Universitat de Barcelona
Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Introducción de las Cadenas de Márkov al Sistema Educativo

Autor: Aleix Rodriguez Zaragoza

Director: Dr. David Márquez Carreras
Realitzado en: Departamento de Matemáticas e Informática.

Barcelona, 11 de junio de 2022

Abstract

Markov chains are discrete stochastic processes without memory. The main goal of this project is to study them in order to achieve a university elective proposal focused on them. As the general topic is so extensive, the main focus will be on the study of discrete-time strings, more specifically on homogeneous strings with discrete time variables and their classification.

Resum

Les cadenes de Màrkov són processos estocàstics discrets sense memòria. L'objectiu principal d'aquest projecte és estudiar-les per aconseguir una proposta d'optativa universitària centrada en aquestes. Com que el tema general és molt extens, el treball se centra en l'estudi de les cadenes a temps discret, més concretament en les cadenes homogènies de variable temporal discreta i la seva classificació.

Resumen

Las cadenas de Márkov son procesos estocásticos discretos sin memoria. El objetivo principal de este proyecto es estudiarlas para conseguir una propuesta de optativa universitaria centrada en estas. Como el tema general es muy extenso, el trabajo se centra en el estudio de las cadenas a tiempo discreto, más concretamente en las cadenas homogéneas de variable temporal discreta y su clasificación.

2020 Mathematics Subject Classification. 11G05, 11G10, 14G10

Agradecimientos

Me gustaría agradecer a mi tutor, David Márquez, por su apoyo en todo el proceso de realización de este proyecto.

También quiero aprovechar esta sección para agradecer a toda la gente que me ha apoyado, familiares, amigos y compañeros, ya que sin ellos tampoco hubiera sido posible llegar hasta aquí

Índice

1. Introducción

En la teoría de la probabilidad, se denomina cadena de Márkov a un tipo especial de proceso estocástico en el que la probabilidad de que ocurra un evento depende únicamente del evento inmediatamente anterior.

El objetivo principal de este trabajo es crear una propuesta de optativa universitaria centrada en las cadenas de Márkov. Como estas dan para un estudio mucho más extenso, el proyecto se centra en el estudio de las cadenas con tiempo discreto y en intentar clasificarlas. La clasificación se hará mediante un programa que, en la propuesta de asignatura, crearían los propios alumnos. El trabajo se estructura de la siguiente forma:

1. Se hará un repaso de algunos conceptos de probabilidad que serán útiles para el estudio de las cadenas.
2. Se dará la definición de cadena de Márkov y que significa que esta sea homogénea. Posteriormente, se estudiarán las probabilidades fundamentales de estas.
3. Se explicará como se clasifican los diferentes estados de las cadenas y algunas propiedades de estos.
4. Se verán las distintas clasificaciones de las cadenas y su comportamiento a largo plazo para llegar a obtener un método programable de clasificación.
5. Estudiaremos en profundidad conceptos previos vistos por encima e introduciremos de nuevos como, por ejemplo, la estadística de las cadenas.
6. Se creará un programa en lenguaje VBA que clasifique automáticamente la cadena cuando se le otorgue la matriz de transición de la misma.

El objetivo de la asignatura propuesta es que los alumnos entiendan lo que es una cadena de Márkov, vean sus propiedades principales y aprendan a clasificarlas. El programa se ha añadido al temario porque es una buena manera de que los alumnos relacionen las matrices de transición y la clasificación de las cadenas a la vez que demuestran que han entendido realmente como clasificarlas.

El proyecto se estructurará directamente como si fuera el temario de una optativa universitaria. En ocasiones también se añadirán comentarios extra o explicaciones que no se darían en un temario real. Por ejemplo, en este primer tema introductorio se dan conceptos que un alumno del Grado de Matemáticas ya ha visto en asignaturas previas, por tanto, no es necesario explicarlo en esta optativa, aunque si sería idóneo al menos compartirlo como documento a los alumnos por si quisieran repasar rápidamente los conceptos.

También, en la sección de programación se da una breve explicación de como se ha realizado el programa, mientras que a los alumnos en su lugar se les daría clases en las que se les ayudaría a realizar su propio programa.

2. Definiciones previas

Se tomará un repaso de la teoría de probabilidades que será útil durante el proyecto. Empezaremos introduciendo conceptos básicos.

Definición 2.1. *Sea Ω un conjunto y $\mathcal{A} \subseteq \mathcal{P}(\Omega)$. Se dice que $\mathcal{A}$ es una σ-álgebra de $\mathcal{P}(\Omega)$ si satisface:*

1. $\Omega \in \mathcal{A}$.
2. $A \in \mathcal{A} \Longrightarrow A^c \in \mathcal{A}$.
3. *Si $\{A_n : n \geq 1\}$ es una familia numerable de elementos de $\mathcal{A}$, entonces* $\bigcup\limits_{n\geq 1} A_n \in \mathcal{A}$.

Definición 2.2. *Sea Ω un conjunto y $\mathcal{F}$ una σ-álgebra sobre $\mathcal{P}(\Omega)$. Entonces el conjunto $(\Omega, \mathcal{F})$ es un espacio medible.*

Definición 2.3. *Un espacio de probabilidad es una terna $(\Omega, \mathcal{A}, P)$ tal que:*

- *Ω es el espacio muestral, un conjunto que contiene los resultados posibles.*
- *$\mathcal{A} \subseteq \mathcal{P}(\Omega)$ es una σ-álgebra que nos permite describir todos los sucesos posibles.*
- *P es la probabilidad, una aplicación P: $\mathcal{A} \longrightarrow [0,1]$ tal que:*
  1. $P(\Omega) = 1$.
  2. *Si $\{A_n : n \geq 1\}$ es una sucesión de conjuntos de $\mathcal{A}$ disjuntos por pares, entonces:* $P(\bigcup\limits_{n=1}^{\infty} A_n) = \sum\limits_{n=1}^{\infty} P(A_n)$.

Observación 2.4. Si Ω es finito, entonces a $(\Omega, \mathcal{A}, P)$ se le llama espacio de probabilidad finito.

Definición 2.5. *Dos sucesos A y B son independientes si $P(A \cap B) = P(A)P(B)$.*

Observación 2.6. Normalmente se usa P(A,B) en lugar de $P(A \cap B)$.

Definición 2.7. *Una familia de sucesos $\{A_i : i \in I\}$, es independiente si para toda combinación finita $A_{i_1}, ..., A_{i_k}$ de conjuntos diferentes se tiene que:*

$$P(\bigcap_{j=1}^{k} A_{i_j}) = \prod_{j=1}^{k} P(A_{i_j}).$$

Definición 2.8. *La probabilidad condicionada del suceso A dado el suceso B es la probabilidad de que suceda A sabiendo que sucedió B. Para $P(B) > 0$ se define como:*

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

Observación 2.9. De la definición de probabilidad condicionada se decuce que:

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A).$$

Siempre que $P(A) > 0$ y $P(B) > 0$.

A continuación veremos diversos teoremas que serán útiles durante el proyecto, sobre todo a la hora de las demostraciones.

Teorema 2.10. *(Teorema de las Probabilidades totales) Sea $\{B_i : i \in I\}$ una partición de Ω tal que $P(B_i) > 0$ para todo $i \in I$ entonces para todo $A \in \mathcal{A}$:*

$$P(A) = \sum_{i \in I} P(A|B_i)P(B_i).$$

Teorema 2.11. *(Regla de Bayes) Sea $\{B_i : i \in I\}$ una partición de Ω tal que $P(B_i) > 0$ para todo $i \in I$ entonces sea $A \in \mathcal{A}$ con $P(A) > 0$:*

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum\limits_{i \in I} P(A|B_i)P(B_i)}.$$

Teorema 2.12. *(Teorema de las Probabilidades compuestas) Sean $A_1, A_2, ..., A_n$ sucesos con $P(A_1 \cap A_2 \cap ... \cap A_{n-1}) > 0$ entonces:*

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

Ahora veremos lo que es una variable aleatoria y definiremos su distribución y su esperanza.

Definición 2.13. *Dados dos espacios medibles $(\Omega, \mathcal{A}), (F, \mathcal{F})$, decimos que una aplicación $X : \Omega \longrightarrow F$ es medible si $X^{-1}(B) \in \mathcal{A}$ para todo $B \in \mathcal{F}$.*

Definición 2.14. *Una variable aleatoria es una aplicación $X : (\Omega, \mathcal{A}) \longrightarrow (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ medible, donde $\mathscr{B}(\mathbb{R})$ significa los borelianos de $\mathbb{R}$.*

A continuación se da una definición equivalente más sencilla de variable aleatoria.

Definición 2.15. *Una aplicación $X : \Omega \longrightarrow \mathbb{R}$ es una variable aleatoria si $\forall x \in \mathbb{R}$ el suceso $\{X \leqslant x\} = \{\omega \in \Omega : X(\omega) \leqslant x\}$ tiene una probabilidad asignada lo que equivale a que $\{X \leqslant x\} \in \mathcal{A}$.*

Si la aplicación va a E , un conjunto numerable, entonces $X : \Omega \longrightarrow E$ es una variable aleatoria discreta si $\forall a \in E$ se verifica que $\{X = a\} \in \mathcal{A}$

Definición 2.16. *Sea X una variable aleatoria. La función de distribución de X es:*

$$\begin{aligned} F_X : \mathbb{R} &\longrightarrow [0, 1] \\ x &\longmapsto F_X(x) = P(X \leq x). \end{aligned}$$

Definición 2.17. *Sea X una variable aleatoria discreta con valores en E verificando que $\sum\limits_{x \in E} |x|P(X = x) < \infty$. Se define la esperanza de X como:*

$$E[X] = \sum_{x \in E} x \cdot P(X = x).$$

Definición 2.18. *Sea X una variable aleatoria discreta con valores en E y sea $g : E \longrightarrow \mathbb{R}$ verificando que $\sum\limits_{x \in E} |g(x)|P(X = x) < \infty$. Se define la esperanza de g(X) de la siguiente forma:*

$$E[g(X)] = \sum_{x \in E} \sum g(x)P(X = x).$$

3. Cadenas de Márkov

Para poder definir formalmente el concepto de Cadena de Márkov aún son necesarios algunos conceptos más:

Definición 3.1. *Un proceso estocástico es una colección de variables aleatorias indexadas:*

$$X_t : \Omega \longrightarrow E.$$

Se llaman estados a los posibles valores que toman las variables aleatorias y conjunto de estados a E.

Observación 3.2. El índice $t \in T$ normalmente se identifica con el tiempo. Se habla de proceso estocástico discreto cuando $T \subseteq \mathbb{N}$ y es continuo cuando $T \subseteq [0, \infty]$.

Este curso se centrará en procesos de tiempo discreto. Debido a esto muchas veces se usará el subíndice n en lugar de t.

Ejemplo 3.3. Para $X_t = i, i \in E$ se dice que el proceso esta en el estado i en el instante t.

