_{j\notin A} \sum_{k\notin A} p_{ij} p_{jk} x_k.
\end{aligned}$$

Fent la substitució anterior repetidament n vegades, obtenim l'expressió següent

$$\begin{aligned}
x_i &= P(X_1 \in A \mid X_0 = i) + \ldots + P(X_1 \notin A, \ldots, X_{n-1} \notin A, X_n \in A \mid X_0 = i) + \\
&\qquad + \sum_{j_1\notin A} \cdots \sum_{j_n\notin A} p_{ij_1} p_{j_1 j_2} \ldots p_{j_{n-1} j_n} x_{j_n} = \\
&= P(H_A = 1 \mid X_0 = i) + \ldots + P(H_A = n \mid X_0 = i) + \sum_{j_1\notin A} \cdots \sum_{j_n\notin A} p_{ij_1} p_{j_1 j_2} \ldots p_{j_{n-1} j_n} x_{j_n} \\
&= P(H_A \leq n \mid X_0 = i) + \sum_{j_1\notin A} \cdots \sum_{j_n\notin A} p_{ij_1} p_{j_1 j_2} \ldots p_{j_{n-1} j_n} x_{j_n}.
\end{aligned}$$

Com que $x_{j_n} \geq 0$, tenim $x_i \geq P(H_A \leq n \mid X_0 = i)$, per a tot n. Així,

$$x_i \geq \lim_{n\to\infty} P(H_A \leq n \mid X_0 = i) = P(H_A < \infty \mid X_0 = i) = \lambda_i^A.$$

□

Passem a veure el resultat anàleg aplicat als temps mig d'absorció.

Teorema 4.3.2. *El temps mig d'absorció m^A és la solució minimal no negativa del sistema d'equacions lineal*

$$\begin{cases}
m_i^A = 0, & i \in A \\
m_i^A = 1 + \sum_{j\notin A} p_{ij} m_j^A, & i \notin A
\end{cases}$$

Demostració. Primer, cal provar que m_A satisfà el sistema d'equacions anterior. Suposem $X_0 = i \in A$, aleshores $H_A = 0$, que implica $m_i^A = 0$, obtenint així la primera expressió del sistema. Assumim, ara, que $X_0 = i \notin A$. Poden donar-se dos casos en funció del valor de $P(H_A = \infty \mid X_0 = i)$:

Suposem que $P(H_A = \infty \mid X_0 = i) > 0$, aleshores $m_i^A = \infty$. Si $i \notin A$ i $p_{ij} > 0$, tenim $m_j^A = \infty$, per tant se satisfà la igualtat $m_i^A = 1 + \sum_{j\notin A} p_{ij} m_j^A = \infty$.

Considerem, ara, el cas que $P(H_A = \infty \mid X_0 = i) = 0$. D'aquesta manera, tenim

$$P(H_A < \infty \mid X_0 = i) = 1 = \sum_{n\geq 1} P(H_A = n \mid X_0 = i).$$

Aplicant-ho a la definició de m_i^A, ens queda

$$\begin{aligned}
m_i^A &= \sum_{n\geq 1} n P(H_A = n \mid X_0 = i) = 1 + \sum_{n\geq 1} (n-1) P(H_A = n \mid X_0 = i) \\
&= 1 + \sum_{n\geq 1} (n-1) \frac{P(H_A = n, X_0 = i)}{P(X_0 = i)} = 1 + \sum_{j\in I} \sum_{n\geq 1} (n-1) \frac{P(H_A = n, X_1 = j, X_0 = i)}{P(X_0 = i)}
\end{aligned}$$

Ara, usant la definició de probabilitat condicionada i la propietat de Markov tenim

$$\begin{aligned} m_i^A &= 1 + \sum_{j \in I} \sum_{n \geq 1} (n-1) \frac{P(H_A = n, X_0 = i, X_1 = j)}{P(X_0 = i, X_1 = j)} \cdot \frac{P(X_0 = i, X_1 = j)}{P(X_0 = i)} \\ &= 1 + \sum_{j \in I} \sum_{n \geq 1} (n-1) P(H_A = n \mid X_1 = j) \cdot p_{ij} = 1 + \sum_{j \in I} m_j^A p_{ij} = 1 + \sum_{j \notin A} m_j^A p_{ij}. \end{aligned}$$

Ara només falta veure que m^A és minimal. Sigui $y = \{y_i \geq 0 : i \in I\}$ una altra solució del sistema. Se satisfà $m_i^A = y_i = 0$, per a qualsevol $i \in A$. Suposem, ara, $i \notin A$, aleshores

$$\begin{aligned} y_i &= 1 + \sum_{j \notin A} p_{ij} y_j = 1 + \sum_{j \notin A} p_{ij} \left(1 + \sum_{k \notin A} p_{jk} y_k \right) \\ &= P(H_A \geq 1 \mid X_0 = i) + P(H_A \geq 2 \mid X_0 = i) + \sum_{j \notin A} \sum_{k \notin A} p_{ij} p_{jk} y_k. \end{aligned}$$

Repetint el procediment anterior n vegades, amb $n \geq 1$, obtenim

$$y_i = P(H_A \geq 1 \mid X_0 = i) + \ldots + P(H_A \geq n \mid X_0 = i) + \sum_{j_1 \notin A} \cdots \sum_{j_n \notin A} p_{j_1 j_2} \ldots p_{j_{n-1} j_n} y_{j_n}.$$

Com que $y_{j_n} \geq 0$, tenim $y_i \geq P(H_A \geq 1 \mid X_0 = i) + \ldots + P(H_A \geq n \mid X_0 = i)$, per tant si fem $n \to \infty$, queda

$$y_i \geq \sum_{n=1}^{\infty} P(H_A \geq n \mid X_0 = i) = E(H_A \mid X_0 = i) = m_i^A.$$

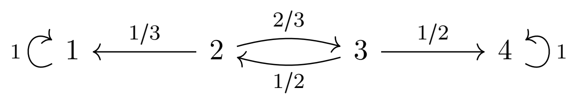
□

Passem a veure alguns exemples per aplicar els resultats estudiats.

Exemples.

(i) Considerem la cadena de Markov que té per matriu de transició i diagrama associat següents:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$



Ens proposem respondre les preguntes:

- Partint de l'estat 2, quina es la probabilitat d'absorció de l'estat 4?
- Quantes etapes, de mitjana, són necessàries per tal que la cadena sigui absorbida pels estats 1 o 4, si l'estat inicial és 2?

Notem que té sentit parlar de probabilitat d'absorció de 4, ja que és un estat absorbent, i per tant una classe tancada. Escriurem $\lambda = \lambda^{\{4\}}$.
Per resoldre el primer problema, ens cal calcular λ_2. El sistema lineal d'equacions descrit al Teorema 4.3.1 per λ és el següent:

$$\begin{cases} \lambda_2 = \frac{1}{3}\lambda_1 + \frac{2}{3}\lambda_3 \\ \lambda_3 = \frac{1}{2}\lambda_2 + \frac{1}{2}\lambda_4 \\ \lambda_4 = 1 \end{cases}$$

Així, substituint l'expressió de λ_3 a λ_2 queda: $\lambda_2 = \frac{1}{2}\lambda_1 + \frac{1}{2}$.

El sistema no ens ha donat informació sobre λ_1, però com que sabem que λ és la solució no negativa minimal, necessàriament $\lambda_1 = 0$. D'aquesta manera, obtenim $\lambda_2 = \frac{1}{2}$. Acabem així, responent la primera pregunta. Passem a respondre la segona.

Escrivint $m = m^{\{1,4\}}$, el segon problema es tradueix a calcular el valor de m_2. Considerem el sistema d'equacions lineal descrit al Teorema 4.3.2 aplicat a m:

$$\begin{cases} m_1 = 0 \\ m_2 = 1 + \frac{2}{3}m_3 \\ m_3 = 1 + \frac{1}{2}m_2 \\ m_4 = 0 \end{cases}$$

Tenim $m_2 = 1 + \frac{2}{3}\left(1 + \frac{1}{2}m_2\right)$, per tant $m_2 = \frac{5}{2}$.

(ii) Un jugador és a un casino amb un capital inicial de i euros. Aposta una unitat en cada etapa, de manera que guanya un altre euro amb probabilitat $p \in (0,1)$, i perd l'aposta amb probabilitat $q = 1 - p$. Suposant que el capital del casino és il·limitat, es tracta d'una variació de l'exemple *Passejada aleatòria sobre* $\mathbb{Z}$ *amb barreres absorbents*, tenint com a única barrera l'estat 0. Estudiem la cadena que determina l'evolució del capital que té per diagrama associat el següent:

$$\begin{tikzcd} 0 \arrow[loop left, "1"] & 1 \arrow[l, "q"] \arrow[r, bend left, "p"] & 2 \arrow[l, bend left, "q"] & \dots & i-1 \arrow[r, bend left, "p"] & \boldsymbol{i} \arrow[l, bend left, "q"] \arrow[r, bend left, "p"] & i+1 \arrow[l, bend left, "q"] & \dots \end{tikzcd}$$

Volem calcular la probabilitat que el jugador s'arruïni, és a dir, calcular la probabilitat d'absorció de l'estat 0, $\lambda_i^{\{0\}}$. Per això, usarem el Teorema 4.3.1 per $\lambda = \lambda^{\{0\}}$:

$$\begin{cases} \lambda_0 = 1 \\ \lambda_i = p\lambda_{i+1} + q\lambda_{i-1}, \qquad i = 1,2,3\dots \end{cases}$$

La solució general d'una recurrència del tipus $ax_{n+1} + bx_n + cx_{n-1} = 0$ és

$$x_n = \begin{cases} A\alpha^n + B\beta^n, & \alpha \neq \beta \\ (A + nB)\alpha^n, & \alpha = \beta \end{cases}$$

amb α i β arrels de $ay^2 + by + c = 0$ i A i B constants que venen donades per les condicions inicials de la recurrència. Aplicant-ho al nostre cas, obtenim

- Si $p \neq q$, tenim com a solució

$$\lambda_i = A + B\left(\frac{q}{p}\right)^i.$$

  Estudiem els dos casos:

  - Si $p < q$, aleshores $\frac{q}{p} > 1$, però $\lambda_i \leq 1$, per tant $B = 0$, i conseqüentment, $\lambda_i = A = 1$, per a qualsevol $i \in I$. Això significa que independentment del capital inicial, el jugador acabarà arruïnat.

- Si $p > q$: Sabem $\lambda_0 = 1 = A + B$, per tant substituint $B = 1 - A$ a l'expressió de la solució queda

$$\lambda_i = A + (1 - A)\left(\frac{q}{p}\right)^i = A\left(1 - \left(\frac{q}{p}\right)^i\right) + \left(\frac{q}{p}\right)^i$$

  Per tal que λ sigui solució no negativa, necessàriament cal que $A \geq 0$, però com que, a més, és minimal, tenim que $A = 0$. Així, $\lambda_i = \left(\frac{q}{p}\right)^i$.

- Si $p = q = \frac{1}{2}$, tenim com a solució

$$\lambda_i = A + Bi.$$

Sabem $\lambda_0 = A + 0 \cdot B = A = 1$. A més, com que $0 \leq \lambda_i \leq 1$, cal que $B = 0$. Per tant, $\lambda_i = 1$, per a qualsevol $i \in I$. D'aquesta manera, tot i que es tracti d'un joc just, el jugador també acabarà arruïnat.

5 Estats recurrents i transitoris

Aquest capítol està dividit en tres parts: en la primera presentem una nova manera de classificar els estats d'una cadena de Markov; estats recurrents i transitoris. Definim, entre d'altres, conceptes com el primer instant en que la cadena visita un cert estat o la probabilitat de retorn a l'estat inicial. Seguidament, en la segona part, veurem com determinar si un estat és recurrent o transitori mitjançant la probabilitat de retorn i els fets que en resulten. Finalment, la tercera i última part del capítol, consisteix en una sèrie d'exemples per acabar de completar l'estudi.

Considerem l'espai de probabilitat $(\Omega, \mathcal{A}, P)$. Sigui $\{X_n : n \geq 0\}$ una $CMH(\gamma, \Pi)$, on cada variable aleatòria discreta X_n pren valors a I finit o numerable i $\Pi = (p_{ij})_{i,j\in I}$.