Definición 3.4. *Una matriz $\mathcal{M} = (p_{ij})_{i,j \in E}$, con E conjunto finito o numerable, es una matriz estocástica si verifica que:*

1. $p_{ij} \in [0, 1]$.
2. $\sum\limits_{j \in E} p_{ij} = 1$.

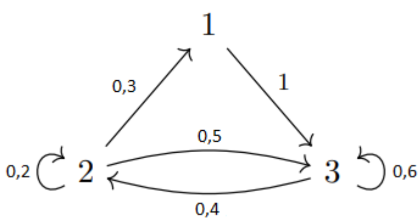
Observación 3.5. Cada fila $(p_{ij})_{j \in E}$ es una probabilidad. Como E puede ser infinito la matriz $\mathcal{M}$ tambien puede tener dimensión infinita.

Existe una correspondencia biyectiva entre las matrices estocásticas y los diagramas generados identificando los estados con vértices y cada probabilidad no nula con arestas.

Ejemplo 3.6. La matriz estocástica:

$$\begin{pmatrix} 0 & 0 & 1 \\ 0,3 & 0,2 & 0,5 \\ 0 & 0,4 & 0,6 \end{pmatrix}$$

se corresponde con el siguiente diagrama:



Gracias a toda la teoría introducida previamente ya se puede obtener una definición formal de las cadenas de Márkov.

Definición 3.7. *Un proceso estocástico* $\{X_n : n \geq 0\}$ *que toma valores en un conjunto de estados* E *es una cadena de Márkov si para todo* $i_0, ..., i_{n+1} \in E$ *con* $n \geq 0$ *se verifica la propiedad de Márkov:*

$$P(X_{n+1} = i_{n+1} \mid X_n = i_n, ..., X_0 = i_0) = P(X_{n+1} = i_{n+1} \mid X_n = i_n). \tag{3.1}$$

Observación 3.8. La propiedad de Márkov es equivalente a que la medición del comportamiento futuro del sistema depende únicamente del estado presente.

3.1. Cadenas de Márkov Homogéneas

Definición 3.9. *Una cadena de Márkov es homogénea si además verifica que* $P(X_{n+1} = i_{n+1} \mid X_n = i_n)$ *no depende de* n.

Definición 3.10. *La matriz estocástica asociada a la cadena,* $\mathcal{M} = (p_{ij})_{i,j\in E}$*, es nombrada matriz de transición.*

Una vez introducidas las cadenas de Márkov homogéneas, el siguiente objetivo es caracterizarlas.

Teorema 3.11. *Sea* $\{X_n : n \geq 0\}$ *un proceso estocástico con valores en* E*. Decimos que es una cadena de Márkov homogénea con distribución de probabilidad inicial* γ *y matriz de transición* $\mathcal{M}$ *si, y solo si, para todo* $i_0, ...i_n \in E$ *y* $\forall n \geq 0$*:*

$$P(X_0 = i_0, X_1 = i_1, ..., X_n = i_n) = \gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}. \tag{3.2}$$

Demostración. Se demostrarán ambas implicaciones por separado:

1. Suponiendo que tenemos una cadena de Márkov homogénea, $\{X_n : n \geq 0\}$, con distribución de probabilidad inicial γ y matriz de transición $\mathcal{M}$. Usando el teorema de las probabilidades compuestas (Teorema 2.12) y posteriormente aplicando la propiedad de Márkov (2.1):

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

2. Sea $\{X_n : n \geq 0\}$ un proceso estocástico verificando la igualdad (2.2).
   Para todo $i \in E$ se verifica que $P(X_0 = i) = \gamma_i$, lo que implica que la funcion de distribución de X_0 es γ. Por tanto solo falta ver que $\{X_n : n \geq 0\}$ es una cadena de Márkov homogénea.

   Usando la definición de probabilidad condicionada:

$$\begin{aligned}
P(X_n = i_n|X_{n-1} = i_{n-1}, ..., X_0 = i_0) &= \frac{P(X_0 = i_0, X_1 = i_1, ..., X_n = i_n)}{P(X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1})} \\
&= \frac{\gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}}{\gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{n-2} i_{n-1}}} \\
&= p_{i_{n-1} i_n}.
\end{aligned}$$

A su vez, usando el teorema de las probabilidades totales (Teorema 2.10):

$$\begin{aligned}
P(X_{n-1} = i_{n-1}) &= \sum_{i_0,\dots,i_{n-2}\in E^{n-1}} P(X_{n-1} = i_{n-1}|X_0 = i_0, \dots, X_{n-2} = i_{n-2})\times \\
&\quad \times P(X_{n-1} = i_{n-1}|X_0 = i_0, \dots, X_{n-2} = i_{n-2}) \\
&= \sum_{i_0,\dots,i_{n-2}\in E^{n-1}} P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}) \\
&= \sum_{i_0,\dots,i_{n-2}\in E^{n-1}} \gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{n-2} i_{n-1}}.
\end{aligned}$$

También de forma similar:

$$\begin{aligned}
P(X_{n-1} = i_{n-1}, X_n = i_n) &= \sum_{i_0,\dots,i_{n-2}\in E^{n-1}} P(X_0 = i_0, \dots, X_n = i_n) \\
&= \sum_{i_0,\dots,i_{n-2}\in E^{n-1}} \gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n} \\
&= p_{i_{n-1} i_n} \sum_{i_0,\dots,i_{n-2}\in E^{n-1}} \gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{n-2} i_{n-1}}.
\end{aligned}$$

Entonces:

$$P(X_n = i_n|X_{n-1} = i_{n-1}) = \frac{P(X_{n-1} = i_{n-1}, X_n = i_n)}{P(X_{n-1} = i_{n-1})} = p_{i_{n-1} i_n}.$$

De donde se deduce que $\{X_n : n \geq 0\}$ es una cadena de Márkov homogénea ya que:

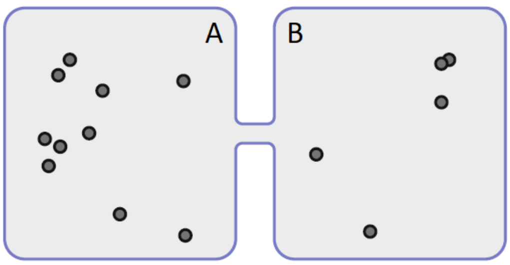
$$P(X_n = i_n|X_0 = i_0, \dots X_{n-1} = i_{n-1}) = P(X_n = i_n|X_{n-1} = i_{n-1}) = p_{i_{n-1} i_n}.$$

$\square$

Notación 3.12. *La cadena de Márkov homogénea con matriz de transición $\mathcal{M}$ y distribución de probabilidades inicial γ la denotaremos $CMH(\gamma, \mathcal{M})$.*

Veamos ejemplos de cadenas de Márkov homogéneas.

Ejemplo 3.13. Cadena de Ehrenfest.
Supongamos que tenemos dos urnas conectadas por un tubo conteniendo un total de N bolas. En cada instante de tiempo n una bola pasa de un volumen al otro aleatoriamente.



Entonces llamaremos Z_n al numero de bolas en la urna A a tiempo n. Las probabilidades $P(Z_{n+1} = j|Z_n = i)$, $0 \leq i, j \leq N$, vienen dadas por:

$$P(Z_{n+1} = k+1|Z_n = k) = \frac{N-k}{N}, \;\; k = 0, ..., N-1.$$
$$P(Z_{n+1} = k-1|Z_n = k) = \frac{k}{N}, \;\; k = 1, ..., N.$$

Por tanto, su matriz de transición será:

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ \frac{1}{n} & 0 & \frac{N-1}{N} & 0 & \cdots & 0 \\ 0 & \frac{2}{n} & 0 & \frac{N-2}{N} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{N-1}{N} & 0 & \frac{1}{n} \\ 0 & \cdots & \cdots & 0 & 1 & 0 \end{pmatrix}$$

Ejemplo 3.14. Paseo aleatorio en $E = \mathbb{Z}^d$.
Sea $\{Y_n : n \geq 1\}$ una sucesión de variables aleatorias con valores en Z^d y distribución común λ. Sea X_0 una variable aleatoria con valores en $\mathbb{Z}^d$, independiente de Y_n. Entonces la secuencia aleatoria $\{X_n : n \geq 0\}$ definida por

$$X_{n+1} = X_n + Y_{n+1}, n \in \mathbb{N},$$

es una $CMH(\mu, P)$, con μ la ley de X_0 y $P_{xy} = \lambda_{y-x}$. Un ejemplo clásico es el del paseo aleatorio simétrico empezando en el 0. Con nuestra notación este es:

$$\mu = \delta_0, \;\; \lambda_{\pm e_i} = \frac{1}{2d},$$

donde $(e_1, ..., e_d)$ es la base ortonormal de $\mathbb{R}^d$.

Comprovemos que $\{X_n : n \geq 0\}$ verifica las condiciones para ser una cadena de Márkov homogénea. Usando que $\{Y_n : n \geq 1\}$ son variables aleatorias independientes y identicamente distribuidas, $\forall n$ y sean $i_0, ..., i_{n-1}, i, j \in E$, podemos ver que:

$$\begin{aligned} P(X_{n+1} = j \mid X_n = i, &X_{n-1} = i_{n-1}, ..., X_0 = i_0) = \\ &= \frac{P(X_{n+1} = j, X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0)}{P(X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0)} \\ &= \frac{P(Y_{n+1} = j-i, Y_n = i - i_{n-1}, ..., Y_1 = i_1 - i_0, X_0 = i_0)}{P(Y_n = i - i_{n-1}, ..., Y_1 = i_1 - i_0, X_0 = i_0)} \\ &= \frac{P(Y_{n+1} = j-i)P(Y_n = i - i_{n-1}, ..., Y_1 = i_1 - i_0, X_0 = i_0)}{P(Y_n = i - i_{n-1}, ..., Y_1 = i_1 - i_0, X_0 = i_0)} \\ &= P(Y_{n+1} = j-i). \end{aligned}$$

También que:

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

Por tanto, se cumple la propiedad de Márkov y, como $P(Y_{n+1} = j - i) := p_{ij}$ no depende de n, tenemos que $\{X_n : n \geq 0\}$ es una cadena de Márkov homogénea.

A continuación se presentan 2 proposiciones de casos concretos en los que se pueden obtener CMH siempre que se verifiquen las condiciones.