5.1 Introducció

L'objectiu principal d'aquesta secció és definir els conceptes i presentar els resultats tècnics més rellevants necessaris per l'estudi que es farà en les seccions següents.

Definició. Un estat $i \in I$ és *recurrent* si

$$P(X_n = i \text{ per infinits } n \mid X_0 = i) = 1.$$

Alternativament, diem que és *transitori* si

$$P(X_n = i \text{ per infinits } n \mid X_0 = i) = 0.$$

En altres paraules, sempre s'acaba retornant a un estat recurrent, però a partir d'un cert moment, ja no es torna més a un estat transitori. Notem, doncs, que un estat absorbent és recurrent.

Un dels propòsits del capítol és provar que la definició anterior parteix el conjunt d'estats en dos de disjunts, és a dir, que un estat és transitori, o bé, recurrent. Com que aquests conceptes tracten amb la cadena a llarg termini, és necessari estudiar-ne l'evolució temporal. D'aquesta manera, començarem definint diferents conceptes i demostrant petits resultats fins a tenir totes les eines necessàries per complir el nostre objectiu.

Entenem per *primer instant* en que la cadena visita l'estat $i \in I$ com la variable aleatòria definida per

$$T_i(\omega) = \inf\{n \geq 1 : X_n(\omega) = i\},$$

de manera que $\omega \in \Omega$ i $\inf\{\varnothing\} = \infty$. Inductivament, podem definir el *r-èssim instant* en que l'estat i és visitat per la cadena com

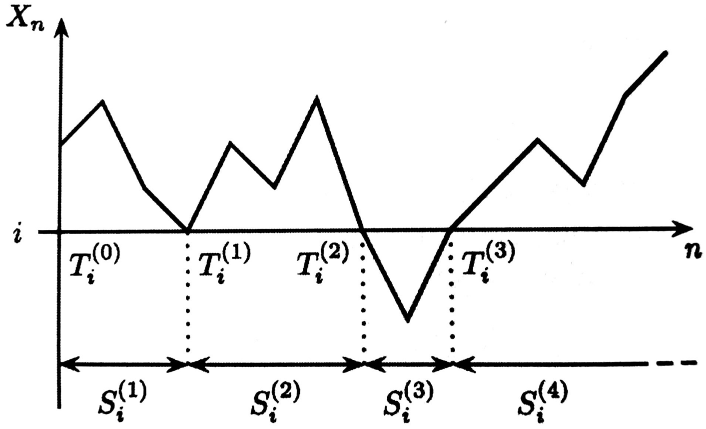
$$\begin{aligned}
T_i^{(0)}(\omega) &= 0,\\
T_i^{(1)}(\omega) &= T_i(\omega),\\
T_i^{(r)}(\omega) &= \inf\{n \geq T_i^{(r-1)} + 1 : X_n(\omega) = i\}, r \geq 2.
\end{aligned}$$

Aleshores, es pot definir el *temps de la r-èssima excursió* com

$$S_i^{(r)} = \begin{cases} T_i^{(r)} - T_i^{(r-1)}, & T_i^{(r-1)} < \infty. \\ 0, & \text{altrament.} \end{cases}$$

Notem que si $T_i^{(r-1)} < \infty$, podem escriure $S_i^{(r)} = \inf\{n \geq 1 : X_{T_i^{(r-1)}+n} = i\}$.

La següent gràfica, que podrem trobar també a [1], ajuda a acabar assimilar aquestes definicions de manera més visual.



Observació. Notem que a diferència del temps d'entrada a i, $H^{\{i\}}$, el primer instant que la cadena visita i, T_i, **no** considera l'estat inicial com a visita. És a dir, si la cadena parteix de l'estat i, T_i serà la següent etapa que la cadena es trobi a i.

A continuació, ens proposem descriure la propietat forta de Markov. Abans, però, cal definir el concepte d'instant d'aturada d'un procés estocàstic.

Definició. Una variable aleatòria $T : \Omega \longrightarrow \{1, 2, 3, \ldots\} \cup \{\infty\}$ diem que és un *instant d'aturada* d'un procés estocàstic $\{X_n : n \geq 0\}$ si, per a cada $n \geq 0$, l'esdeveniment $\{T = n\}$ només depèn de $X_0, \ldots, X_n$.

Observació. Les variables aleatòries $T_i^{(r)}$ i $S_i^{(r)}$ són instants d'aturada. En efecte, $\left\{T_i^{(r)} = n\right\}$ significa que $X_n(\omega) = i$, i que abans d'arribar a l'etapa n, la cadena ha passat per l'estat i exactament $r - 1$ vegades més. Per tant, $\left\{T_i^{(r)} = n\right\}$ depèn només de $X_0, \ldots, X_n$. Conseqüentment, $S_i^{(r)}$ és un instant d'aturada.

En la secció **3.1** descrivíem per primera vegada la propietat de Markov, dient que per a cada instant m, condicionalment a $X_m = i$, el procés després d'aquest instant, comença de nou des de l'estat i. La propietat forta de Markov ens mostra que la propietat anterior es manté per instants d'aturada.

Teorema 5.1.1 (Propietat forta de Markov)**.** *Sigui $\{X_n : n \geq 0\}$ una $CMH(\gamma, \Pi)$ i T un instant d'aturada de la cadena. Aleshores, condicionalment a $T < \infty$ i $X_T = i$, es té que $\{X_{T+n} : n \geq 0\}$ és una $CMH(\delta_i, \Pi)$ i independent de $X_0, X_1, \ldots, X_T$, tal que $\delta_i = (\delta_{ij})_{j \in I}$ complint $\delta_{ij} = 1$, si $i = j$, i $\delta_{ij} = 0$, si $i \neq j$.*

Demostració. Sigui B un esdeveniment determinat per $X_0, X_1, \ldots, X_T$. Aleshores, es té que $B \cap \{T = m\}$ dependrà només de $X_0, X_1, \ldots, X_m$. Així, tenim

$$\begin{aligned}
&P(\{X_T = j_0, X_{T+1} = j_1, \ldots, X_{T+n} = j_n\} \cap B \cap \{T = m\} \cap \{X_T = i\}) \\
&\quad = P(\{X_m = j_0, X_{m+1} = j_1, \ldots, X_{m+n} = j_n\} \cap B \cap \{T = m\} \cap \{X_m = i\}) \\
&\quad = P(X_m = j_0, \ldots, X_{m+n} = j_n \mid B \cap \{T = m\} \cap \{X_m = i\}) \cdot P(B \cap \{T = m\} \cap \{X_m = i\}).
\end{aligned}$$

On en la última igualtat hem fet servir la definició de probabilitat condicionada. Ara, com que l'expressió del primer factor del producte està en termes de X_m, podem aplicar-hi la propietat de Markov a l'instant m, i posteriorment, la propietat d'homogeneïtat. D'aquesta manera, ens queda

$$\begin{aligned} P(\{X_T = j_0, X_{T+1} = j_1, \dots, X_{T+n} = j_n\} \cap B \cap \{T = m\} \cap \{X_T = i\}) \\ = P(X_m = j_0, \dots, X_{m+n} = j_n \mid X_m = i) \cdot P(B \cap \{T = m\} \cap \{X_m = i\}) \\ = P(X_0 = j_0, \dots, X_n = j_n \mid X_0 = i) \cdot P(B \cap \{T = m\} \cap \{X_m = i\}). \end{aligned}$$

Seguidament, sumarem des de $m = 0$ fins a infinit als dos cantons de la igualtat, obtenint

$$\begin{aligned} P(\{X_T = j_0, X_{T+1} = j_1, \dots, X_{T+n} = j_n\} \cap B \cap \{T < \infty\} \cap \{X_T = i\}) \\ = P(X_0 = j_0, \dots, X_n = j_n \mid X_0 = i) \cdot P(B \cap \{T < \infty\} \cap \{X_T = i\}). \end{aligned}$$

Finalment, dividint l'expressió per $P(\{T < \infty\} \cap \{X_T = i\})$ queda

$$\begin{aligned} P(\{X_T = j_0, X_{T+1} = j_1, \dots, X_{T+n} = j_n\} \cap B \mid T < \infty, X_T = i) \\ = P(X_0 = j_0, \dots, X_n = j_n \mid X_0 = i) \cdot P(B \mid T < \infty, X_T = i). \end{aligned}$$

D'aquest resultat en traiem que $\{X_{T+n} : n \geq 0\}$ és una cadena de Markov homogènia amb matriu de transició Π i que és independent de $X_0, X_1, \dots, X_T$. □

Passem a enunciar un lema que utilitza els conceptes que hem definit i que usarem en demostracions més endavant. Per a demostrar-lo, farem servir la propietat forta de Markov que acabem de descriure.

Lema 5.1.2. *Per cada $r \geq 2$, condicionalment a $T_i^{(r-1)} < \infty$, $S_i^{(r)}$ és independent de $\{X_m : m \leq T_i^{(r-1)}\}$ i se satisfà*

$$P(S_i^{(r)} = n \mid T_i^{(r-1)} < \infty) = P(T_i = n \mid X_0 = i).$$

Demostració.

És evident que si $T = T_i^{(r-1)} < \infty$, aleshores $X_T = i$. Ara, apliquem la propietat forta de Markov a l'instant d'aturada T i obtenim que, condicionalment a $T < \infty$, $\{X_{T+n} : n \geq 0\}$ és una $CMH(\delta_i, \Pi)$ i independent de $X_0, X_1, \dots, X_T$. D'aquesta manera, com que $S_i^{(r)}$ només depèn de X_{T+n} amb $n \geq 1$, tenim que $S_i^{(r)}$ és independent de $X_0, X_1, \dots, X_T$.

A més, tenim

$$S_i^{(r)} = \inf\{n \geq 1 : X_{T+n} = i\}.$$

Per tant, estem dient que $S_i^{(r)}$ és el primer instant en que el procés $\{X_{T+n} : n \geq 0\}$ es troba a l'estat i, és a dir,

$$P(S_i^{(r)} = n \mid T_i^{(r-1)} < \infty) = P(T_i = n \mid X_0 = i).$$

□

Una altra noció imprescindible per a demostrar el que ens havíem proposat al principi del capítol és la de *probabilitat de retorn* a un estat.

Definició. Partint de l'estat i, la probabilitat que la cadena visiti l'estat j és

$$\rho_{ij} = P(T_j < \infty \mid X_0 = i).$$

Així, podem definir la *probabilitat de retorn* a l'estat i com $\rho_{ii} = P(T_i < \infty \mid X_0 = i)$.

Un concepte que va molt lligat amb la definició anterior és el següent:

Definició. El *nombre de visites*, N_i, que fa la cadena a l'estat i es defineix com

$$N_i(\omega) = \#\{n \geq 1 : X_n = i\} = \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n = i\}}.$$

Corol·lari 5.1.3. *A partir de la definició anterior obtenim*

$$\begin{aligned} E(N_i \mid X_0 = j) &= E\left(\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=i\}} \mid X_0 = j\right) = \sum_{n=1}^{\infty} E\left(\mathbb{1}_{\{X_n=i\}} \mid X_0 = j\right) = \sum_{n=1}^{\infty} P(X_n = i \mid X_0 = j) \\ &= \sum_{n=1}^{\infty} p_{ji}^{(n)}. \end{aligned}$$

Per tant, $E(N_i \mid X_0 = i) = \sum_{n=1}^{\infty} p_{ii}^{(n)}$.

La manera de relacionar els últims dos conceptes definits és a partir del lema següent.

Lema 5.1.4. *Per* $r \geq 1$*, la probabilitat que partint de l'estat* i *la cadena passi per un estat* j *almenys* r *vegades és*

$$P(N_j \geq r \mid X_0 = i) = \rho_{ij} \rho_{jj}^{r-1}.$$

Demostració. Notem que si $X_0 = i$, aleshores $\{N_j \geq r\} = \{T_j^{(r)} < \infty\}$. Dit això, provarem el resultat anterior per inducció.

Considerem el cas inicial $r = 1$. Efectivament,

$$P(N_j \geq 1 \mid X_0 = i) = P(T_j < \infty \mid X_0 = i) = \rho_{ij}.$$

Suposem que la igualtat se satisfà per r i passem a veure que es compleix per $r + 1$. Fent servir la definició de $T_i^{(r)}$ i probabilitat condicionada obtenim

$$\begin{aligned} P(N_j \geq r + 1 \mid X_0 = i) &= P\left(T_j^{(r+1)} < \infty \mid X_0 = i\right) = P\left(T_j^{(r)} < \infty, S_j^{(r+1)} < \infty \mid X_0 = i\right) \\ &= \frac{P\left(T_j^{(r)} < \infty, S_j^{(r+1)} < \infty, X_0 = i\right)}{P(X_0 = i)} \\ &= \frac{P\left(T_j^{(r)} < \infty, S_j^{(r+1)} < \infty, X_0 = i\right)}{P\left(T_j^{(r)} < \infty, X_0 = i\right)} \cdot \frac{P\left(T_j^{(r)} < \infty, X_0 = i\right)}{P(X_0 = i)} \\ &= P\left(S_j^{(r+1)} < \infty \mid T_j^{(r)} < \infty, X_0 = i\right) \cdot P\left(T_j^{(r)} < \infty \mid X_0 = i\right) \\ &= \left[\sum_{n=1}^{\infty} P\left(S_j^{(r+1)} = n \mid T_j^{(r)} < \infty\right)\right] \cdot P(N_j \geq r \mid X_0 = i) \end{aligned}$$

Ara, usant el Lema 5.1.2 i la hipòtesi d'inducció, obtenim

$$\begin{aligned} P(N_j \geq r + 1 \mid X_0 = i) &= \left[\sum_{n=1}^{\infty} P(T_j = n \mid X_0 = j)\right] \cdot \rho_{ij} \rho_{jj}^{r-1} = P(T_j < \infty \mid X_0 = j) \cdot \rho_{ij} \rho_{jj}^{r-1} \\ &= \rho_{ij} \rho_{jj}^{r}. \end{aligned}$$

□

Observació. En particular, tenim

$$P(N_i \geq 0 \mid X_0 = i) = 1,$$
$$P(N_i \geq r \mid X_0 = i) = \rho_{ii}^r, \qquad r \geq 1.$$

5.2 Criteris de classificació

Una vegada assumides els nocions definides en l'anterior secció, passem a treballar amb els resultats més destacables relacionats amb aquests conceptes. El que interessa, sobretot, és que una vegada acabat aquest apartat siguem capaços d'identificar fàcilment el conjunt d'estats recurrents (resp. transitoris) d'una cadena de Markov homogènia.

Teorema 5.2.1. *Sigui $\{X_n : n \geq 1\}$ una cadena de Markov homogènia, tenim*

(i) Si $\rho_{ii} = P(T_i < \infty \mid X_0 = i) = 1$, llavors i és un estat recurrent i, a més, $\sum_{n\geq 1} p_{ii}^{(n)} = \infty$.

(ii) Si $\rho_{ii} = P(T_i < \infty \mid X_0 = i) < 1$, llavors i és un estat transitori i, a més, $\sum_{n\geq 1} p_{ii}^{(n)} < \infty$.

Demostració. Suposem que $\rho_{ii} = 1$. Usant la propietat de continuïtat seqüencial de la probabilitat i el Lema 5.1.4, obtenim

$$\begin{aligned} P(N_i = \infty \mid X_0 = i) &= \lim_{r\to\infty} P(N_i > r \mid X_0 = i) = \lim_{r\to\infty} P(N_i \geq r+1 \mid X_0 = i) = \lim_{r\to\infty} \rho_{ii}^{r+1} \\ &= \lim_{r\to\infty} 1 = 1. \end{aligned}$$

Estem dient, doncs, que el nombre de vegades que la cadena visita l'estat i és infinit, per tant i és recurrent i, a més,

$$\sum_{n\geq 1} p_{ii}^{(n)} = E(N_i \mid X_0 = i) = \infty.$$

Suposem, ara, que $\rho_{ii} < 1$. Utilitzant la Proposició 2.5 i el Lema 5.1.4, obtenim

$$\sum_{n\geq 1} p_{ii}^{(n)} = E(N_i \mid X_0 = i) = \sum_{r\geq 1} P(N_i \geq r \mid X_0 = i) = \sum_{r\geq 1} \rho_{ii}^r = \frac{\rho_{ii}}{1-\rho_{ii}} < \infty.$$

El fet que $\sum_{n\geq 1} p_{ii}^{(n)} < \infty$, significa que hi ha un nombre finit de vegades que la cadena visita l'estat i, per tant $P(N_i = \infty \mid X_0 = i) = 0$ i l'estat i és transitori. □

Notem que a partir de les definicions d'estat recurrent i transitori, obtenim que el recíproc també és cert. A partir d'aquí, podem afirmar que tot estat d'una cadena de Markov homogènia és transitori o recurrent. A més, hem obtingut dos criteris per determinar la recurrència o transitorietat d'un estat: un en termes de probabilitat de retorn i un altre en probabilitats de transició.

A continuació, citarem diferents eines que ens poden ser útils a l'hora d'identificar quins estats d'una cadena són recurrents i quins transitoris.

Teorema 5.2.2. *Sigui C una classe. Tots els estats de C són recurrents, o bé, tots són transitoris.*

Demostració.

Siguin $i, j \in C$. Suposem que i és un estat transitori, es tracta de veure que j també ho és. Com que i i j es comuniquen, existeixen naturals n i m tal que $p_{ij}^{(n)} > 0$ i $p_{ji}^{(m)} > 0$. Sigui $r \geq 1$, tenim

$$p_{ij}^{(n+r+m)} \geq p_{ij}^{(n)} p_{jj}^{(r)} p_{ji}^{(m)} \implies p_{jj}^{(r)} \leq \frac{p_{ij}^{(n+r+m)}}{p_{ij}^{(n)} p_{ji}^{(m)}}$$

Llavors,

$$\sum_{r\geq 1} p_{jj}^{(r)} \leq \frac{1}{p_{ij}^{(n)} p_{ji}^{(m)}} \cdot \sum_{r\geq 1} p_{ii}^{(n+r+m)} < \infty$$

On la última desigualtat és certa pel teorema anterior aplicat a l'estat i. Ara, com que $\sum_{r\geq 1} p_{jj}^{(r)} < \infty$, tornant a aplicar el Teorema 5.2.1 per j, en traiem que j és transitori.

Conseqüentment, com que ser recurrent és complementari de ser transitori, si un estat d'una classe és recurrent, necessàriament la resta ho han de ser. □

Hem vist, doncs, que el fet que un estat sigui recurrent o transitori és una propietat de classe. Així, a partir d'ara, podem parlar de *classes recurrents* o *transitòries*.

Teorema 5.2.3. *Sigui $i \in I$ un estat recurrent. Si $i \longrightarrow j$, aleshores j és un estat recurrent. En altres paraules, un estat transitori no pot ser accessible des d'un estat recurrent.*

Demostració. Com que j és accessible des de i, sabem que existeix un natural n tal que $p_{ij}^{(n)} > 0$. Per tant, com que i és recurrent, la cadena hi retorna infinites vegades, i en cada una d'aquestes, existeix una probabilitat fixa positiva tal que després de n etapes, la cadena es trobarà a j. Així, la cadena passarà infinites vegades per j, és a dir, j és recurrent. □

Teorema 5.2.4. *Tota classe C recurrent és tancada.*

Demostració. Suposem que C és una classe no tancada, volem arribar a contradicció. Si C no és tancada, significa que existeix un estat $i \in C$ tal que $\sum_{j\in C} p_{ij} < 1$ i existeix $k \notin C$ tal que $p_{ik} > 0$. Altrament dit, que sortim de la classe C. Com que i és recurrent i $i \longrightarrow k$, pel Teorema 5.2.3 tenim que $k \longrightarrow i$, i per tant $k \in C$, arribant així a contradicció. □

Teorema 5.2.5. *Tota classe C finita i tancada és recurrent.*

Demostració. Sigui C una classe finita i tancada. Suposem que partim d'un estat d'aquesta classe. Per definició, la cadena no pot sortir-ne. Si tots els estats són transitoris, significa que cada un és visitat cap o un nombre finit de vegades, però això no és possible tenint en compte que la cadena es queda a C per sempre. Per tant, existeix almenys un estat recurrent a C. Sabem, però, que la recurrència és una propietat de classe, així C és una classe recurrent. □

Observacions. A partir dels resultats anteriors i els plantejaments que s'han fet servir per demostrar-los, podem afirmar:

- Tota cadena de Markov homogènia finita té almenys un estat recurrent.
- Si existeix un estat $i \in I$ tal que $i \longrightarrow j$, però $j \not\longrightarrow i$, aleshores j és transitori.
- Si I és finit, aleshores $j \in I$ és transitori, si, i només si, existeix un estat $i \in I$ tal que $i \longrightarrow j$, però $j \not\longrightarrow i$.

5.3 Exemples

Tanquem el capítol amb una sèrie exemples per acabar d'entendre el que s'ha explicat durant aquest.

(i) *Passejada aleatòria sobre* $\mathbb{Z}$

Sabem que la cadena és irreductible, és a dir, que té una única classe. Com que el conjunt d'estats és infinit, no podem afirmar que la classe és recurrent, per això estudiarem el cas de l'estat 0, i usant el Teorema 5.2.2, podrem estendre el resultat a la classe. I consegüentment a la cadena, ja que la classe és tota ella.

Observem que per tal que la cadena torni a 0 cal que faci els mateixos passos a l'esquerra que a la dreta. Per tant, per $n = 1, 2, \dots$ tenim

$$\begin{aligned} p_{00}^{(2n-1)} &= 0, \\ p_{00}^{(2n)} &= \binom{2n}{n} \cdot p^n \cdot (1-p)^n. \end{aligned}$$

Per la complexitat de les expressions, donarem una demostració poc rigorosa, fent servir càlculs aproximats. Així, fent ús de la fórmula de Starling[1] tenim

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2} \sim \frac{(2n)^{2n} e^{-2n} \sqrt{4\pi n}}{n^{2n} e^{-2n} 2\pi n} = \frac{2^{2n}}{\sqrt{n\pi}}.$$

Llavors,

$$p_{00}^{(2n)} \sim \frac{2^{2n}}{\sqrt{n\pi}} p^n \cdot (1-p)^n = \frac{1}{\sqrt{n\pi}} (4p(1-p))^n.$$

Volem calcular $\sum_{n=1}^{\infty} p_{00}^{(n)}$ per poder aplicar el Teorema 5.2.1 i determinar si l'estat 0 és recurrent o transitori. Com que $p_{00}^{(m)} = 0$ per m senar, només cal considerar els m parells.

- Suposem que $p = \frac{1}{2}$, tenim

$$p_{00}^{(2n)} \sim \frac{1}{\sqrt{n\pi}},$$

aleshores

$$\sum_{n=1}^{\infty} p_{00}^{(n)} = \sum_{n=1}^{\infty} p_{00}^{(2n)} = \infty.$$

Així, 0 és un estat recurrent, i per tant la cadena també.

[1] Fórmula de *Starling*: $n! \sim n^n e^{-n} \sqrt{2\pi n}$ quan n s'acosta a infinit.

- Suposant que $p \neq \frac{1}{2}$ tenim que $4p(1-p) < 1$. Aplicant el criteri del quocient[2] amb $a_n = p_{00}^{(2n)}$ obtenim que la sèrie convergeix, és a dir,

$$\sum_{n=1}^{\infty} p_{00}^{(n)} = \sum_{n=1}^{\infty} p_{00}^{(2n)} < \infty.$$

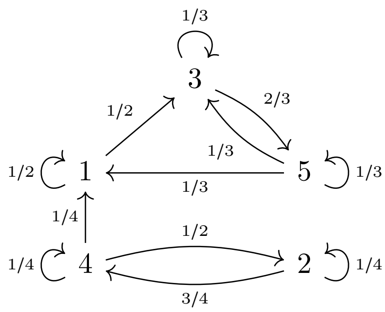
Per tant, la cadena és transitòria, per ser-ho l'estat 0.