Proposición 3.15. *Sea $\{X_n : n \geq 0\}$ una $CMH(\gamma, \mathcal{M})$ entonces $\{X_{n+m} : m \geq 0\}$ es una $CMH(\mathscr{L}(X_m), \mathcal{M})$, siendo $\mathscr{L}(X_m)$ la función de probabilidad de X_m.*

Demostración. Usando el teorema de las probabilidades totales (Teorema 2.10) y aplicando el Teorema 3.11 a $CMH(\gamma, \mathcal{M})$ se obtiene que para todo $m \geq 0$:

$$
\begin{aligned}
P(X_{0+m} = j_0, ... X_{n+m} = j_n) = & \\
&= \sum_{i_0,...,i_{m-1} \in E^m} P(X_0 = i_0, ..., X_{m-1} = i_{m-1}, X_m = j_0, ..., X_{m+n} = j_n) \\
&= \sum_{i_0,...,i_{m-1} \in E^m} \gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{m-2} i_{m-1}} p_{i_{m-1} j_0} p_{j_0 j_1} \cdots p_{j_{n-1} j_n} \\
&= p_{j_0 j_1} \cdots p_{j_{n-1} j_n} \sum_{i_0,...,i_{m-1} \in E^m} \gamma_{i_0} p_{i_0 i_1} \cdots p_{i_{m-2} i_{m-1}} p_{i_{m-1} j_0} \\
&= P(X_m = j_0) p_{j_0 j_1} \cdots p_{j_{n-1} j_n}.
\end{aligned}
$$

□

En la siguiente proposición estudiaremos un caso particular que puede ser útil para construir cadenas de Márkov homogéneas.

Proposición 3.16. *Sea $\{Z_n : n \geq 1\}$ una sucesión de variables aleatorias independientes idénticamente distribuidas, $Z_n : \Omega \longrightarrow E$ y $f : E \times E \longrightarrow E$. Sea X_0 una variable aleatoria con valores en E independiente de $\{Z_n : n \geq 1\}$, entonces definiendo $X_{n+1} = f(X_n, Z_{n+1})$ para $n \geq 0$ se tiene que $\{X_n : n \geq 0\}$ es una cadena de Márkov homogénea.*

Demostración. Viendo que:

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

y que su valor no depende de n queda demostrado que $\{X_n : n \geq 0\}$ es una cadena de Márkov homogénea bajo las condiciones de la proposición. Entonces veámoslo:

1. Utilizando la independencia:

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

2. $\{X_0 = i_0, ..., X_n = i_n\}$ se puede expresar en términos de $X_0, Z_1, ..., Z_n$ por lo que es independiente de Z_{n+1}. Usando esto:

$$
\begin{aligned}
P(X_{n+1} = j | X_0 = i_0, ..., X_{n-1} = i_{n-1}, X_n = i) = \\
&= \frac{P(X_0 = i_0, ..., X_{n-1} = i_{n-1}, X_n = i, X_{n+1} = j)}{P(X_0 = i_0, ..., X_{n-1} = i_{n-1}, X_n = i)} \\
&= \frac{P(X_0 = i_0, f(i_0, Z_1) = i_1, ..., f(i_{n-1}, Z_n) = i, f(i, Z_{n+1}) = j)}{P(X_0 = i_0, f(i_0, Z_1) = i_1, ..., f(i_{n-1}, Z_n) = i)} \\
&= \frac{P(f(i, Z_{n+1}) = j) P(X_0 = i_0, f(i_0, Z_1) = i_1, ..., f(i_{n-1}, Z_n) = i)}{P(X_0 = i_0, f(i_0, Z_1) = i_1, ..., f(i_{n-1}, Z_n) = i)} \\
&= P(f(i, Z_{n+1}) = j) = p_{ij}
\end{aligned}
$$

Claramente son iguales e independientes de n.

□

3.2. Probabilidades fundamentales

El siguiente objetivo es estudiar diversas probabilidades, referentes a las cadenas de Márkov, que serán de gran utilidad para el estudio de las mismas.

3.2.1. Probabilidad de transición

Definición 3.17. *Sea $\{X_n; n \geq 0\}$ una cadena de Márkov homogénea que se encuentra en el estado i. Definimos la* ***probabilidad de transición*** *como la probabilidad de que m pasos después se encuentre en el estado j. La expresión formal es:*

$$p_{ij}^{(m)} := P(X_{n+m} = j | X_n = i) \tag{3.3}$$

con $i, j \in E$ y n,m enteros positivos. Esta sección se centrará en encontrar dicha probabilidad.

Observación 3.18. El caso $m = 1$ es el que ya se ha estudiado: $p_{ij}^{(1)} = p_{ij}$.

Pasemos a ver algunos procesos recursivos que facilitan el cálculo de estas probabilidades.

Proposición 3.19. *Se pueden obtener las probabilidades m-ésimas $p_{ij}^{(m)}$ con $m \geq 2$ recursivamente:*

$$p_{ij}^{(m)} = \sum_{k \in E} p_{ik}^{(m-1)} p_{kj}. \tag{3.4}$$

Demostración. Usando la definición de probabilidad condicionada (Definición 2.8) y el teorema de las probabilidades totales (Teorema 2.10):

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

□

Se define la matriz $\mathcal{M}_m := (p_{ij}^{(m)})_{i,j \in E}$. Entonces:

$$\mathcal{M}_m = \mathcal{M}^m = \mathcal{M} \times \underset{(m)}{\dots} \times \mathcal{M}.$$

Usando la Proposición 3.19 se obtienen las siguientes propiedades:

1. $\mathcal{M}_m$ es una matriz estocástica.
2. La siquiente equiación se llama **Equación de Chapman-Kolmogorov:**

$$p_{ij}^{(l+k)} = \sum_{k \in E} p_{ik}^{(l)} p_{kj}^{(k)}. \tag{3.5}$$

   Se ha utilizado la propiedad del producto de matrices: $\mathcal{M}^{l+k} = \mathcal{M}^l \mathcal{M}^k$.

3. Iterando la equación de Champman-Kolmogorov:

$$p_{ij}^{(m)} = \sum_{i_1 \in E} p_{ii_1} p_{i_1 j}^{(m-1)} = \sum_{i_1, \dots, i_{m-1} \in E} p_{ii_1} p_{i_1 i_2}, \dots, p_{i_{m-1} j}. \tag{3.6}$$

El siguiente paso es determinar la ley del proceso de una cadena de Márkov homogénea con distribución de probabilidad inicial γ y matriz de transición $\mathcal{M}$.

Proposición 3.20. *Sea* $\{X_n; n \geq 0\}$ *una* $CMH(\gamma, \mathcal{M})$. *Entonces* $\forall k \in E$*:*

$$P(X_n = k) = \gamma_k^{(n)},$$

donde $\gamma^{(n)} = \gamma \mathcal{M}^n$.

Demostración. Usando el teorema de las probabilidades totales (Teorema 2.10):

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

□

Corolario 3.21. *Sea* $\{X_n; n \geq 0\}$ *una* $CMH(\gamma, \mathcal{M})$ *se verifica:*

1. $\gamma^{(l+k)} = \gamma^{(l)}\mathcal{M}^k$.
2. $\gamma_j^{(n)} = P(X_n = j) = \sum_{h\in E} \gamma_h p_{hj}^{(n)} = \sum_{h\in E}(\sum_{i_1,\ldots,i_{n-1}\in E} \gamma_h p_{hi_1}p_{i_1i_2}\cdots p_{i_{n-1}j})$.

Demostración. Volviendo a usar el teorema de las probabilidades totales:

1. $\gamma_i^{(l+k)} = P(X_{l+k} = i) = \sum_{j\in E} P(X_{l+k} = i|X_l = j)P(X_l = j) = \sum_{j\in E} \gamma_j^{(l)} p_{ji}^{(k)}$.

   Por lo tanto $\gamma^{(l+k)} = \gamma^{(l)}\mathcal{M}^k$.

2. Usando la Proposición 3.20:

   $$\gamma_j^{(n)} = P(X_n = j) = \sum_{h\in E} \gamma_h p_{hj}^{(n)}.$$

   Y aplicando la igualdad (2.6):

   $$\sum_{h\in E} \gamma_h p_{hj}^{(n)} = \sum_{h\in E}(\sum_{i_1,\ldots,i_{n-1}\in E} \gamma_h p_{hi_1}p_{i_1i_2}\cdots p_{i_{n-1}j}).$$

□

Proposición 3.22. *La distribución inicial* γ *y la matriz de probabilidades de tranción* $\mathcal{M}$ *determinan la ley del vector aleatorio* $(X_{n_0}, ..., X_{n_k})$, $0 \leq n_1 < n_2 < ... < n_k$.

Demostración. Se demostrará primero que se verifica para $n_0 = 0, ..., n_k = k$:

$$\begin{aligned} P(X_0 = i_0, ..., X_k = i_k) &= P(X_0 = i_0)P(X_1 = i_1|X_0 = i_0)P(X_2 = i_2|X_0 = i_0, X_1 = i_1)\times \\ &\quad ... \times P(X_k = i_k|X_0 = i_0, ..., X_{k-1} = i_{k-1}) \\ &= \gamma_{i_0}p_{i_0i_1}\cdots p_{i_{k-1}i_k}. \end{aligned}$$

Utilizándolo a continuación:

$$\begin{aligned} P(X_{n_0} = i_0, ..., X_{n_k} = i_k) &= P(X_{n_0} = i_0)P(X_{n_1} = i_1|X_{n_0} = i_0) \times ... \\ &\quad \times P(X_{n_k} = i_k|X_{n_0} = i_0, ..., X_{n_{k-1}} = i_{k-1}) \\ &= \sum_{h\in E} \gamma_h p_{hi_0}^{(n_0)} p_{i_0i_1}^{(n_1-n_0)} \cdots p_{i_{k-1}i_k}^{(n_k-n_{k-1})}. \end{aligned}$$

□

3.2.2. Probabilidad de pasada

Definición 3.23. *El tiempo de primer paso al estado j se define como:*

$$T_j := \text{ínf}\{n \geq 1 : X_n = j\},$$

con

$$T_j := \infty \text{ si } X_n \neq j \text{ para toda } n \geq 1.$$

Inductivamente, definimos el tiempo de r-ésimo paso al estado j como:

$$\begin{aligned}
T_j^{(0)} &= 0. \\
T_j^{(1)} &= T_j \\
T_j^{(r)} &= \text{ínf}\{n \geq T_j^{(r-1)} + 1 : X_n = j\}, \text{ con } r \geq 2.
\end{aligned}$$

Gracias a la definición anterior podemos representar formalmente el tiempo que se tarda en volver a j por r-ésima vez.

Definición 3.24. *El tiempo de r-ésimo retorno es, $S_i^{(r)}$, tal que:*

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

Definición 3.25. *Diremos que una variable aleatoria $T : \Omega \longrightarrow \mathbb{N}$ es un tiempo de parada de un proceso estocástico $\{X_n : n \geq 0\}$ si, $\forall n \geq 0$, el suceso $\{T = n\}$ solo depende de $X_0, ...X_n$.*

Veremos ahora que algunos conceptos vistos previamente eran tiempos de parada.

Corolario 3.26. *Las variables aleatorias $T_i^{(r)}$ y $S_i^{(r)}$ son tiempos de parada.*

Demostración. $\{T_i^{(r)} = n\}$ equivale a que $X_n = i$ y en las etapas anteriores la cadena se ha encontrado en el estado i $r - 1$ veces. Por lo que claramente $\{T_i^{(r)} = n\}$ depende únicamente de $X_0, ..., X_n$. Evidentemente, que $T_i^{(r)}$ sea un tiempo de parada, implica que $S_i^{(r)}$ también lo es.

□

En la Definición 3.7 vimos el concepto de propiedad de Márkov. A continuación veremos que esta propiedad se mantiene para tiempos de parada.

Teorema 3.27. *(Propiedad fuerte de Márkov). Sean $\{X_n : n \geq 0\}$ una $CMH(\gamma, \mathcal{M})$ y T un tiempo de parada de la cadena. Suponiendo que $T < \infty$ y $X_T = i$, obtenemos que $\{X_{T+n} : n \geq 0\}$ es una $CMH(\delta_i, \mathcal{M})$ independiente de $X_0, X_1, ...X_T$ con $\delta_i := (\delta_{ij})_{j \in E}$ verificando que, para $i = j$, $\delta_{ij} = 1$ y, para $i \neq j$, $\delta_{ij} = 0$. En otras palablas, para todo suceso A determinado por $X_0, X_1, ..., X_T$ y todo $m > 0$ se verifica que:*

$$\begin{aligned}
&P(\{X_{T+1} = j_1, X_{T+2} = j_2, ..., X_{T+m} = j_m\} \cap A | T < \infty, X_T = i) \\
&\qquad = P(X_1 = j_1, ..., X_m = j_m | X_0 = i) \cdot P(A | T < \infty, X_T = i).
\end{aligned}$$

Demostración. Si tenemos un suceso A determinado por $X_0, X_1, ..., X_T$, evidentemente $A \cap \{T = m\}$ depende únicamente de $X_0, X_1, ..., X_m$. Por tanto, usando la Definición 2.8 de probabilidad condicionada:

$$
\begin{aligned}
P(\{X_{T+1} = j_1, &X_{T+2} = j_2, ..., X_{T+n} = j_n\} \cap A \cap \{T = m\} \cap \{X_T = i\}) \\
&= P(\{X_{T+1} = j_1, X_{T+2} = j_2 ..., X_{m+n} = j_n\} \cap A \cap \{T = m\} \cap \{X_m = i\}) \\
&= P(X_{m+1} = j_1, ..., X_{m+n} = j_n | A \cap \{T = m\} \cap \{X_m = i\}) \times \\
&\quad \times P(A \cap \{T = m\} \cap \{X_m = i\}).
\end{aligned}
$$

Como tenemos el primer factor del producto escrito en términos de X_m, podemos aplicar la propiedad de Márkov al tiempo m. Usando esto y la probabilidad de homogeneidad obtenemos que:

$$
\begin{aligned}
P(\{X_{T+1} = j_1, &X_{T+2} = j_2, ..., X_{T+n} = j_n\} \cap A \cap \{T = m\} \cap \{X_T = i\}) \\
&= P(X_{m+1} = j_1, ..., X_{m+n} = j_n | X_m = i) \cdot P(A \cap \{T = m\} \cap \{X_m = i\}) \\
&= P(X_1 = j_1, ..., X_n = j_n | X_0 = i) \cdot P(A \cap \{T = m\} \cap \{X_m = i\}).
\end{aligned}
$$

Sumando de $m = 0$ hasta infinito en ambos lados de la igualdad:

$$
\begin{aligned}
P(\{X_{T+1} = j_1, &X_{T+2} = j_2, ..., X_{T+n} = j_n\} \cap A \cap \{T < \infty\} \cap \{X_T = i\}) \\
&= P(X_1 = j_1, ..., X_n = j_n | X_0 = i) \cdot P(A \cap \{T < \infty\} \cap \{X_T = i\}).
\end{aligned}
$$

Finalmente, dividiendo la expresión por $P(\{T < \infty\} \cap \{X_T = i\})$ obtenemos:

$$
\begin{aligned}
P(\{X_{T+1} = j_1, &X_{T+2} = j_2, ..., X_{T+n} = j_n\} \cap A | T < \infty, X_T = i) \\
&= P(X_1 = j_1, ..., X_n = j_n | X_0 = i) \cdot P(A | T < \infty, X_T = i).
\end{aligned}
$$

□

Lema 3.28. *Supongamos que $T_i^{(r-1)} < \infty$. Entonces, $\forall r \geq 2$, $S_i^{(r)}$ es independiente de $\{X_m : m \leq T_i^{(r-1)}\}$ y se verifica que:*

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

Demostración. Evidentemente, si $T = T_i^{(r-1)} < \infty$, entonces $X_T = i$. Aplicando la propiedad fuerte de Márkov en el tiempo de parada T obtenemos que, si $T < \infty$, $\{X_{T+n} : n \geq 0\}$ es una $CMH(\delta_i, \mathcal{M})$ independiente de $X_0, X_1, ..., X_T$. Como $S_i^{(r)}$ solo depende de X_{T+n}, con $n \geq 1$, tenemos que $S_i^{(r)}$ es independiente de $X_0, X_1, ..., X_T$. Además obtenemos que:

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

Lo que equivale a que $S_i^{(r)}$ es el primer tiempo en el que el proceso $\{X_{T+n} : n \geq 0\}$ se encuentra en el estado i. o equivalentemente:

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

□

Definición 3.29. *La* ***probabilidad de primera pasada***, *$f_{ij}^{(n)}$, es la probabilidad de pasar de un estado i a otro estado j en n etapas por primera vez, es decir, para $n \geq 1$:*

$$f_{ij}^{(n)} = P(X_n = j, X_{n-1} \neq j, X_{n-2} \neq j, ..., X_1 \neq j | X_0 = i).$$

Definición 3.30. *Suponiendo que nos encontramos en el estado i. La probabilidad de que la cadena pase al estado j es:*

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

*Nombraremos a ρ_{ii} la **probabilidad de retorno** y a ρ_{ij} para $i \neq j$ la **probabilidad de pasada***

Definición 3.31. *Se define el número de veces que la cadena pasa por un estado i como:*

$$N_i := \#\{n \geqslant 1 : X_n = i\} = \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n = i\}}.$$

El siguiente lema relaciona las dos últimas definiciones:

Lema 3.32. *Sea $r \geq 1$.La probabilidad de que una cadena que parte del estado i se encuentre por lo menos r veces en el estado j es:*

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

Demostración. Para $X_0 = i$ se verifica que $\{N_j \geq r\}$=$\{T_j^{(r)} < \infty\}$. Con esto en cuenta iniciamos la demostración por inducción.

Para $r = 1$ se tiene que:

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

Supongamos, ahora que ya disponemos de caso inicial, que la igualdad es cierta para r y veremos que se cumple también para $r + 1$. Usando la Definición 3.23 y la probabilidad condicionada (Definición 2.8):

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

Usando la hipótesis de inducción y el Lema 3.28:

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

□

A continuación veremos un par de ejemplos de cadenas de Márkov sencillas sobre las cuales calcularemos algunas de estas propiedades.

Ejemplo 3.33. Supongamos que cuando hace un día soleado, es igual de probable que al día siguiente haya tanto nieve como lluvia, que jamás hay 2 días soleados seguidos y que si un día nieva o llueve, al día siguiente hay la misma probabilidad de que el tiempo siga igual o de que cambie. Si cambia, solo la mitad de las veces cambia a soleado.

1. Si hoy hace soleado ¿cuál es la probabilidad de que en dos días vuelva a hacer soleado?
2. Si hoy es un día de sol ¿cuál es la probabilidad de que en tres días haya un día de lluvia?
3. ¿cuál es la probabilidad de retorno del estado de nieve?

En este caso nuestro espacio de estados es:

$$E = \{S, L, N\},$$

donde $S = Soleado, L = Lluvia, N = nieve$.

La matriz de transición asociada es:

$$\mathcal{M} = \begin{array}{c} \\ S \\ L \\ N \end{array} \begin{array}{c} \begin{array}{ccc} S & L & N \end{array} \\ \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.25 & 0.5 & 0.25 \\ 0.25 & 0.25 & 0.5 \end{pmatrix} \end{array}$$

Pasemos a contestar las preguntas:
1. En realidad, como es imposible que haga soleado 2 días esta es la misma probabilidad que la probabilidad de primera pasada:

$$f_{ii}^{(2)} = P(X_2 = S, X_1 \neq S | X_0 = S) = P(X_2 = S | X_0 = S).$$

Pero para calcularlo con mayor facilidad usaremos la Proposición 3.20:

$$P(X_2 = S) = \gamma_1^{(2)}$$

Como sabemos que $X_0 = S$, $\gamma = (1, 0, 0)$, solo falta ver el valor de $\mathcal{M}^2$:

$$\mathcal{M}^2 = \begin{pmatrix} 0{,}25 & 0{,}375 & 0{,}375 \\ 0{,}1875 & 0{,}4375 & 0{,}375 \\ 0{,}1875 & 0{,}375 & 0{,}4375 \end{pmatrix}$$

Por tanto:

$$\gamma \mathcal{M}^2 = (0{,}25; 0{,}375; 0{,}375).$$

De donde fácilmente vemos que:

$$P(X_2 = S) = \gamma_1^{(2)} = 0, 25.$$

2. Usando el mismo procedimiento:

$$P(X_3 = L|X_0 = S) = \gamma_2^{(3)}.$$

Calculamos $\mathcal{M}^3$:

$$\mathcal{M}^3 = \begin{pmatrix} 0{,}1875 & 0{,}40625 & 0{,}40625 \\ 0{,}203125 & 0{,}40625 & 0{,}390625 \\ 0{,}203125 & 0{,}390625 & 0{,}40625 \end{pmatrix}$$

Por tanto:

$$\gamma\mathcal{M}^3 = (0{,}1875; 0{,}40625; 0{,}40625).$$

De donde fácilmente vemos que:

$$P(X_3 = L) = \gamma_2^{(3)} = 0{,}40625.$$

3. Tarde o temprano se volverá al estado de nieve, ya que $p_{iN}^{(1)} > 0$ para todo estado i, de hecho se volverá varias veces, por tanto:

$$\rho_{NN} = P(T_N < \infty|X_0 = N) = 1.$$

Ejemplo 3.34. Supongamos que estamos haciendo un estudio sobre la fidelidad de un usuario sobre ciertas marcas. Supongamos que tenemos tres marcas A, B y C. Ahora supongamos que el 90 % de los usuarios que de la marca A se mantienen fieles a la misma y que el 10 % restante se pasa a la marca B mientras que a ninguno le interesa la marca C. El 50 % de los usuarios de la marca B se acaba pasando a la marca A mientras que el resto se mantiene fiel a la marca. Por último, los usuarios de la marca C no parecen muy contentos con la marca, ya que solo el 10 % de sus usuarios se mantiene fiel a la marca mientras que el resto se cambian a la marca A o la B con la misma probabilidad. Queremos saber:

1. ¿Cuál es la probabilidad de que un usuario de la Marca A acabe consumiendo la marca C?
2. ¿Cuál es la probabilidad de que un usuario de la Marca C acabe consumiendo la marca A? Es la misma probabilidad de que acabe consumiendo la marca B?
3. ¿Cuál es la probabilidad de retorno de la Marca C?
4. Suponiendo que la variable tiempo n esta en años, es decir, que el cliente valora cambiarse de forma anual. ¿Cuál es la probabilidad de que un usuario de la Marca A consuma esta marca por lo menos durante 5 años? y la de que un usuario de la marca C la consuma por lo menos durante 3 años?