(ii) *Passejada aleatòria sobre $\mathbb{Z}$ amb barreres reflectants*

En capítols anteriors hem vist que la cadena té tres classes: $C = \{1, 2, \ldots, N-1\}$, $\{0\}$ i $\{N\}$. Tant 0 com N són estats absorbents, per tant són recurrents i les seves respectives classes també. Pel que fa a la classe C, és suficient en veure el comportament d'un únic estat. Considerant $i = 1$, tenim que $i \longrightarrow 0$, però $0 \not\longrightarrow i$, per tant 1 és un estat transitori, i consegüentment, la classe C també.

(iii) Considerem la cadena de Markov que té per matriu de probabilitats i diagrama associat els següents:

$$\Pi = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/4 & 0 & 3/4 & 0 \\ 0 & 0 & 1/3 & 0 & 2/3 \\ 1/4 & 1/2 & 0 & 1/4 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \end{pmatrix}$$



Tenim dues classes: $\{2, 4\}$ i $\{1, 3, 5\}$. Tenim que 1 és accessible des de 4, però 4 no ho és des d'1. Així, 4 és transitori i, per tant, $\{2, 4\}$ també. Com que la cadena és finita, no pot ser que tots els estats siguin transitoris, així, doncs, $\{1, 3, 5\}$ és una classe recurrent.

Ara, ens preguntem: Si la cadena inicialment es troba a l'estat 4, quina és la probabilitat que la cadena retorni a l'estat 4 almenys 15 vegades?

Es tracta de saber quant val $P(N_4 \geq 15 \mid X_0 = 4)$. Pel Lema 5.1.4, tenim que $P(N_4 \geq 15 \mid X_0 = 4) = \rho_{44}^{15}$, per tant només ens cal determinar ρ_{44}.

Denotem $\rho_{44}^{(n)}$ com la probabilitat de retornar per primera vegada a l'estat 4 en exactament n etapes havent sortit de 4. Així,

$$\rho_{44} = \sum_{n=1}^{\infty} \rho_{44}^{(n)}.$$

Estudiem el valor de $\rho_{44}^{(n)}$ pels diferents valors de n.

- $n = 1$: Partint de 4, per tal de retornar-hi en una única etapa cal que no s'hi mogui, és a dir, $\rho_{44}^{(1)} = p_{44} = \frac{1}{4}$.
- $n = 2$: Per fer-ho en dues etapes, és necessari primer passar a 2 i després, tornar a 4. És a dir,

$$\rho_{44}^{(2)} = p_{42} \cdot p_{24} = \frac{1}{2} \cdot \frac{3}{4}.$$

[2]Criteri del quocient: Si existeix $L = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$, aleshores la sèrie $\sum_{n=0}^{\infty} a_n$ convergeix absolutament.

- $n \geq 3$: El primer retorn a l'estat 4 es pot fer en n etapes si: primer es fa una transició de 4 a 2, en les següents $n-2$ etapes ens mantenim a 2, i finalment en l'última etapa retornem a 4. Per tant,

$$\rho_{44}^{(n)} = p_{42} \cdot p_{22}^{n-2} \cdot p_{24} = \frac{1}{2} \cdot \left(\frac{1}{4}\right)^{n-2} \cdot \frac{3}{4}.$$

D'aquesta manera, ens quedaria

$$\begin{aligned}\rho_{44} &= \sum_{n=1}^{\infty} \rho_{44}^{(n)} = \rho_{44}^{(1)} + \sum_{n=2}^{\infty} \rho_{44}^{(n)} = \frac{1}{4} + \sum_{n=2}^{\infty} \frac{1}{2} \cdot \left(\frac{1}{4}\right)^{n-2} \cdot \frac{3}{4} = \frac{1}{4} + \frac{1}{2} \cdot \frac{3}{4} \sum_{n=2}^{\infty} \left(\frac{1}{4}\right)^{n-2} \\ &= \frac{1}{4} + \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{1}{1-\frac{1}{4}} = \frac{3}{4}.\end{aligned}$$

Per tant, partint de l'estat 4, la probabilitat que la cadena hi retorni almenys 15 vegades és $\rho_{44}^{15} = \left(\frac{3}{4}\right)^{15} \approx 0.013$.

6 Distribucions invariants

Com es comporta una cadena de Markov a llarg termini? Generalment, la cadena no pot convergir a un únic estat degut a l'atzar que presenta i que hem descrit a la matriu de transició. Podem pensar, però, que la cadena *tendeix* a una distribució en concret. Efectivament, sota unes condicions específiques, la distribució de X_n s'*estabilitza.*

En aquest capítol, treballarem amb un tipus de distribució: les distribucions invariants. N'estudiarem l'existència i unicitat, presentant els conceptes de regularitat i ergodicitat d'una cadena de Markov homogènia. Veurem, finalment, que la distribució que *equilibra* la cadena és única i invariant. En tot moment, complementarem els conceptes teòrics amb exemples per a un millor seguiment.

Com en altres capítols i fins que no es digui el contrari, considerem l'espai de probabilitat $(\Omega, \mathcal{A}, P)$. Sigui $\{X_n : n \geq 0\}$ una $CMH(\gamma, \Pi)$, on cada variable aleatòria discreta X_n pren valors a I finit o numerable i $\Pi = (p_{ij})_{i,j \in I}$.

6.1 Distribucions invariants

En aquesta secció ens proposem fer un estudi preliminar d'una cadena de Markov homogènia a llarg temps, és a dir, treballar amb

$$\lim_{n\to\infty} p_{ij}^{(n)}.$$

Definirem el concepte de distribució invariant, n'estudiarem l'existència i veurem la relació que té amb el límit anterior.

Definició. Sigui Π una matriu de transició. Una distribució de probabilitat ν sobre I diem que és una *distribució invariant* (o *estacionària*) *per* Π si se satisfà

$$\nu = \nu\Pi.$$

D'aquesta manera si considerem el vector fila $\nu = (\nu_i)_{i\in I}$, tenim per a cada $i \in I$,

$$\nu_i = \sum_{j\in I} \nu_j \cdot p_{ji}.$$

Iterant n vegades la primera expressió de la definició obtenim $\nu = \nu\Pi_n$ per a qualsevol $n \geq 1$. A més, donada una cadena de Markov homogènia $\{X_n : n \geq 0\}$, a la Proposició 3.2.2, vam poder provar que la llei de cada X_n era

$$P(X_n = k) = \gamma_k^{(n)} = (\gamma\Pi_n)_k.$$

Per tant, si tenim una $CMH(\gamma, \Pi)$ on la distribució inicial és invariant per Π, és a dir, $\gamma = \gamma\Pi_n$, obtenim

$$P(X_n = k) = \gamma_k^{(n)} = (\gamma\Pi_n)_k = \gamma_k.$$

Així X_n té distribució γ independentment de n. Amb això estem dient que si una cadena comença amb una distribució estacionària, continua amb la mateixa distribució per sempre. D'aquí se'n dedueix el següent resultat:

Teorema 6.1.1. *Sigui* $\{X_n : n \geq 0\}$ *una* $CMH(\gamma, \Pi)$ *amb* γ *una distribució invariant per* Π. *Aleshores,* $\{X_{n+m} : n \geq 0\}$ *és una* $CMH(\gamma, \Pi)$ *amb* $m \geq 1$.

Hem definit un tipus de distribució, però no tenim la certesa que existeixin. Amb el següent teorema passem a estudiar-ne l'existència.

Teorema 6.1.2. *Tota matriu de transició* Π *corresponent a una cadena de Markov homogènia amb conjunt d'estats* I *finit, té almenys una distribució invariant.*

Demostració. Sigui v una probabilitat sobre I. Així, si $v = (v_1, \dots, v_{|I|})$ on $|I| = \#I$, aleshores $v_i \in [0,1]$ i $\sum_{i \in I} v_i = 1$. Ara, per a cada $n \geq 1$ definim

$$v^{(n)} = \frac{1}{n} \sum_{k=0}^{n-1} v \Pi_k.$$

Anem a veure que $v^{(n)}$ també és una probabilitat sobre I. És clar que $v_i^{(n)} \geq 0$ per a tot $i \in I$. Cal veure, també, que la suma de $v_i^{(n)}$ sigui 1:

$$\sum_{i \in I} v_i^{(n)} = \frac{1}{n} \sum_{i \in I} \sum_{k=0}^{n-1} \sum_{h \in I} v_h \cdot p_{hi}^{(k)} = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{h \in I} v_h \left[\sum_{i \in I} p_{hi}^{(k)} \right] = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{h \in I} v_h = \frac{1}{n} \sum_{k=0}^{n-1} 1 = 1.$$

A més, com que el conjunt de probabilitats sobre I és un conjunt tancat i acotat de $[0,1]^{|I|}$, existeix una subsuccessió que convergeix feblement[3] a un element del mateix conjunt, pel Teorema de Bolzano-Weierstrass[4]. Altrament dit, existeix una probabilitat ν sobre I tal que per alguna subsuccessió $\{v^{(n_k)}\}_{k \geq 1}$ es compleix

$$v^{(n_k)} \xrightarrow[k \to \infty]{} \nu.$$

Llavors,

$$v^{(n_k)} - v^{(n_k)} \Pi = \frac{1}{n_k} \left[\sum_{l=0}^{n_k - 1} v \Pi_l - \sum_{l=0}^{n_k - 1} v \Pi_{l+1} \right] = \frac{1}{n_k} (v - v \Pi_{n_k}).$$

Ara, fent tendir k a infinit obtenim

$$\nu - \nu \Pi = \lim_{k \to \infty} \frac{1}{n_k} (v - v \Pi_{n_k}) = 0.$$

On en l'última igualtat hem fet servir el fet que $\{v - v\Pi_{n_k}\}_{k \geq 1}$ és una successió acotada. Així, ν és una distribució invariant. □

Gràcies al resultat anterior podem afirmar que el sistema $\nu = \nu \Pi$ té solució sempre i quan Π sigui finita. No podem assegurar, però, que aquesta solució sigui una probabilitat ni que sigui única. Pel que fa el cas infinit, no podem dir-ne res.

Exemples.

(i) Si ν_1 i ν_2 són distribucions invariants per Π, aleshores qualsevol combinació lineal convexa de ν_1 i ν_2 també ho és. És a dir, per a $\lambda \in [0,1]$, $\lambda \nu_1 + (1 - \lambda) \nu_2$ és una distribució invariant.

[3] Una successió de variables aleatòries $\{X_n\}_{n \geq 1}$ amb distribucions $\{F_n\}_{n \geq 1}$ convergeix feblement (o en llei) a una variable aleatòria X amb distribució F si $\lim_{n \to \infty} F_n(x) = F(x)$, als punts $x \in \mathbb{R}$ on F és contínua.

[4] Teorema de Bolzano-Weierstrass: Tota successió que pren valors a un conjunt $S \subset \mathbb{R}^n$ tancat i acotat, té una subsuccessió convergent a un element de S.

(ii) Si $\Pi = Id$, llavors tota distribució és invariant.

(iii) *Passejada aleatòria sobre $\mathbb{Z}$ amb barreres absorbents*

Com en altres capítols, estudiem aquest exemple considerant el conjunt d'estats $I = \{0, 1, \dots, N\}$ i matriu de transició $(N+1) \times (N+1)$ següent

$$\Pi = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ q & 0 & p & \dots & 0 & 0 & 0 \\ 0 & q & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & p & 0 \\ 0 & 0 & 0 & \dots & q & 0 & p \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}$$

Volem trobar les distribucions invariants per aquesta matriu. Per això, imposem la igualtat $\nu = \nu\Pi$ i obtenim

$$\begin{cases} \nu_0 = \nu_0 + \nu_1 \cdot q \\ \nu_1 = \nu_2 \cdot q \\ \nu_j = \nu_{j-1} \cdot p + \nu_{j+1} \cdot q, \qquad j = 2, \dots, N-2 \\ \nu_{N-1} = \nu_{N-2} \cdot p \\ \nu_N = \nu_N + \nu_{N-1} \cdot p \end{cases}$$

D'on se'n treu que $\nu_1 = \nu_2 = \dots = \nu_{N-1} = 0$. No tenim cap restricció per ν_0 i ν_N, però sabem que per tal que ν sigui una distribució s'ha de complir $\nu_j \geq 0$ per a tot element j del conjunt d'estats i, a més, $\sum_{j \in I} \nu_j = 1$. Per tant, $\nu_0 + \nu_N = 1$ amb $\nu_0, \nu_N \geq 0$. Així, per a $\lambda \in [0,1]$, obtenim la família de distribucions invariants per Π següent:

$$\nu = (\lambda, 0, \dots, 0, 1-\lambda).$$

(iv) Estudiem un altre exemple que també tenim vist, la cadena de dos estats $\alpha, \beta \in (0,1)$ de matriu de transició

$$\Pi = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

Imposant la igualtat que determina una distribució invariant, obtenim les expressions

$$\begin{cases} \nu_1 = \nu_1(1-\alpha) + \nu_2\beta \\ \nu_2 = \nu_1\alpha + \nu_2(1-\beta) \end{cases}$$

que es redueix a l'equació $\nu_1\alpha = \nu_2\beta$. Imposant que ν és una distribució, obtenim una única distribució invariant per Π

$$\nu = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right).$$

El següent teorema ens permetrà veure la relació que hi ha entre les distribucions invariants i $\lim_{n\to\infty} p_{ij}^{(n)}$ quan tractem amb cadenes finites.