En este caso nuestro espacio de estados es:

$$E = \{A, B, C\},$$

La matriz de transición asociada es:

$$\begin{array}{c c} & \begin{array}{ccc} A & B & C \end{array} \\ \begin{array}{c} A \\ B \\ C \end{array} & \begin{pmatrix} 0{,}9 & 0{,}1 & 0 \\ 0,5 & 0,5 & 0 \\ 0,45 & 0,45 & 0,1 \end{pmatrix} \end{array}$$

1. La probabilidad que pide el enunciado es $\rho_{AC} = (T_C < \infty | X_0 = A)$ pero como C no es accesible desde A $\rho_{AC} = 0$.

2. Ahora se nos pide

$$\rho_{CA} = (T_A < \infty | X_0 = C).$$

Una vez salgamos del estado C jamás volveremos a este ya que la probabilidad de pasar tanto de A como B a C es nula. Entonces una vez salgamos de C el usuario se movera entre A y B siguiendo las probabilidades de la matriz y tarde o temprano cambiará de la una a la otra ya que ningua probabilidad entre A y B es nula. Por tanto

$$\rho_{CA} = 1.$$

Siguiendo el mismo argumento vemos que $\rho_{CB} = \rho_{CA} = 1$.

3. La probabilidad de retorno a C $\rho_{CC} = (T_C < \infty | X_0 = C) = 0,1$ ya que la única manera de volver a C es desde el propio C en el primer paso con probabilidad 0,1.

4. La probabilidad de que un usuario de la Marca A la consuma por lo menos 5 años es

$$P(N_A \geq 5 | X_0 = A) = \rho_{AA} \rho_{AA}^4 = 1,$$

ya que $\rho_{AA} = 1$ siguiendo la misma logica que en la respuesta 2.

La probabilidad de que un usuario de la Marca C la consuma por lo menos 3 años es

$$P(N_C \geq 5 | X_0 = C) = \rho_{CC} \rho_{CC}^2 = 0,001,$$

ya que como hemos visto en la respuesta 3 $\rho_{CC} = 0,1$.

4. Clasificación de los estados

Sea $CMH(\gamma, \mathcal{M})$ con conjunto de estados E, estos se pueden clasificar. El objetivo de esta sección es estudiar la clasificación de estados para posteriormente poder estudiar la de las cadenas.

Definición 4.1. *Un estado j se dice **accesible** desde el estado i si y sólo si para algún n entero positivo:*

$$p_{ij}^{(n)} > 0.$$

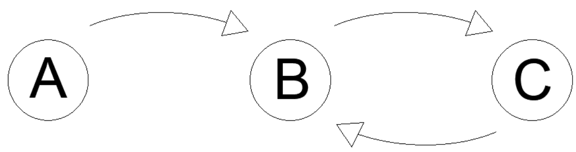
En otras palabras, es posible viajar del estado i al estado j con una probabilidad no nula en n pasos.

Notación 4.2. *En el caso en que tanto j es accesible desde i como i es accesible desde j se dice que i y j se comunican. En tal caso usaremos la siguiente notación:*

$$i \longleftrightarrow j.$$

Observación 4.3. La comunicación, $\longleftrightarrow$, es una relación de equivalencia.

Ejemplo 4.4. Veamos un ejemplo de accesibilidad y comunicación:



B y C se comunican. En cambio, A no es accesible desde B ni desde C, pero ambos son accesibles desde A.

La comunicación genera una partición disjunta del conjunto de estados en clases de equivalencia. Los elementos de una clase son los que se comunican entre sí. Si tenemos dos clases es posible que un elemento de una sea accesible desde la otra, en cambio, es imposible que se comuniquen ya que si se comunicaran pertenecerían a la misma clase y las clases son disjuntas.

Definición 4.5. *Sea C una clase y $i \in C$. Se dice que C es cerrada si todo elemento j accesible desde i pertenece a C.*

Definición 4.6. *Un estado i es **absorbente** cuando una vez que se entra en él no se puede salir, es decir, si para todo estado j tal que $i \neq j$:*

$$p_{ii} = 1.$$
$$p_{ij} = 0.$$

Definición 4.7. *Un estado i es **esencial** si verifica que todo estado accesible desde i se comunica con i.*

Teorema 4.8. *Un estado no esencial no puede ser accesible desde un estado esencial, por lo tanto, la esencialidad es una propiedad de clase.*

Demostración. Sea i un estado esencial que se comunica con un estado cualquiera j. Es suficiente con probar que j es esencial. Si k es accesible desde j, por la esencialidad de i, i es accesible desde k, como i se comunica con j, j se comunica con k. Por tanto j es esencial.

□

Antes de pasar al siguiente tipo de estado veamos una notación que nos será de utilidad.

Notación 4.9. *$n \mid m$ denota que m es divisor de n mientras que $n \nmid m$ denota que m no es divisor de n.*

Definición 4.10. *Un estado i es **periódico** si verifica que:*

$$p_{ii}^{(n)} = 0 \text{ para } n \nmid t.$$
$$p_{ii}^{(n)} \neq 0 \text{ para } n \mid t.$$

Siendo t un número entero mayor que 1.

En el caso de la definición anterior usaremos la siguiente notación:

i es periódico de periodo t.

Si para un estado no existe t entonces diremos que el estado es aperiódico.

Proposición 4.11. *Si el estado i tiene periodo t y j se comunica con i, entonces j tiene periodo t, es decir, la periodicidad es una propiedad de clase.*

Demostración. Supongamos que i tiene periodicidad d y j tiene periodicidad t, con $d \neq t$.Entonces:

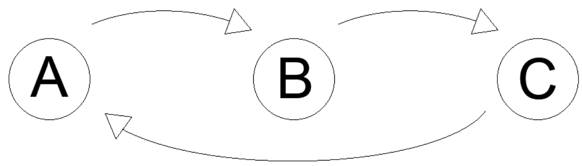
$$p_{jj}^{(n)} = 0 \text{ para } n \nmid t.$$
$$p_{jj}^{(n)} \neq 0 \text{ para } n \mid t.$$
$$p_{ii}^{(n)} = 0 \text{ para } n \nmid d.$$
$$p_{ii}^{(n)} \neq 0 \text{ para } n \mid d.$$

Podemos suponer que $d > t$ sin perdida de generalidad. Entonces, es imposible que $p_{ii}^{(n)} = 0$ para $n \nmid d$ con $p_{jj}^{(n)} \neq 0$ para $n \mid t$ si i y j se comunican, ya que $p_{jj}^{(n)} \neq 0$ implica que $p_{ii}^{(n)} \neq 0$.
Por tanto queda demostrado por contradicción que la periodicidad es una propiedad de clase.

□

Observación 4.12. Fijemonos en que no tiene sentido el periodo para estados no esenciales. Si no hay probabilidad positiva de volver al estado, este no puede tener un periodo. Por tanto, solo se estudiará la periodicidad de un estado si este es esencial.

Ejemplo 4.13. Veamos un ejemplo de estado periódico:



A, B y C son periódicos de periodo 3. En este ejemplo, como solo hay una clase, si un estado era periódico todos tenían que serlo.

4.1. Recurrencia y transitoriedad.

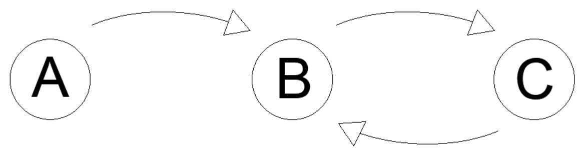
Ahora veremos 2 tipos de estados opuestos, los recurrentes y los transitorios, los cuales nos indicarán diversas propiedades de la cadena.

Definición 4.14. *Un estado i es* ***recurrente*** *si verifica que:*

$$P(X_n = i \text{ para infinitos } n \,|X_0 = i) = 1.$$

Una explicación más intuitiva es que siempre se vuelve a un estado recurrente.

Ejemplo 4.15. Veamos un ejemplo de estado recurrente: B y C son recurrentes. En



cambio A no es recurrente.

Observación 4.16. Evidentemente, un estado absorbente siempre es recurrente.

Proposición 4.17. *Para cualquier estado i, las siguientes afirmaciones son equivalentes:*

1. *El estado i es recurrente.*
2. *El número de veces que se vuelve a i es casi seguramente infinito, es decir:*

$$P(N_i = \infty|X_0 = i) = 1, \text{ o equivalente } P(N_i < \infty|X_0 = i) = 0.$$

3. *La esperanza del número de veces que se vuelve a i es infinita, es decir:*

$$\mathbb{E}[N_i|X_0 = i] = \infty$$

4. *Se verifica que:*

$$\sum_{n=1}^{\infty} f_{ii}^{(n)} = 1.$$

Demostración. Veremos las implicaciones individualmente hasta queden demostradas las equivalencias:
$1 \iff 2$

$$\text{i es recurrente.} \iff \rho_{ii} = 1 \iff P(N_i = \infty | X_0 = i) = 1.$$

$1 \implies 3$, como $1 \implies \rho_{ii} = 1$:

$$\mathbb{E}[N_i | X_0 = i] = \sum_{m=1}^{\infty} mP(N_i = m | X_0 = i) = (1 - \rho_{ii})\rho_{ii} \sum_{m=1}^{\infty} m(\rho_{ii}^{m-1}) = \tfrac{\rho_{ii}}{1-\rho ii} = \infty.$$

$3 \implies 1$

$$\mathbb{E}[N_i | X_0 = i] = \tfrac{\rho_{ii}}{1-\rho ii} = \infty \implies \rho_{ii} = 1 \implies \text{i es recurrente.}$$

$4 \iff 1$

$$\begin{aligned}
\rho_{ii} &= P(X_n = i \text{ para algún } n \geq 1 | X_0 = i) \\
&= P(\bigcup_{n\geq 1} \{X_n = i\} | X_0 = i) \\
&= \sum_{n=1}^{\infty} P(X_n = i, X_{n-1} \neq i, ..., X_1 \neq i | X_0 = i) \\
&= \sum_{n=1}^{\infty} f_{ii}^{(n)}.
\end{aligned}$$

Por tanto,

$$\sum_{n=1}^{\infty} f_{ii}^{(n)} = 1 \iff \rho_{ii} = 1 \iff \text{i es recurrente.}$$

□

Corolario 4.18. *Un estado i es recurrente si y solo si:*

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} = \sum_{n=1}^{\infty} P(X_n = i | X_0 = i) = \infty.$$

Demostración. Para todo i recurrente:

$$\begin{aligned}
\infty = \mathbb{E}[N_i | X_0 = i] &= \mathbb{E}[\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n = i\}} | X_0 = i] \\
= \sum_{n=1}^{\infty} \mathbb{E}[\mathbb{1}_{\{X_n = i\}} | X_0 = i] &= \sum_{n=1}^{\infty} P(X_n = i | X_0 = i).
\end{aligned}$$

Supongamos ahora que $\sum_{n=1}^{\infty} P(X_n = i | X_0 = i) = \infty$. Entonces, $\mathbb{E}[N_i | X_0 = i] = \infty$ lo que implica que i es recurrente.