Teorema 6.1.3. *Sigui I un conjunt finit. Suposem que per a algun $i \in I$ se satisfà*

$$p_{ij}^{(n)} \xrightarrow[n\to\infty]{} \nu_j, \text{ per a tot } j \in I.$$

Aleshores, $\nu = \{\nu_j : j \in I\}$ és una distribució invariant per Π.

Demostració. Cal veure que ν és distribució i que és invariant per Π. Veiem, primer, que és distribució sobre I. Tenim

$$\nu_j = \lim_{n\to\infty} p_{ij}^{(n)} \geq 0.$$

I a més,

$$\sum_{j\in I} \nu_j = \sum_{j\in I} \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \sum_{j\in I} p_{ij}^{(n)} = 1.$$

On en la segona igualtat hem pogut fer l'intercanvi de límit i sumatori gràcies a la finitud de I. Ara només falta demostrar que ν és invariant. Per això ens cal veure que $\nu = \nu\Pi$. Efectivament, obtenim

$$\nu_j = \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \sum_{k\in I} p_{ik}^{(n-1)} p_{kj} = \sum_{k\in I} \lim_{n\to\infty} p_{ik}^{(n-1)} p_{kj} = \sum_{k\in I} \nu_k p_{kj}.$$

Hem obtingut, doncs, que ν és una distribució invariant per Π. □

Exemple.

Considerem l'exemple de la cadena de dos estats $\alpha, \beta \in (0, 1)$. Anteriorment, hem vist que hi ha una única distribució invariant per la matriu de transició que determina aquesta cadena:

$$\nu = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta} \right)$$

Passem a veure-ho fent ús del Teorema 6.1.3. Hem vist en altres capítols que

$$p_{11}^{(n)} = \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n,$$
$$p_{12}^{(n)} = \frac{\alpha}{\alpha+\beta} + \frac{-\alpha}{\alpha+\beta}(1-\alpha-\beta)^n.$$

Ara, fent tendir n a infinit, ens queda

$$\lim_{n\to\infty} p_{11}^{(n)} = \frac{\beta}{\alpha+\beta},$$
$$\lim_{n\to\infty} p_{12}^{(n)} = \frac{\alpha}{\alpha+\beta}.$$

Aplicant el Teorema 6.1.3 per $i = 1$, obtenim

$$\nu = (\nu_1, \nu_2) = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta} \right)$$

és distribució invariant per Π.

6.2 Ergodicitat

En la secció anterior hem provat l'existència de distribucions invariants quan tractàvem amb cadenes de Markov homogènies amb conjunt d'estats finit. En aquest apartat n'estudiarem la unicitat. Per això, cal que prèviament tinguem assumides les següents definicions.

Definició. Una matriu de transició Π és *regular* si existeix un nombre enter $n_0 \geq 1$ tal que
$$\min_{i,j\in I} p_{ij}^{(n_0)} > 0.$$
Així, diem que una $CMH(\gamma,\Pi)$ és *regular* si la seva matriu de transició ho és.

Una manera d'interpretar la definició anterior és que direm que una cadena de Markov homogènia és regular si existeix una potència de la matriu de transició Π de manera que tots els elements són estrictament majors a 0.

Definició. Una cadena de Markov homogènia es diu que és *ergòdica* si per a tot $i \in I$ existeix el límit
$$\lim_{n\to\infty} p_{ij}^{(n)}$$
que és independent de i i defineix una distribució no degenerada sobre I. En altres paraules, que se satisfà
$$\lim_{n\to\infty} p_{ij}^{(n)} = \nu_j \text{ tal que } \sum_{j\in I} \nu_j = 1 \text{ i } \nu_j > 0 \text{ per a tot } j \in I.$$

Cal observar que si apliquem el Teorema 6.1.3 a la definició anterior, obtenim que sempre i quan el conjunt d'estats de la cadena, I, sigui finit, la distribució que ve definida per l'ergodicitat és invariant.

Dit això, podem enunciar el teorema que es coneix per *Teorema ergòdic per cadenes de Markov homogènies finites.*

Teorema 6.2.1. *Sigui $\{X_n : n \geq 0\}$ una $CMH(\gamma,\Pi)$ amb conjunt d'estats I finit. La cadena és regular si, i només si, és ergòdica.*

Demostració. Comencem demostrant la implicació directa. Vegem, primer, que existeix $\nu = (\nu_j)_{j\in I}$ tal que per a tot $i \in I$, $\lim_{n\to\infty} p_{ij}^{(n)} = \nu_j$. Per això, definim per $n \geq 1$,
$$m_j^{(n)} = \min_{i\in I} p_{ij}^{(n)} \text{ i } M_j^{(n)} = \max_{i\in I} p_{ij}^{(n)}.$$
Utilitzant la l'equació de Chapman-Kolmogorov obtenim
$$m_j^{(n+1)} = \min_{i\in I} p_{ij}^{(n+1)} = \min_{i\in I} \sum_{k\in I} p_{ik} \cdot p_{kj}^{(n)} \geq \min_{i\in I} \sum_{k\in I} p_{ik} \min_{k\in I} p_{kj}^{(n)} = \min_{k\in I} p_{kj}^{(n)} = m_j^{(n)}.$$

De manera anàloga podem obtenir que $M_j^{(n+1)} \leq M_j^{(n)}$. Per tant, $\{m_j^{(n)}\}_{n\geq 1}$ és una successió creixent acotada superiorment i $\{M_j^{(n)}\}_{n\geq 1}$ és una successió decreixent acotada inferiorment. Al ser successions acotades, sabem que existeix límit que denotarem per m_j i M_j, respectivament. Notem que és suficient veure que per a $j \in I$,
$$\lim_{n\to\infty} (M_j^{(n)} - m_j^{(n)}) = 0. \tag{6.2.1}$$

En efecte, si es complís la igualtat descrita, obtindríem que $m_j = M_j$, i com que per definició tenim $m_j^{(n)} \leq p_{ij}^{(n)} \leq M_j^{(n)}$, passant al límit quan n s'apropa a infinit, quedaria
$$\lim_{n\to\infty} p_{ij}^{(n)} = \nu_j,$$

on $\nu_j = m_j = M_j$, per $j \in I = \{1, \dots, N\}$.

Passem a veure, doncs, que se satisfà la igualtat (6.2.1). Com que la cadena és regular, existeix un enter $n_0 \geq 1$ tal que $\min_{i,j\in I} p_{ij}^{(n_0)} > 0$. Sigui ϵ aquest valor. Aleshores, fent servir l'equació de Chapman-Kolmogorov, obtenim

$$\begin{aligned} p_{ij}^{(n_0+n)} &= \sum_{k\in I} p_{ik}^{(n_0)} p_{kj}^{(n)} = \sum_{k\in I} \left(p_{ik}^{(n_0)} - \epsilon p_{jk}^{(n)}\right) p_{kj}^{(n)} + \epsilon \sum_{k\in I} p_{jk}^{(n)} p_{kj}^{(n)} \\ &= \sum_{k\in I} \left(p_{ik}^{(n_0)} - \epsilon p_{jk}^{(n)}\right) p_{kj}^{(n)} + \epsilon p_{jj}^{(2n)}. \end{aligned}$$

Com que, per definició, $\epsilon \leq p_{ik}^{(n_0)}$ i $p_{jk}^{(n)} \leq 1$, tenim $p_{ik}^{(n_0)} - \epsilon p_{jk}^{(n)} \geq 0$. Així, podem afirmar

$$p_{ij}^{(n_0+n)} \geq m_j^{(n)} \sum_{k\in I} \left(p_{ik}^{(n_0)} - \epsilon p_{jk}^{(n)}\right) + \epsilon p_{jj}^{(2n)} = (1-\epsilon) m_j^{(n)} + \epsilon p_{jj}^{(2n)}.$$

Aplicant mínims a l'expressió resultant queda

$$m_j^{(n_0+n)} \geq m_j^{(n)}(1-\epsilon) + \epsilon p_{jj}^{(2n)}.$$

Anàlogament, obtenim la següent expressió pel màxim

$$M_j^{(n_0+n)} \leq M_j^{(n)}(1-\epsilon) + \epsilon p_{jj}^{(2n)}.$$

Per tant,

$$M_j^{(n_0+n)} - m_j^{(n_0+n)} \leq (1-\epsilon)\left(M_j^{(n)} - m_j^{(n)}\right).$$

Ara, iterant i usant que ϵ és positiu i acotat per 1, obtenim

$$0 \leq M_j^{(kn_0+n)} - m_j^{(kn_0+n)} \leq (1-\epsilon)^k \left(M_j^{(n)} - m_j^{(n)}\right) \xrightarrow[n\to\infty]{} 0.$$

Tenim, doncs, que existeix una subsuccessió $\{n_k\}_{k\geq 1}$ tal que

$$\lim_{k\to\infty} (M_j^{(n_k)} - m_j^{(n_k)}) = 0.$$

Ara bé, la successió $\{M_j^{(n)} - m_j^{(n)}\}_{n\geq 1}$ és monòtona decreixent, així obtenim el resultat que volíem

$$\lim_{n\to\infty} (M_j^{(n)} - m_j^{(n)}) = 0.$$

Per acabar de provar la implicació directa, ens cal veure que efectivament ν és una distribució no degenerada sobre I.

Com que hem vist $\lim_{n\to\infty} p_{ij}^{(n)} = \nu_j$, tenim

$$\sum_{j\in I} p_{ij}^{(n)} \xrightarrow[n\to\infty]{} \sum_{j\in I} \nu_j,$$

i, a més, $\sum_{j\in I} p_{ij}^{(n)} = 1$, per tant $\sum_{j\in I} \nu_j = 1$.

A més, tenim que per a tot $n \geq n_0$, se satisfà $m_j^{(n)} \geq m_j^{(n_0)} \geq \epsilon > 0$, per tant

$$\nu_j = \lim_{n\to\infty} m_j^{(n)} \geq \epsilon > 0.$$

Hem obtingut que ν és una distribució no degenerada sobre I. A més, com que I és finit, ν és invariant. Finalitzem, així, la demostració de la implicació directa. Ara, suposem que la cadena és ergòdica i volem veure que és regular.

Com que, per a tot $i \in I$ se satisfà $\lim_{n\to\infty} p_{ij}^{(n)} = \nu_j > 0$, llavors per a qualsevol $j \in I$ existeix un enter n_j tal que si $n \geq n_j$, tenim $p_{ij}^{(n)} > 0$, $\forall i \in I$. En particular, $\min_{i\in I} p_{ij}^{(n_j)} > 0$. Ara, si considerem $n_0 = \max(n_1, \ldots, n_N)$, tenim

$$\min_{i\in I} p_{ij}^{(n_0)} > 0,$$

per qualsevol $j \in I$. Per tant, Π és regular, i conseqüentment, ho és la cadena. □

La proposició següent ens confirma la unicitat d'una distribució invariant sota les hipòtesis imposades en l'anterior teorema.

Proposició 6.2.2. *Sigui $\{X_n : n \geq 0\}$ una $CMH(\gamma, \Pi)$ regular amb I finit. Aleshores,*

(a) La distribució que defineix l'anterior teorema és única.

(b) La llei de X_n convergeix feblement a la distribució invariant de la cadena quan n s'acosta a ∞, independentment de la γ considerada.

Demostració.