□

Corolario 4.19. *Sea j un estado recurrente. Entonces todo estado i que se comunique con j es también un estado recurrente.*

Demostración. Como $i \longleftrightarrow j$, existen $a \geq 1$ y $b \geq 1$ tales que:

$$P(X_a = j|X_0 = i) > 0 \text{ y también } P(X_b = i|X_0 = j) > 0.$$

Entonces:

$$\begin{aligned}\sum_{n=a+b}^{\infty} P(X_n = i|X_0 = i) \geq P(X_a = j|X_0 = i)P(X_b = i|X_0 = j)\times \\ \times \sum_{n=a+b}^{\infty} P(X_{n-a-b} = j|X_0 = j) \\ = P(X_a = j|X_0 = i)P(X_b = i|X_0 = j)\sum_{n=0}^{\infty}P(X_n = j|X_0 = j) \\ = \infty\end{aligned}$$

Lo que demuestra que el estado i es recurrente si suponemos que j lo es y usamos el colorario 3.10.

□

Teorema 4.20. *Sea $\{X_n : n \geq 1\}$ una cadena de Márkov homogénea:*
Si $p_{ii} = P(T_i < \infty|X_0 = i) = 1$, entonces i es un estado recurrente y se verifica la igualdad:

$$\sum_{n\geq 1} p_{ii}^{(n)} = \infty.$$

Demostración. Supondremos que $\rho_{ii} = 1$. Aplicando el Lema 3.32 y la propiedad de continuidad secuencial de la probabilidad:

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

Con esto queda probado que el número de veces que la cadena se encuentra en el estado i es infinito o, equivalentemente, que i es recurente. Veamos ahora que se verifica la igualdad:

$$\sum_{n\geq 1} p_{ii}^{(n)} = \sum_{n\geq 1} P(X_n = i|X_0 = i) = E(N_i|X_0 = i) = \infty$$

□

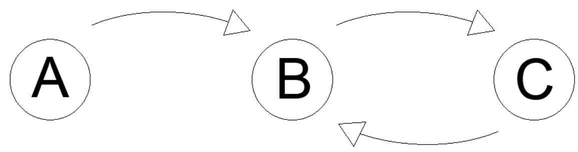
Definición 4.21. *Un estado i es* ***transitorio*** *si verifica que:*

$$P(X_n = i \text{ para infinitos } n \ |X_0 = i) = 0.$$

Una explicación más intuitiva es que a partir de cierto momento jamás se vuelve a un estado transitorio.

Observación 4.22. Un estado i no puede ser transitorio y recurrente. De hecho, se puede definir el estado transitorio como aquel que no es recurrente.

Ejemplo 4.23. Veamos un ejemplo de estado transitorio:



A es transitorio. B y C no lo son ya que como vimos en el Ejemplo 4.15 eran recurrentes.

Proposición 4.24. *Para cualquier estado i, las siguientes afirmaciones son equivalentes:*

1. *El estado i es transitorio.*
2. *El número de veces que se vuelve a i es casi seguramente finito, es decir:*
$$P(N_i = \infty | X_0 = i) = 0, \text{ o equivalente } P(N_i < \infty | X_0 = i) = 1.$$
3. *La esperanza del número de veces que se vuelve a i es finita, es decir:*
$$\mathbb{E}[N_i | X_0 = i] < \infty$$
4. *Se verifica que:*
$$\sum_{n=1}^{\infty} f_{ii}^{(n)} < 1.$$

Demostración. Esta proposición es una consecuencia directa de la Proposición 4.17 y de la Observación 4.22.

□

Corolario 4.25. *Un estado i es transitorio si y solo si:*
$$\sum_{n=1}^{\infty} P(X_n = i | X_0 = i) < \infty.$$

Demostración. Para todo i transitorio:
$$\infty > \mathbb{E}[N_i | X_0 = i] = \mathbb{E}[\sum_{n=1}^{\infty} \mathbb{1}_{X_n = i} | X_0 = i] = \sum_{n=1}^{\infty} \mathbb{E}[\mathbb{1}_{X_n = i} | X_0 = i] = \sum_{n=1}^{\infty} P(X_n = i | X_0 = i).$$

Supongamos ahora que $\sum_{n=1}^{\infty} P(X_n = i | X_0 = i) < \infty$. Entonces, $\mathbb{E}[N_i | X_0 = i] < \infty$ lo que implica que i es transitorio.

□

Corolario 4.26. *Sea j un estado transitorio. Entonces todo estado i que se comunique con j es también un estado transitorio.*

Demostración. Si suponemos que esta afirmación es falsa, entonces por la Observación 4.22 el estado i sería recurrente, pero esto contradice el corolario 4.19.

□

Teorema 4.27. *Sea* $\{X_n : n \geq 1\}$ *una cadena de Márkov homogénea:*
Si $p_{ii} = P(T_i < \infty | X_0 = i) < 1$, *entonces i es un transitorio y se verifica la igualdad:*

$$\sum_{n \geq 1} p_{ii}^{(n)} < \infty.$$

Demostración. Es consecuencia directa del Teorema 4.20 y la Observación 4.22.

□

Teorema 4.28. *Los elementos de una clase C son todos transitorios o todos recurrentes.*

Demostración. Es consecuencia directa de los Corolarios 4.19 y 4.26.

□

Diremos que una clase es transitoria o recurrente en función de la clasificación de sus elementos.

Teorema 4.29. *Sea* $\{X_n; n \geq 0\}$ *una cadena de Márkov con un número de estados finito. Entonces* $\{X_n; n \geq 0\}$ *tiene al menos un estado recurrente.*

Demostración. Viendo que:

$$\mathbb{E}[N_j | X_0 = i] = (1 - \rho_{jj})\rho_{ij} \sum_{m=1}^{\infty} m(\rho_{jj}^{m-1}) = \frac{\rho_{ij}}{1 - \rho jj}$$

Sean i y j dos estados. Supongamos que j es transitorio, entonces $\rho_{jj} < 1$ y por la proposición 3.15:

$$\mathbb{E}[N_j | X_0 = i] = \sum_{n=1}^{\infty} P(X_n = j | X_0 = i) < \infty$$

lo que implica que:

$$\lim_{n \to \infty} P(X_n = j | X_0 = i) = 0$$

para cualquier j estado transitorio. En el caso en el que todos los estados sean transitorios, como hay finitos estados, por la ley de las probabilidades totales, tenemos que:

$$0 = \sum_j \lim_{n \to \infty} P(X_n = j | X_0 = i) = \lim_{n \to \infty} \sum_j P(X_n = j | X_0 = i) = \lim_{n \to \infty} 1 = 1$$

lo que es una contradicción. Por tanto, es imposible que todos los estados sean transitorios, lo que implica que al menos hay uno de recurrente.

□

Teorema 4.30. *Un estado transitorio no puede ser accesible desde un estado recurrente.*

Demostración. $i \longrightarrow j$ implica que existe n verificando que $p_{ij}^{(n)} > 0$. Como i es recurrente, la cadena vuelve a i infinitas veces. Como cada vez que la cadena se encuentra en el estado i tenemos que $p_{ij}^{(n)} > 0$, la cadena se encontrara un número infinito de veces en el estado j o, equivalentemente, j es recurrente.

□

Teorema 4.31. *Toda clase C recurrente es cerrada.*

Demostración. Supongamos que existe C recurrente pero no cerrada. Como C es no cerrada, existe $i \in C$ que se comunica con un estado $j \notin C$. i es recurrente lo que, por el teorema anterior, fuerza a que j sea recurrente lo que contradice que $j \notin C$.

□

Teorema 4.32. *Toda clase C finita y cerrada es recurrente.*

Demostración. C finita y cerrada implica que no se puede salir de la clase. Si C fuera transitoria, tendríamos que sus elementos se visitan un número finito de veces, pero como tenemos un número finito de elementos, entonces habría que salir forzosamente de la clase. Por tanto C es recurrente.

□

Observación 4.33. En el caso finito que un estado sea recurrente es equivalente a que sea esencial. Esto no es necesariamente cierto para cadenas infinitas.

Veremos ahora un ejemplo infinito en el que hay una sola clase, la cual es cerrada y transitoria. El objetivo de este ejemplo es mostrar que para clases infinitas no es cierto que si es cerrada es recurrente. También observaremos que los estados de la clase son esenciales, probando que en el infinito si existen los estados esenciales y transitorios.

Ejemplo 4.34. Supongamos que tenemos una cadena de Márkov infinita $\{X_n : n > 0\}$, la cual verifica que para todo i:

$$P(X_1 = i+1|X_0 = i) = \frac{1}{4}.$$
$$P(X_1 = i-1|X_0 = i) = \frac{3}{4}.$$

Llamaremos a pasar del estado i al $i+1$ moverse a la derecha y pasar de i a $i-1$ moverse a la izquierda. Como la probabilidad de moverse a la izquierda es superior a la de moverse a la derecha, la cadena tiende a ir cada vez más a la izquierda, por tanto, sus estados son transitorios. Como todos los estados están conectados, ya que aunque la cadena tiende a la izquierda hay probabilidad positiva de acceder a todos los estados de la derecha, hay una única clase y todos los estados son esenciales.

Este ejemplo es un paseo aleatorio con valores en $\mathbb{Z}$.

4.2. Tipos de recurrencia

Para todo estado recurrente i, llamaremos tiempo medio de recurrencia de i a

$$\mu_i(i) := E(T_i|X_0 = i) = \sum_{n=1}^{\infty} nP(T_i = n|X_0 = i) = \sum_{n=1}^{\infty} nf_{ii}^{(n)}.$$

Observación 4.35. El tiempo medio de recurrencia no tiene sentido para estados transitorios, ya que estos se visitan un número finito de veces.

Entonces podemos clasificar los estados recurrentes en función de su tiempo medio de recurrencia.

Definición 4.36. *Sea i un estado recurrente. Diremos que es:*

- *recurrente positivo, si su tiempo medio de recurrencia es finito, es decir, si*

$$\mu_i(i) = E(T_i|X_0 = i) < \infty$$

- *recurrente nulo, si su tiempo medio de recurrencia es infinito, es decir, si*

$$\mu_i(i) = E(T_i|X_0 = i) = \infty$$

Teorema 4.37. *Sea $\{X_n : n \geq 0\}$ una cadena de Márkov con conjunto de estados E finito. Entonces, todos los estados recurrentes pertenecientes a E son estados recurrentes positivos.*

Demostración. Como E es finito en realidad

$$\mu_i(i) = E(T_i|X_0 = i) = \sum_{n=1}^{\infty} n f_{ii}^{(n)} = \sum_{n=1}^{\infty} nP(X_n = i, X_{n-1} \neq i, X_{n-2} \neq i..., X_1 \neq i|X_0 = i)$$
$$= \sum_{n=1}^{k} nP(X_n = i, , X_{n-1} \neq i, X_{n-2} \neq i..., X_1 \neq i|X_0 = i) < \infty$$

donde k=#E.

□

Observación 4.38. Cuando trabajemos con conjuntos de estados finitos no hará falta especificar el tipo de recurrencia, ya que esta será siempre positiva.

Con todo esto podemos introducir la noción de otro tipo de estado.

Definición 4.39. *Un estado recurrente se dice que es ergódico si es recurrente positivo y aperiódico.*

Por tanto, en el caso finito solo hace falta que sea recurrente y aperiódico.

5. Clasificación de cadenas

En esta sección se clasificarán las cadenas en diversos grupos en función de las propiedades de sus estados.

Definición 5.1. *Una cadena es* ***irreducible*** *si todos sus estados son alcanzables desde cualquier otro en un número finito de pasos. Esta condición es equivalente a que cualquier estado j es alcanzable desde otro estado i. En caso contrario decimos que es* ***reducible****.*

Ejemplo 5.2. La matriz de transición

$$\begin{pmatrix} 0,2 & 0,3 & 0,5 \\ 0,3 & 0,3 & 0,4 \\ 0,1 & 0,2 & 0,7 \end{pmatrix}$$

es irreducible ya que de hecho todos sus estados son alcanzables desde cualquier otro en un único paso.

Ejemplo 5.3. La matriz de transición

$$\begin{pmatrix} 0,3 & 0,2 & 0,5 \\ 0 & 0,6 & 0,4 \\ 0 & 0,3 & 0,7 \end{pmatrix}$$

es reducible ya que de hecho su primer estado no es alcanzable desde ningún otro.

Definición 5.4. *Decimos que una cadena es* **recurrente** *si todos sus estados son recurrentes.*

Teorema 5.5. *Toda cadena finita y irreducible es recurrente.*

Demostración. Al ser la cadena finita, por el Teorema 4.29, tiene al menos un estado recurrente. Como es irreducible, todos sus estados son alcanzables por cualquier otro en un número finito de pasos, pero como por el Teorema 4.30 un estado transitorio no puede ser accesible desde uno recurrente, obtenemos que todos los estados son recurrentes.

□

Definición 5.6. *Sean $i, j \in \mathbb{N}$. Una matriz $A = [a_{ij}]$ se dice que es positiva si $a_{ij}^{(n)} > 0$ para todos los i,j.*

Definición 5.7. *Una matriz de transición $\mathcal{M}$ se dice que es regular si existe un número entero N tal que $\mathcal{M}^N$ es positivo.*

Notación 5.8. *Una cadena con matriz de transición regular se llama cadena* ***regular****.*

Una cadena regular evidentemente es irreducible. El caso contrario no tiene por qué ser necesariamente cierto.

Definición 5.9. *Decimos que una cadena es* ***ergódica*** *si todos sus estados son ergódicos.*

Definición 5.10. *Decimos que una cadena es* ***periódica*** *si todos sus estados son periódicos. De igual forma decimos que es aperiódica si todos sus estados lo son.*

Observación 5.11. Las cadenas periódicas son irreducibles.

5.1. Metodo de clasificación

El siguiente objetivo es encontrar un método para clasificar fácilmente las cadenas. Para cumplir esa meta, estudiaremos el comportamiento de la cadena a largo plazo. Por norma general, las cadenas de Márkov no suelen tender a un único estado, peró si podemos pensar que, en general, cuando $n \longrightarrow \infty$, la probabilidad de estar en un estado i es independiente del estado inicial y se conoce como probabilidad de estado estable π_i, es decir, la distribución de la cadena se estabiliza. Entonces, usando la Proposición 3.20 nos basta calcular $\mathcal{M}^n$, con n suficientemente grande, para saber el comportamiento de los estados y obtener una clasificación de la matriz.

5.1.1. Distribuciones invariantes

Para estudiar el comportamiento a largo plazo de la cadena estudiaremos el concepto de distribución invariante y veremos que la distribución que estabiliza la cadena es única y invariante.

Definición 5.12. *Una distribución de probabilidad en el conjunto de estados E es cualquier familia π con valores en $[0,1]$ tal que:*

$$\sum_{i \in E} \pi_i = 1.$$

A continuación estudiaremos la distribución de la matriz cuando n tiende a infinito.

Definición 5.13. *Una cadena de Márkov $\{X_n : n \geq 0\}$ se dice que admite una distribución de probabilidad límite si satisface que existe*

$$\lim_{n \to \infty} P(X_n = j | X_0 = i), \tag{5.1}$$

$\forall i, j$ estados de la cadena y además forma una probabilidad de distribución en el conjuto de estados E, es decir:

$$\sum_{j \in E} \lim_{n \to \infty} P(X_n = j | X_0 = i) = 1 \tag{5.2}$$

La ecuación 4.2 se cumple siempre que exista el límite 4.1 y el conjunto de estados E sea finito.

Observación 5.14. Fijemonos en que $P(X_n = j | X_0 = i) = p_{ij}^{(n)}$.

Proposición 5.15. *Una cadena de Márkov $\{X_n : n \geq 0\}$ con conjunto de estados finito E y matriz de transición regular, admite una distribución de probabilidad límite $\pi = (\pi_i)_{i=0,1,\ldots,\#E}$ definida por:*

$$\pi_j = \lim_{n \to \infty} P(X_n = j | X_0 = i) = \lim_{n \to \infty} p_{ij}^{(n)}$$

La demostración nos la saltaremos ya que esto mismo lo veremos en la demostración del Teorema 5.25.

Proposición 5.16. *Sea $\{X_n : n \geq 0\}$ una cadena de Márkov con un estado transitorio j. Entonces $\forall i \in E$:*

$$\lim_{n \to \infty} P(X_n = j | X_0 = i) = 0.$$

Demostración. Como j es transitorio, la probabilidad de retorno de j, ρ_{jj}, en tiempo finito es menor estricta que 1 por definición. Entonces:

$$\begin{aligned} E[N_j|X_0 = i] &= E\left[\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=j\}} \middle| X_0 = i\right] \\ &= \sum_{n=1}^{\infty} E[\mathbb{1}_{\{X_n=j\}}|X_0 = i] \\ &= \sum_{n=1}^{\infty} P(X_n = j|X_0 = i) \\ &= \frac{\rho_{ij}}{1-\rho_{jj}} < \infty. \end{aligned}$$

La convergencia de la serie implica que su término general tiende a 0, es decir:

$$\lim_{n\to\infty} P(X_n = j|X_0 = i) = 0.$$

Como esto se verifica para todo $i \in E$, ya hemos demostrado la proposición.

□

Proposición 5.17. *Sea $\{X_n : n \geq 0\}$ una cadena de Márkov con un estado recurrente j. Entonces $\forall i \in E$:*

$$\lim_{n\to\infty} P(X_n = j|X_0 = i) > 0.$$

Demostración. Es una consecuencia inmediata de la proposición anterior y de la Observación 4.22.

□

Una medida sobre el conjunto E es un vector $\{\gamma_x : x \in E\}$ tal que $0 \leq \gamma_x \leq \infty$, para todo x. Cuando esta es finita, $\sum_{x\in E} \gamma_x < \infty$, esta se puede normalizar y de esta forma obtener una probabilidad en E:

$$\left(\frac{\gamma_x}{\sum_z \gamma_z}, x \in E\right).$$

Una medida γ se dice invariante, respecto a la matriz $\mathcal{M}$ cuando

$$\gamma\mathcal{M} = \gamma,$$

o equivalentemente:

$$\sum_{y\in E} \gamma_y \mathcal{M}_{yx} = \gamma_x, \qquad x \in E.$$

Diremos que una medida γ es estrictamente positiva si $\gamma_x > 0$, para toda $x \in E$.

Definición 5.18. *Sea $\mathcal{M}$ una matriz de transición y ν una distribución de probabilidad sobre el conjunto de estados E. ν es una distribución invariante para $\mathcal{M}$ si se cumple que:*

$$\nu = \nu\mathcal{M}$$

Iterando la expresión de la definición se obtiene que $\nu = \nu\mathcal{M}^n$ para toda $n \geq 1$. Entonces usando la Proposición 3.20 donde vimos que:

$$P(X_n = k) = \gamma_k^{(n)} = (\gamma\mathcal{M}^n)_k$$

Podemos determinar que si tenemos una $CMH(\gamma, \mathcal{M})$ donde la distribución inicial es invariante para $\mathcal{M}$, se obtiene que

$$P(X_n = k) = \gamma_k^{(n)} = (\gamma\mathcal{M}^n)_k = \gamma_k$$

y entonces, X_n tiene una distribución constante. De este resultado se deduce el siguiente teorema:

Teorema 5.19. *Sea $\{X_n : n \geq 0\}$ una $CMH(\gamma, \mathcal{M})$ con γ distribución invariante para $\mathcal{M}$. Entonces, $\{X_{n+m} : n \geq 0\}$, con $m \geq 1$ es una $CMH(\gamma, \mathcal{M})$.*

Nos falta comprobar la existencia de las distribuciones invariantes.

Teorema 5.20. *Toda cadena de Márkov homogénea con un conjunto de estados finitos y matriz de transición $\mathcal{M}$, tiene al menos una distribución invariante.*

Demostración. Sea $\nu = (\nu_1, ..., \nu_{\#E})$ una probabilidad sobre el conjunto de estados E verificando que $\nu_i \in [0,1]$ y $\sum\limits_{i\in E} \nu_i = 1$. Entonces para cualquier $n \geq 1$ se define

$$\nu^{(n)} = \frac{1}{n}\sum_{k=0}^{n-1} \nu\mathcal{M}^k.$$

Evidentemente, $\nu_i^{(n)} \geq 0$, $\forall i \in E$. Si vemos que $\sum\limits_{i\in E} \nu_i^{(n)} = 1$, entonces $\nu^{(n)}$ es una probabilidad sobre E. Entonces como

$$\sum_{i\in E} \nu_i^{(n)} = \frac{1}{n}\sum_{i\in E}\sum_{k=0}^{n-1}\sum_{j\in E} \nu_j \cdot p_{ji}^{(k)} = \frac{1}{n}\sum_{k=0}^{n-1}\sum_{j\in E} \nu_j \left[\sum_{i\in E} p_{ji}^{(k)}\right] = \frac{1}{n}\sum_{k=0}^{n-1}\sum_{j\in E} \nu_j = \frac{1}{n}\sum_{k=0}^{n-1} 1 = 1$$

tenemos que $\nu^{(n)}$ es una probabilidad sobre E. Como el conjunto de probabilidades sobre E es un conjunto cerrado y acotado de $[0,1]^{\#E}$, existe, por el Teorema de Bolzano-Weierstrass, una subsucesión que converge débilmente a un elemento del mismo conjunto. Equivalentemente, existe una probabilidad ν sobre E tal que para alguna subsucesión $\{\nu^{n_k}\}_{k\geq 1}$ se cumple que

$$\nu^{(n_k)} \underset{k\to\infty}{\longrightarrow} \nu$$

Por tanto,

$$\nu^{(n_k)} - \nu^{(n_k)}\mathcal{M} = \frac{1}{n_k}\left[\sum_{m=0}^{n_k-1} \nu\mathcal{M}_m - \sum_{m=0}^{n_k-1} \nu\mathcal{M}_{m+1}\right] = \frac{1}{n_k}(\nu - \nu\mathcal{M}_{n_k}).$$

Haciendo tender k a ∞ y sabiendo que $\{\nu - \nu\mathcal{M}_{n_k}\}_{k\geq 1}$ es una sucesión acotada:

$$\nu - \nu\mathcal{M} = \lim_{k\to\infty} \frac{1}{n_k}(\nu - \nu\mathcal{M}_{n_k}) = 0.$$

Con lo que queda demostrado que ν es una distribución invariante.