(a) Com que la cadena és regular i té conjunt d'estats finit, aplicant el Teorema 6.2.1, tenim que la cadena és ergòdica i defineix una distribució invariant. Volem veure que és única, per això suposem que n'hi ha una altra $\hat{\nu} = (\hat{\nu}_j)_{j\in I}$. Ha de complir

$$\hat{\nu} = \hat{\nu}\Pi_n,$$

és a dir, per a tot $j \in I$ tenim

$$\hat{\nu}_j = \sum_{k\in I} \hat{\nu}_k p_{kj}^{(n)}.$$

Fent el límit de l'última expressió quan n s'apropa a infinit i fent servir que I és finit, obtenim

$$\hat{\nu}_j = \sum_{k\in I} \hat{\nu}_k \lim_{n\to\infty} p_{kj}^{(n)} = \sum_{k\in I} \hat{\nu}_k \nu_j = \nu_j.$$

(b) Usant que I és finit, la llei de X_n descrita a la Proposició 3.2.2 i el primer resultat d'aquesta la proposició, obtenim el que volíem

$$\lim_{n\to\infty} P(X_n = j) = \lim_{n\to\infty} (\gamma\Pi_n)_j = \lim_{n\to\infty} \sum_{i\in I} \gamma_i p_{ij}^{(n)} = \left(\sum_{i\in I} \gamma_i\right) \lim_{n\to\infty} p_{ij}^{(n)} = \nu_j.$$

□

Cal destacar que el recíproc del primer resultat de la Proposició 6.2.2 no és cert. És a dir, pot donar-se el cas que existeixi una única distribució invariant per Π finita, però que la

$CMH(\gamma, \Pi)$ no sigui ergòdica. Per exemple, considerem la cadena de Markov homogènia que tingui per matriu de transició la següent:

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

Imposant la condició $\nu = \nu\Pi$ amb $\nu = (\nu_1, \nu_2)$, obtenim

$$\nu_1 = \nu_2.$$

Per tal que sigui distribució cal que $\nu_1 + \nu_2 = 1$. D'aquí, obtenim que la cadena té una única distribució invariant $\nu = (\frac{1}{2}, \frac{1}{2})$.

D'altra banda, per inducció podem veure que per a tot $n \geq 0$,

$$\Pi_{2n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{i} \qquad \Pi_{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Així, doncs, $\lim\limits_{n\to\infty} p_{ij}^{(n)}$ no existeix i la cadena no és ergòdica.

Hem vist, doncs, que una cadena de Markov amb I finit, té una única distribució invariant, però no és ergòdica.

D'altra banda, del segon resultat de la proposició anterior podem extreure'n que si tenim una cadena de Markov homogènia finita i regular, podem trobar la llei de X_n de manera aproximada calculant la distribució invariant de la cadena, ν. És a dir, resolent el sistema lineal $\nu(Id - \Pi) = 0$.

De vegades pot ser complicat determinar si una cadena és regular o no. El següent criteri ens serà útil quan tractem amb cadenes irreductibles.

Lema 6.2.3. *Sigui $\{X_n : n \geq 0\}$ una cadena de Markov homogènia finita i irreductible. Si existeix $k \in I$ tal que $p_{kk} > 0$, aleshores la cadena és regular.*

Demostració. Fixem dos estats $i, j \in I$. Com que la cadena és finita i irreductible, existeix un enter $n_{ij} = n(i,j) \geq 1$ tal que $p_{ij}^{(n_{ij})} > 0$. Considerem $n = \max\limits_{i,j\in I} n_{ij}$ i $m = 2n - n_{ik} - n_{kj}$. Es compleix

$$p_{ij}^{(2n)} = p_{ik}^{(n_{ik})} p_{kk} \overset{m)}{\cdots} p_{kk} p_{kj}^{(n_{kj})} > 0.$$

Així, tots els elements de la matriu de transició Π_{2m} són estrictament positius, per tant la cadena és regular. □

Tanquem la part teòrica de la secció estudiant la relació entre estats transitoris i distribucions invariants.

Proposició 6.2.4. *Sigui $\{X_n : n \geq 0\}$ una cadena Markov homogènia finita. Si $i \in I$ és un estat transitori i ν una distribució invariant, aleshores $\nu_i = 0$.*

Demostració. Com que ν és una distribució invariant, tenim

$$\nu_i = \sum_{j\in I} \nu_j p_{ji}^{(n)}.$$

A més, el fet que i sigui un estat transitori, implica $\lim\limits_{n\to\infty} p_{ji}^{(n)} = 0$, ja que a partir d'un instant, una vegada sortim de i, ja no hi tornem més.

Usant això i que I és finit, obtenim

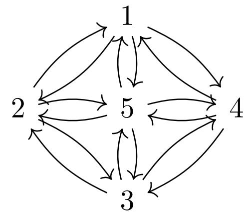
$$\nu_i = \lim_{n\to\infty} \nu_i = \lim_{n\to\infty} \sum_{j\in I} \nu_j p_{ji}^{(n)} = \sum_{j\in I} \nu_j \lim_{n\to\infty} p_{ji}^{(n)} = 0.$$

□

Es pot demostrar que el resultat anterior s'estén a I numerable. D'aquí, podem concloure que si una cadena de Markov homogènia finita té algun estat transitori, llavors no pot ser ergòdica. Podem observar, també, que si una una cadena és irreductible i existeix una distribució invariant, necessàriament ha de ser recurrent.

Exemple. Considerem la cadena de Markov homogènia que té com a matriu de transició associada i diagrama següents:

$$\Pi = \begin{pmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 \end{pmatrix}$$



La cadena és finita, així que té almenys una distribució invariant pel Teorema 6.1.2. Per afirmar que la distribució és única cal veure que la cadena és regular. D'aquesta manera, aplicant el Teorema 6.2.1, obtindrem que la cadena és ergòdica, i conseqüentment, que la distribució que defineix és única per la Proposició 6.2.2 (a). Vegem, doncs, que la cadena és regular.

Notem que encara que la cadena sigui irreductible, no podem fer servir el criteri que hem donat al Lema 6.2.3, perquè tots els elements de la diagonal són 0. Tenim, però, que

$$\Pi_2 = \begin{pmatrix} 11/36 & 1/12 & 11/36 & 1/12 & 2/9 \\ 1/12 & 11/36 & 1/12 & 11/36 & 2/9 \\ 11/36 & 1/12 & 11/36 & 1/12 & 2/9 \\ 1/12 & 11/36 & 1/12 & 11/36 & 2/9 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/3 \end{pmatrix}$$

Hem obtingut, doncs, una potència de Π amb tots els elements estrictament positius, per tant la cadena és regular. Ara, la distribució que trobem resolent el sistema $\nu = \nu\Pi$ serà única, i a més, la llei de X_n tendirà feblement a aquesta quan n s'acosti a infinit. Imposant $\nu = \nu\Pi$ queda

$$\begin{cases} \nu_1 = \frac{1}{3}(\nu_2 + \nu_4) + \frac{1}{4}\nu_5, \\ \nu_2 = \frac{1}{3}(\nu_1 + \nu_3) + \frac{1}{4}\nu_5, \\ \nu_3 = \frac{1}{3}(\nu_2 + \nu_4) + \frac{1}{4}\nu_5, \\ \nu_4 = \frac{1}{3}(\nu_1 + \nu_3) + \frac{1}{4}\nu_5, \\ \nu_5 = \frac{1}{3}(\nu_1 + \nu_2 + \nu_3 + \nu_4). \end{cases}$$

Resolent el sistema imposant que ν compleix la condició de probabilitat $\sum_{i\in I} \nu_i = 1$, obtenim que la única distribució invariant per aquesta cadena és

$$\nu = (\nu_1, \nu_2, \nu_3, \nu_4, \nu_5) = \left(\frac{3}{16}, \frac{3}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{4}\right).$$

7 Comportament al límit

A llarg d'aquest capítol continuarem amb l'estudi del comportament a llarg termini d'una cadena de Markov homogènia, que ja havíem iniciat prèviament. Més concretament, definirem el concepte de temps mitjà de recurrència d'un estat, tot veient-ne la relació amb les distribucions invariants definides en l'anterior capítol.

Recordem que hem vist que sota unes condicions concretes, la distribució límit d'una Cadena de Markov homogènia coincideix amb l'única distribució invariant per Π. En aquest apartat, veurem la generalització d'aquest fet.

7.1 Comportament en mitjana

Sigui $\{X_n : n \geq 0\}$ una cadena de Markov homogènia irreductible i recurrent. El propòsit principal d'aquesta secció és fer un estudi del comportament mitjà de la cadena a llarg temps, o sigui, analitzar el següent límit

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} p_{ij}^{(k)}.$$

Al Corol·lari 5.1.3, provàvem que el nombre de vegades esperat que la cadena visiti l'estat j és $\sum_{k\geq 0} p_{ij}^{(k)}$. D'aquesta manera, intuïtivament, podem pensar que la successió que hem definit anteriorment, pot estar relacionada amb la *versió finita* de N_j. Efectivament, si definim el *temps d'ocupació* de l'estat j en les n primeres etapes com

$$N_j^{(n)} = \#\{0 \leq k \leq n : X_k = j\} = \sum_{k=0}^{n} \mathbb{1}_{\{X_k = j\}},$$

tenim

$$E(N_j^{(n)} \mid X_0 = i) = \sum_{k=0}^{n} E(\mathbb{1}_{\{X_k=j\}} \mid X_0 = i) = \sum_{k=0}^{n} P(X_k = j \mid X_0 = i) = \sum_{k=0}^{n} p_{ij}^{(k)}.$$

Així, podem dir que s'espera que, en mitjana, el temps d'ocupació de l'estat j sigui

$$\frac{1}{n+1} E(N_j^{(n)} \mid X_0 = i) = \frac{1}{n+1} \sum_{k=0}^{n} p_{ij}^{(k)}.$$

Veurem, més endavant, que per complir l'objectiu que ens hem proposat a l'inici de l'apartat és necessari que tinguem entesa la següent noció:

Definició. Es defineix el *temps de retorn esperat* (o *temps mitjà de recurrència*) de l'estat j com

$$m_{jj} = E(T_j \mid X_0 = j) = \sum_{k=1}^{\infty} k P(T_j = k \mid X_0 = j) = \sum_{k=1}^{\infty} k \rho_{jj}^{(k)}.$$

Recordem que $\rho_{ij}^{(n)}$ designa la probabilitat de visitar per primera vegada l'estat j en exactament n etapes partint de l'estat i, és a dir, $\rho_{ij}^{(n)} = P(T_j = n \mid X_0 = i)$.

Donem un resultat relacionat amb el concepte anterior i que necessitarem més tard.

Proposició 7.1.1. *Per a tot $n \geq 1$, se satisfà*

$$p_{ij}^{(n)} = \sum_{k=1}^{n} \rho_{ij}^{(k)} p_{jj}^{(n-k)}.$$

Demostració. Usant la definició de T_j i de probabilitat condicionada, obtenim

$$\begin{aligned} p_{ij}^{(n)} &= P(X_n = j \mid X_0 = i) = P(X_n = j, T_j \leq n \mid X_0 = i) = \sum_{k=1}^{n} P(X_n = j, T_j = k \mid X_0 = i) \\ &= \sum_{k=1}^{n} P(X_n = j \mid T_j = k, X_0 = i) P(T_j = k \mid X_0 = i) \\ &= \sum_{k=1}^{n} P(X_n = j \mid X_0 = i, X_1 \neq j, \ldots, X_{k-1} \neq j, X_k = j) P(T_j = k \mid X_0 = i). \end{aligned}$$

Ara, aplicant la propietat de Markov al primer factor de l'expressió resultant, ens queda

$$p_{ij}^{(n)} = \sum_{k=1}^{n} P(X_n = j \mid X_k = j) P(T_j = k \mid X_0 = i) = \sum_{k=1}^{n} p_{jj}^{(n-k)} \rho_{ij}^{(k)}.$$

□

Convé destacar que m_{jj} indica el temps esperat que tarda la cadena en retornar a j per primera vegada. Així, de manera poc rigorosa, podem dir que en n etapes, el nombre de retorns a j serà $\frac{n}{m_{jj}}$. D'aquesta manera, si considerem n prou gran, tindrem

$$E(N_j^{(n)} \mid X_0 = j) \sim \frac{n}{m_{jj}}.$$

Per tant,

$$\frac{1}{n+1} E(N_j^{(n)} \mid X_0 = j) \sim \frac{1}{m_{jj}}.$$

Cal recalcar que l'estat de partida de la cadena és irrellevant, ja que una vegada hem fet la primera visita a l'estat j, el passat deixa de tenir importància.

Seguidament, passem a veure formalment que el plantejament anterior és vàlid. Per a poder-ho demostrar, és necessari que abans enunciem el següent resultat de sèries de potències (conegut com un dels teoremes de Hardy i Littlewood):

Teorema 7.1.2. *Sigui $A(z) = \sum_{n=0}^{\infty} a_n z^n$ amb $a_n \geq 0$. Suposem que la sèrie convergeix per $0 < z < 1$, aleshores*

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{k=0}^{n} a_k = \lim_{z \to 1} A(z)(1-z).$$

Ara sí, podem enunciar el teorema que ens proporcionarà el valor del límit que hem presentat al principi de la secció.

Teorema 7.1.3. *Sigui $\{X_n : n \geq 0\}$ una cadena de Markov homogènia irreductible i recurrent. Llavors, per a $j \in I$ se satisfà*

$$\lim_{n \to \infty} \frac{1}{n+1} \sum_{k=0}^{n} p_{ij}^{(k)} = \frac{1}{m_{jj}}.$$

Demostració. Definim les següents sèries de potències en $\mathbb{R}$

$$P_{ij}(z) = \sum_{n=0}^{\infty} p_{ij}^{(n)} z^n,$$
$$R_{ij}(z) = \sum_{n=0}^{\infty} \rho_{ij}^{(n)} z^n.$$

Notem que com que $i, j \in I$ són recurrents, usant el Criteri de l'arrel[5] en sèries de potències, obtenim que les dues sèries són convergents per $|z| < 1$.

A més, fent servir la Proposició 7.1.1 obtenim

$$\begin{aligned} P_{ij}(z) &= p_{ij}^{(0)} + \sum_{n=1}^{\infty} \left(\sum_{k=1}^{n} \rho_{ij}^{(k)} p_{jj}^{(n-k)} z^k \right) z^{n-k} = p_{ij}^{(0)} + \sum_{k=1}^{\infty} \rho_{ij}^{(k)} z^k \left(\sum_{n=k}^{\infty} p_{jj}^{(n-k)} z^{n-k} \right) \\ &= p_{ij}^{(0)} + \sum_{k=1}^{\infty} \rho_{ij}^{(k)} z^k \left(\sum_{n=0}^{\infty} p_{jj}^{(n)} z^n \right) = p_{ij}^{(0)} + R_{ij}(z) P_{jj}(z). \end{aligned}$$

Observem que el cas que $i = j$, de l'expressió anterior en traiem $P_{jj}(z) = \frac{1}{1-R_{jj}(z)}$, per tant

$$P_{ij}(z) = p_{ij}^{(0)} + \frac{R_{ij}(z)}{1 - R_{jj}(z)}.$$

Ara bé, si apliquem el Teorema 7.1.2 amb $A(z) = P_{ij}(z)$, obtenim

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} p_{ij}^{(k)} = \lim_{z\to 1} (1-z) P_{ij}(z) = p_{ij}^{(0)} \lim_{z\to 1} (1-z) + \lim_{z\to 1} \frac{1-z}{1 - R_{jj}(z)}. \tag{7.1.1}$$

On en l'última igualtat hem fet servir que tot estat de la cadena és recurrent, així que

$$\lim_{z\to 1} R_{ij}(z) = R_{ij}(1) = \sum_{n\geq 1} \rho_{ij}^{(n)} = \rho_{ij} = 1.$$

De fet, com que el primer sumand de (7.1.1) val 0 per a qualssevol $i, j \in I$, només cal determinar el valor del segon. Notem que com que tota sèrie de potències és infinitament diferenciable en $|x| < R$, amb R el radi de convergència, i les seves derivades resulten de derivar la sèrie original terme a terme, podem aplicar la Regla de l'Hôpital al segon sumand obtenint

$$\lim_{z\to 1} \frac{1-z}{1 - R_{jj}(z)} = \lim_{z\to 1} \frac{1}{R'_{jj}(z)} = \lim_{z\to 1} \frac{1}{\sum\limits_{k\geq 1} k \rho_{jj}^{(k)} z^{k-1}}. \tag{7.1.2}$$

Finalment, aplicant el Teorema d'Abel[6] al denominador de (7.1.2) tenint en compte que la derivada d'una sèrie de potències convergeix en el mateix interval que la sèrie original, ens queda

$$\lim_{z\to 1} \frac{1-z}{1 - R_{jj}(z)} = \frac{1}{\sum\limits_{k\geq 1} k \rho_{jj}^{(k)}} = \frac{1}{m_{jj}}.$$

[5] Criteri de l'arrel: Si existeix radi de convergència $R = \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}}$, aleshores $\sum\limits_{n=0}^{\infty} a_n z^n$ per a $|z| < R$.

[6] Teorema d'Abel: Sigui $F(z) = \sum_{k\geq 0} a_k z^k$ una sèrie de potències de coeficients reals convergent per $|z| < 1$ i tal que $\sum_{k\geq 0} a_k$ és convergent, aleshores $\lim_{z\to 1^-} F(z) = \sum_{k\geq 0} a_k$.

Com que la cadena és recurrent, tenim $m_{jj} > 0$, per tant el límit (7.1.2) existeix i ha tingut sentit aplicar l'Hôpital. Així, doncs, ens queda

$$\lim_{n\to\infty} \frac{1}{n+1}\sum_{k=0}^{n} p_{ij}^{(k)} = \frac{1}{m_{jj}}.$$

□

7.2 Temps mitjà de recurrència i distribucions invariants

L'objectiu d'aquesta secció es veure quina relació hi ha entre el temps mitjà de recurrència i les distribucions invariants que havíem definit al capítol anterior. Així, suposem que tenim una cadena de Markov homogènia amb matriu de transició Π i conjunt d'estats finit $I = \{1, \dots, M\}$. Definim per a cada $j \in I$,

$$\omega_j = \frac{1}{m_{jj}}.$$

Amb el següent teorema, veurem que sota unes condicions específiques, el vector que acabem de definir $\omega = (\omega_j)_{j\in I}$ és la única distribució invariant per Π.

Teorema 7.2.1. *Sigui $\{X_n : n \geq 0\}$ una $CMH(\gamma, \Pi)$ finita i irreductible, aleshores ω és una distribució invariant no degenerada sobre I. A més, qualsevol altra solució de $x = x\Pi$ és múltiple de ω.*

Demostració. Veiem, primer, que efectivament ω és solució del sistema $x = x\Pi$. Usant l'equació de Chapman-Kolmogorov, obtenim

$$\frac{1}{n+1}\sum_{l=0}^{n} p_{ik}^{(l+1)} = \frac{1}{n+1}\sum_{l=0}^{n}\left(\sum_{j\in I} p_{ij}^{(l)} p_{jk}\right) = \sum_{j\in I}\left(\frac{1}{n+1}\sum_{l=0}^{n} p_{ij}^{(l)}\right) p_{jk}. \tag{7.2.1}$$

D'una banda, podem escriure la primera expressió de la següent manera

$$\frac{1}{n+1}\sum_{l=0}^{n} p_{ik}^{(l+1)} = \frac{p_{ik}^{(n+1)}}{n+1} + \frac{1}{n+1}\sum_{l=1}^{n} p_{ik}^{(l)} = \frac{p_{ik}^{(n+1)}}{n+1} + \frac{1}{n+1}\sum_{l=0}^{n} p_{ik}^{(l)} - \frac{p_{ik}^{(0)}}{n+1}$$

Així, si fem el límit quan n s'acosta a infinit resulta

$$\lim_{n\to\infty} \frac{1}{n+1}\sum_{l=0}^{n} p_{ik}^{(l+1)} = \lim_{n\to\infty} \frac{1}{n+1}\left(p_{ik}^{(n+1)} - p_{ik}^{(0)}\right) + \lim_{n\to\infty} \frac{1}{n+1}\sum_{l=0}^{n} p_{ik}^{(l)} = \frac{1}{m_{kk}} = \omega_k.$$

On en la penúltima igualtat hem fet servir que el primer sumand val 0 i el Teorema 7.1.3. D'altra banda, si considerem la segona expressió de (7.2.1) i en fem el límit quan n s'apropa a infinit, tenint en compte que I és finit, obtenim

$$\lim_{n\to\infty} \sum_{j\in I}\left(\frac{1}{n+1}\sum_{l=0}^{n} p_{ij}^{(l)}\right) p_{jk} = \sum_{j\in I}\left(\lim_{n\to\infty} \frac{1}{n+1}\sum_{l=0}^{n} p_{ij}^{(l)}\right) p_{jk}.$$

Ara, aplicant el Teorema 7.1.3 resulta

$$\lim_{n\to\infty} \sum_{j\in I}\left(\frac{1}{n+1}\sum_{l=0}^{n} p_{ij}^{(l)}\right) p_{jk} = \sum_{j\in I} \frac{1}{m_{jj}} p_{jk} = \sum_{j\in I} \omega_j p_{jk}.$$

Per tant, hem obtingut

$$\omega_k = \sum_{j \in I} \omega_j p_{jk}.$$

Veurem, ara, que ω és una distribució no degenerada sobre I. Així, amb el que acabem de veure podrem afirmar que és invariant. Una altra vegada fent servir el Teorema 7.1.3 i la finitud de I tenim

$$\sum_{j \in I} \omega_j = \sum_{j \in I} \left(\lim_{n \to \infty} \frac{1}{n+1} \sum_{l=0}^{n} p_{ij}^{(l)} \right) = \lim_{n \to \infty} \frac{1}{n+1} \sum_{l=0}^{n} \sum_{j \in I} p_{ij}^{(l)} = \lim_{n \to \infty} \frac{1}{n+1} \sum_{l=0}^{n} 1 = 1.$$

Això implica que existeix un estat $j_0 \in I$ tal que $\omega_{j_0} > 0$. A més, com que tots els estats es comuniquen pel fet que la cadena sigui irreductible, tenim que donat un $k \in I$, existeix un enter $n = n(k)$ tal que $p_{j_0 k}^{(n)} > 0$. Per tant, tenim

$$\omega_k = \omega_{j_0} p_{j_0 k}^{(n)} + \sum_{j \neq j_0} p_{jk}^{(n)} \omega_j > 0.$$

Així, doncs, hem vist que ω és una distribució invariant no degenerada per Π. Només, ens falta veure que si $v \in \mathbb{R}^M$ és solució de $x = x\Pi$, llavors v és múltiple de ω. Tenim, per a qualsevol $n \geq 0$,

$$v = \frac{1}{n+1} \sum_{l=0}^{n} v = \frac{1}{n+1} \sum_{l=0}^{n} v \Pi_l.$$

És a dir, per a cada $k \in I$,

$$v_k = \sum_{j \in I} v_j \left(\frac{1}{n+1} \sum_{l=0}^{n} p_{jk}^{(l)} \right).$$

Finalment, fent tendir n a infinit, obtenim

$$v_k = \sum_{j \in I} v_j \left(\lim_{n \to \infty} \frac{1}{n+1} \sum_{l=0}^{n} p_{jk}^{(l)} \right) = \sum_{j \in I} v_j \omega_k = \omega_k \left(\sum_{j \in I} v_j \right).$$

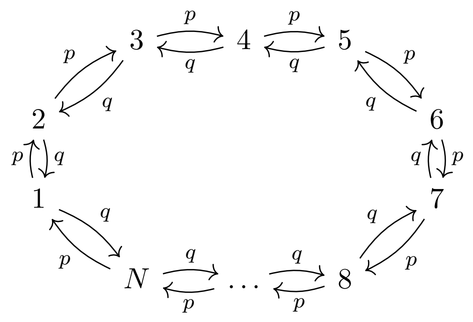
Així, $v_k = a\omega_k$, amb $a = \sum_{j \in I} v_j$. □

Tanquem l'apartat amb un exemple recollint el que s'ha estudiat.

Exemple. En una fira, un nen es troba en la situació que pot pujar infinitament als cavallets, però en cada viatge ha de canviar de cavallet, podent anar al de davant o al de darrere amb probabilitat $p \in (0, 1)$ i $q = 1 - p$, respectivament. Ens preguntem, quant temps, en mitjana, està a cada cavallet si n'hi ha N i s'hi està H hores?

Podem entendre la situació anterior com una passejada aleatòria finita, $I = \{1, \ldots, N\}$, i cíclica amb matriu de transició i diagrama següents:

$$\Pi = \begin{pmatrix} 0 & p & 0 & \ldots & 0 & 0 & q \\ q & 0 & p & \ldots & 0 & 0 & 0 \\ 0 & q & 0 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & p & 0 \\ 0 & 0 & 0 & \ldots & q & 0 & p \\ p & 0 & 0 & \ldots & 0 & q & 0 \end{pmatrix}$$



Notem que tant la suma dels elements de les files com la de les columnes és 1. Aquest tipus de matrius s'anomenen *biestocàstiques* i satisfan que si $N = |I|$, aleshores la distribució $\nu = (\nu_j)_{j\in I}$ amb $\nu_j = \frac{1}{N}$ és invariant. En efecte, com que hem dit que la suma dels elements de les columnes és 1, per a $j \in I$ tenim

$$\nu_j = \frac{1}{N} = \frac{1}{N}\sum_{i\in I} p_{ij} = \sum_{i\in I}\frac{1}{N}p_{ij} = \sum_{i\in I}\nu_i p_{ij}.$$

Destaquem, també, que tots els estats es comuniquen, per tant la cadena és irreductible. Així, doncs, podem aplicar el Teorema 7.2.1, obtenint que la distribució $\omega = (\frac{1}{m_{jj}})_{j\in I}$ és la única invariant per Π. Ara bé, havíem vist que ν també és distribució invariant, per tant

$$\omega_j = \frac{1}{m_{jj}} = \frac{1}{N} = \nu_j.$$

D'altra banda, en la secció **5.2** vèiem que tota cadena de Markov finita ha de tenir almenys un estat recurrent. A més, la recurrència és una propietat de classe, i en el nostre exemple estem tractant amb una cadena irreductible. Així, la nostra cadena és recurrent i podem aplicar-hi el Teorema 7.1.3. D'aquesta manera,

$$\omega_j = \frac{1}{N} = \frac{1}{m_{jj}} = \lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n} p_{ij}^{(k)} = \frac{1}{n+1}E(N_j^{(n)} \mid X_0 = i).$$

Recordem que $\frac{1}{n+1}E(N_j^{(n)} \mid X_0 = i)$ designa, en mitjana, el temps esperat d'ocupació de l'estat j. De manera que si el nen s'està H hores als cavallets, haurà passat, en mitjana, $\frac{1}{N}H$ hores en cada un. Per exemple, si hi ha 12 cavallets i s'hi està 3 hores, direm, que en mitjana, estarà 15 minuts en cada un.

7.3 Distribució límit i distribucions invariants

Com bé indica el títol d'aquesta secció, estudiarem la relació que hi ha entre la distribució límit d'una cadena de Markov homogènia i les distribucions invariants.

A la Proposició 6.2.2 provàvem que si tractem amb una cadena finita i regular (i consegüentment, ergòdica), aleshores la distribució límit de la cadena era l'única distribució invariant. Amb el següent teorema generalitzem aquest resultat.

Teorema 7.3.1. *Sigui $\{X_n : n \geq 0\}$ una $CMH(\gamma,\Pi)$ tal que per a tot $i, j \in I$, existeix $\nu = (\nu_j)_{j\in I}$ independent de i, complint*

$$\lim_{n\to\infty} p_{ij}^{(n)} = \nu_j.$$

Llavors, o bé, $\sum_{j\in I}\nu_j = 0$ i en tal cas no existeix cap distribució invariant, o bé, $\sum_{j\in I}\nu_j = 1$ i ν és l'única distribució invariant.

Demostració. Vegem, primer, que $\sum_{j\in I}\nu_j \in \{0,1\}$. Notem que se satisfà $\nu = \nu\Pi$. En efecte, cal tenir en compte que estem tractant amb valors reals, així, com que existeix el límit $\lim_{n\to\infty} p_{ki}^{(n)} = \nu_i$, aleshores coincideix amb el límit inferior. Aquest resultat es pot trobar a [2]. Dit això, usant, a més, el Lema de Fatou per Sumes[7] i l'equació de

[7]Lema de Fatou per Sumes: Sigui X un conjunt finit o numerable i $f_n : X \longrightarrow [0,\infty)$ una successió de funcions, aleshores

$$\sum_{x\in X}\liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty}\sum_{x\in X} f_n.$$

Chapman-Kolmogorov, obtenim

$$\sum_{i\in I}\nu_i p_{ij} = \sum_{i\in I}\lim_{n\to\infty} p_{ki}^{(n)} p_{ij} = \sum_{i\in I}\liminf_{n\to\infty} p_{ki}^{(n)} p_{ij} \le \liminf_{n\to\infty}\sum_{i\in I} p_{ki}^{(n)} p_{ij} = \liminf_{n\to\infty} p_{kj}^{(n+1)} = \nu_j.$$

Volem veure que és una igualtat, per això suposem que existeix algun $j_0 \in I$ pel qual $\sum_{i\in I}\nu_i p_{ij_0} < \nu_{j_0}$. Tenim

$$\sum_{j\in I}\nu_j > \sum_{j\in I}\sum_{i\in I}\nu_i p_{ij} = \sum_{i\in I}\nu_i\left(\sum_{j\in I} p_{ij}\right) = \sum_{i\in I}\nu_i.$$

Per tant, hem arribat a contradicció i tenim $\sum_{i\in I}\nu_i p_{ij} = \nu_j$. Ara, fent un raonament similar a l'anterior, en resulta

$$0 \le \sum_{i\in I}\nu_i = \sum_{i\in I}\lim_{n\to\infty} p_{ki}^{(n)} = \sum_{i\in I}\liminf_{n\to\infty} p_{ki}^{(n)} \le \liminf_{n\to\infty}\sum_{i\in I} p_{ki}^{(n)} = 1.$$

Fins ara, hem vist que $\sum_{j\in I}\nu_j \in [0,1]$ i que se satisfà $\nu = \nu\Pi_n$. Per tant, en particular, $|\nu_i p_{ij}^{(n)}| \le \nu_j$ i $\sum_{j\in I}\nu_j \le 1 < \infty$. Així, podem fer servir el Teorema de Convergència Dominada per Sumes[8] amb funció dominant $g = \nu_j$, obtenint

$$\nu_j = \lim_{n\to\infty}\nu_j = \lim_{n\to\infty}\sum_{i\in I}\nu_i p_{ij}^{(n)} = \sum_{i\in I}\nu_i \lim_{n\to\infty} p_{ij}^{(n)} = \left(\sum_{i\in I}\nu_i\right)\nu_j.$$

Per tant, $\sum_{i\in I}\nu_i \in \{0,1\}$. Ara, ens interessa estudiar aquests dos casos.

Suposem que $v = (v_j)_{j\in I}$ és una distribució invariant. Per a qualsevol $j \in I$, se satisfà

$$v_j = \lim_{n\to\infty} v_j = \lim_{n\to\infty}\sum_{i\in I} v_i p_{ij}^{(n)} = \sum_{i\in I} v_i \lim_{n\to\infty} p_{ij}^{(n)} = \left(\sum_{i\in I} v_i\right)\nu_j = \nu_j.$$

On hem usat el Teorema de Convergència Dominada per Sumes per poder fer l'intercanvi entre sumatori i límit.

D'una banda, si $\sum_{i\in I}\nu_i = 0$, com que $v = \nu$, tenim que $\sum_{i\in I} v_i = 0$, que contradiu el fet que v és una distribució. D'altra banda, si $\sum_{i\in I}\nu_i = 1$, com que $\nu = \nu\Pi$, llavors ν és una distribució invariant, que és única, ja que havíem considerat v una distribució estacionària qualsevol. $\square$

Abans hem dit que aquest resultat generalitzava la Proposició 6.2.2. Efectivament, si suposem que I és finit i que per a cada $i,j \in I$ existeix $\nu_j = \lim_{n\to\infty} p_{ij}^{(n)}$ independent de i, aleshores

$$\sum_{j\in I}\nu_j = \sum_{j\in I}\lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty}\sum_{j\in I} p_{ij}^{(n)} = 1.$$

D'aquesta manera, aplicant el teorema que acabem de demostrar, tenim que ν és la distribució límit de la cadena i és l'única que és invariant.

Cal destacar que a diferència del que havíem vist en la Proposició 6.2.2, fent servir aquest nou resultat no estem demanant que ν_j sigui estrictament positiva, sinó que és suficient que existeixi i sigui independent de i. Així mateix, aquest teorema és vàlid per cadenes infinites.

[8] Teorema de Convergència Dominada per Sumes: Sigui $f_n : X \longrightarrow \mathbb{C}$ una successió de funcions, tal que per $x \in X$ i $n \ge 0$, existeix el límit $\lim_{n\to\infty} f_n(x)$, i una funció dominant $g : X \longrightarrow [0,\infty)$ tal que $|f_n(x)| \le g(x)$ i $\sum_{x\in X}|g(x)| < \infty$. Aleshores,

$$\lim_{n\to\infty}\sum_{x\in X} f_n(x) = \sum_{x\in X}\lim_{n\to\infty} f_n(x).$$

8 Conclusions

Com es va dir a la introducció, el propòsit del projecte era desenvolupar la teoria de les cadenes de Markov homogènies. Una vegada finalitzat el treball, considero que era un objectiu bastant ambiciós. Tot i tractar exclusivament amb un tipus de cadenes de Markov, i algunes vegades, fins i tot, restringir l'estudi al cas finit, hi ha molts temes que no s'han pogut explorar i documentar, ja que se sobrepassava el que s'espera d'aquest treball. Així mateix, s'hagués pogut estructurar el projecte de manera que es toqués superficialment cada tema, però això hagués implicat no entrar en detall en alguns altres aspectes, fent, possiblement, que es perdi el fil del que s'està estudiant o fer un estudi incomplet.

Cal recordar, també, que en tot moment treballàvem únicament amb processos de variable temporal discreta. Amb tot això, és evident que l'estudi de cadenes (o processos) de Markov és molt extens i el fet que tingui tantes aplicacions en altres ciències, en destaca la seva importància.

M'hagués agradat poder donar més exemples que ens puguem trobar a la vida quotidiana, però degut a la complexitat d'aquests fenòmens aleatoris, m'ha estat difícil i he hagut d'optar per presentar exemples didàctics per complementar la part teòrica.

Durant la memòria, a partir del que ja coneixíem de probabilitats, hem pogut construir i definir el que entenem per cadena de Markov homogènia. Pas a pas, hem edificat una sòlida estructura amb la qual treballar i poder provar grans resultats. Hem estat capaços d'estudiar les propietats dels estats, podent-los classificar segons la comunicació, i posteriorment, segons l'evolució temporal de la cadena; en estats recurrents i transitoris. Un dels grans estudis que s'ha fet és veure'n el comportament a mida que avançàvem en el temps; arribant a la conclusió que per cadenes finites i regulars, la cadena *s'equilibra*, de manera que la distribució límit coincideix amb la única distribució invariant que existeix per la matriu de transició de la cadena.

Pel que fa a l'àmbit personal, a més de presentar-me un nou món el qual desconeixia com és el de les cadenes de Markov, el que destacaria és la millora en la capacitat de formalitzar i descriure amb rigorositat i claredat les idees exposades, que durant els anys de grau no s'havia pogut aprofundir.

Referències

[1] J. R. Norris. *Markov Chains.* Cambridge University Press, Cambridge, 1997.

[2] B. K. Driver. *Analysis Tools with Applications.* Springer, Berlin, 2003.

[3] M. Sanz. *Probabilitats.* Edicions Universitat de Barcelona, Barcelona, 1999.

[4] M. Iosifescu. *Finite Markov Processes and Their Applications.* John Wiley & Sons, New York, 1980.

[5] M. Kijima. *Markov Processes for Stochastic Modeling.* Chapman & Hall, London, 1997.

[6] P. Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues.* Springer, New York, 1999.