$\square$

Teorema 5.21. *Sea E un conjunto de estados finito. Supongamos que para algún $i \in E$ se cumple que*

$$p_{ij}^{(n)} \underset{n\to\infty}{\longrightarrow} \nu_j, \ \forall j \in E.$$

Entonces, $\nu = \{\nu_j : j \in E\}$ es una distribución invariante para $\mathcal{M}$.

Demostración. Por hipótesis sabemos que

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

Además como E es finito:

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

Por tanto ν es una distribución sobre E. Falta ver la invarianza:

$$v_j = \lim_{n\to\infty} p_{ij}^{(n)} = \lim_{n\to\infty} \sum_{k\in E} p_{ik}^{(n-1)} p_{kj} = \sum_{k\in E} \lim_{n\to\infty} p_{ik}^{(n-1)} p_{kj} = \sum_{k\in E} v_k p_{kj}.$$

lo que evidentemente implica que $\nu = \nu\mathcal{M}$.

$\square$

5.1.2. Ergodicidad y estado estable

Acabamos de ver la existencia de distribuciones invariantes para cadenas de Márkov homogéneas con conjunto de estados finito. En este apartado buscamos demostrar que esta es única y como determinarla.

Definición 5.22. *Una cadena de Markov homogénea es ergódica si $\forall i \in E$ existe el limite*

$$\pi_j = \lim_{n\to\infty} p_{ij}^{(n)},$$

este es independiente de i y define una distribución no degenerada sobre E.

Observación 5.23. Esta definición es equivalente a la que ya dimos de cadena ergódica.

Proposición 5.24. *Si el conjunto de estados E es finito, la distribución definida por la definición anterior es invariante.*

Demostración. Es consecuencia directa del Teorema 5.21.

Teorema 5.25. *Sea $\{X_n : n \geq 0\}$ una $CMH(\gamma, \mathcal{M})$ con conjunto de estados finito E. La matriz de transición de la cadena es regular si, y solo si, la cadena es ergódica.*

Demostración. Primero supondremos que la cadena es regular y demostraremos que entonces es ergódica. Veamos que existe $\nu = (\nu_j)_{j\in E}$ verificando que $\forall i \in E$, $\lim_{n\to\infty} p_{ij}^{(n)} = \nu_j$. Para $n \geq 1$, usaremos la siguiente notación:

$$M_j^{(n)} = \max_{i\in E} p_{ij}^{(n)}.$$

$$m_j^{(n)} = \min_{i\in E} p_{ij}^{(n)}.$$

Usando la equación de Champman-Kolmogorov:

$$M_j^{(n+1)} = \max_{i\in E}\sum_{k\in E} p_{ik}\cdot p_{kj}^{(n)} \leq \max_{i\in E}\sum_{k\in E} p_{ik}M_j^{(n)} = M_j^{(n)}.$$

De forma similar:

$$m_j^{(n+1)} = \min_{i\in E}\sum_{k\in E} p_{ik}\cdot p_{kj}^{(n)} \geq \min_{i\in E}\sum_{k\in E} p_{ik}m_j^{(n)} = m_j^{(n)}.$$

Entonces, tenemos que $\{M_j^{(n)}\}_{n\geq 1}$ es una sucesión decreciente acotada inferiormente y $\{m_j^{(n)}\}_{n\geq 1}$ es una sucesión creciente acotada superiormente. Como ambas son sucesiones acotadas, podemos afirmar que existe su límite el cual escribiremos como M_j y m_j respectivamente. Si vemos que para $j \in E$ se verifica que

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

obtendremos que $M_j = m_j$, pero por definición $M_j \geq p_{ij}^{(n)} \geq m_j$ y si hacemos que n tienda a infinito, entonces

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

con $\nu = M_j = m_j$, para $j \in E$.
Veamos entonces que se verifica la igualdad. A causa de la regularidad, existe un entero $m \geq 1$ tal que $\min\limits_{i,j\in E} p_{ij}^{(m)} > 0$. Llamaremos a este valor μ. Usando la equación de Chapman-Kolmogorov:

$$\begin{aligned} p_{ij}^{(m+n)} &= \sum_{k\in E} p_{ik}^{(m)}p_{kj}^{(n)} = \sum_{k\in E}(p_{ik}^{(m)} - \mu p_{jk}^{(n)})p_{kj}^{(n)} + \mu\sum_{k\in E} p_{jk}^{(n)}p_{kj}^{(n)} \\ &= \sum_{k\in E}(p_{ik}^{(m)} - \mu p_{jk}^{(n)})p_{kj}^{(n)} + \mu p_{jj}^{(2n)}. \end{aligned}$$

Como por definición se verifica que $\mu \leq p_{ik}^{(m)}$ y $p_{jk}^{(n)} \leq 1$, claramente, $p_{ik}^{(m)} - \mu p_{jk}^{(n)} \geq 0$. Por tanto:

$$p_{ij}^{(m+n)} \geq m_j^{(n)}\sum_{k\in E}(p_{ik}^{(m)} - \mu p_{jk}^{(n)}) + \mu p_{jj}^{(2n)} = (1-\mu)m_j^{(n)} + \mu p_{jj}^{(2n)}.$$

Aplicando máximos a la expresión que acabamos de obtener llegamos a:

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

De igual manera, pero aplicando mínimos:

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

Entonces:

$$M_j^{(m+n)} - m_j^{(m+n)} \leq (1-\mu)(M_j^{(n)} - m_j^{(n)}) \underset{n\to\infty}{\longrightarrow} 0.$$

Iterando y gracias a que $0 < \mu \leq 1$:

$$M_j^{(km+n)} - m_j^{(km+n)} \leq (1-\mu)^k(M_j^{(n)} - m_j^{(n)}) \underset{n\to\infty}{\longrightarrow} 0.$$

Por tanto, existe una subsucesión $\{n_k\}_{k\geq 1}$ cumpliendo que:

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

Pero como la sucesión $\{M_j^{(n)} - m_j^{(n)}\}_{n\geq 1}$ es monótona decreciente se llega a la expresion deseada:

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

Para concluir, falta ver que ν es una distribución no degenerada sobre E. Como hemos visto que $\lim\limits_{n\to\infty} p_{ij}^{(n)} = \nu_j$:

$$\sum_{j\in E} p_{ij}^{(n)} \underset{n\to\infty}{\longrightarrow} \sum_{j\in E} \nu_j,$$

Con $\sum\limits_{j\in E} p_{ij}^{(n)} = 1$, por lo que $\sum\limits_{j\in E} \nu_j = 1$. También sabemos que $\forall n \geq m$, se cumple que $m_j^{(n)} \geq m_j^{(m)} \geq \mu > 0$, por tanto:

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

Entonces, ν es una distribución no degenerada sobre E. Como E es finito, ν es invariante. Por tanto ya hemos demostrado la primera implicación.
Suponiendo ahora que la cadena es ergódica demostraremos la regularidad.
$\forall i \in E$, se cumple que $\lim\limits_{n\to\infty} p_{ij}^{(n)} = \nu_j > 0$, entonces para cualquier $j \in E$ existe n_j entero verificando que para todo $n \geq n_j$, tenemos que $p_{ij}^{(n)} > 0$. En particular, $\min\limits_{i\in E} p_{ij}^{(n_j)}$.
Entonces, sea $n_0 = \max(n_1, ..., n_{\#E})$:

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

$\forall j \in E$. Por tanto, $\mathcal{M}$ es regular.

$\square$

El siguiente objetivo es ver que la distribución límite es invariante y, para cadenas regulares, única.

Proposición 5.26. *Supongamos E es un conjunto de estados finito y que la distribución límite:*

$$\pi_j^{(i)} := \lim_{n\to\infty} P(X_n = j | X_0 = i)$$

existe para todos los $i, j \in E$. Sea

$$\lim_{n\to\infty} \mathcal{M}^n = \begin{pmatrix} \pi_0^{(0)} & \pi_1^{(0)} & \cdots & \pi_{\#E}^{(0)} \\ \pi_0^{(1)} & \pi_1^{(1)} & \cdots & \pi_{\#E}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_0^{(\#E)} & \pi_1^{(\#E)} & \cdots & \pi_{\#E}^{(\#E)} \end{pmatrix}.$$

Entonces, para todo estado i, el vector $\pi^{(i)} := (\pi_j^{(i)})_{j\in E}$ es una distribución invariante para $\mathcal{M}$.

Demostración. Sean i, j dos estados de E:

$$\begin{aligned}
\pi_j^{(i)} &= \lim_{n\to\infty} P(X_n = j|X_0 = i) \\
&= \lim_{n\to\infty} P(X_{n+1} = j|X_0 = i) \\
&= \lim_{n\to\infty} \sum_{l\in E} P(X_{n+1} = j|X_n = l)P(X_n = l|X_0 = i) \\
&= \sum_{l\in E} p_{lj}^{(1)} \lim_{n\to\infty} P(X_n = l|X_0 = i) \\
&= \sum_{l\in E} \pi_l^{(i)} p_{lj}^{(1)}.
\end{aligned}$$

Donde hemos podido intercambiar la posición del límite y el sumatorio gracias a que E es finito. Por tanto:

$$\pi^{(i)} = \pi^{(i)}\mathcal{M},$$

o equivalentemente, $\pi^{(i)}$ es una distribución invariante de $\mathcal{M}$ para cualquier i de E.

□

Proposición 5.27. *Supongamos que tenemos $\{X_n; n \geq 0\}$ una $CMH(\gamma, \mathcal{M})$ regular con conjunto de estados E finito. Entonces, la distribución invariante $\pi^{(i)}$ es única.*

Demostración. Usando el Teorema 5.25, como E es finito entonces que la cadena sea regular es equivalente a que esta es ergódica y, por tanto, define una distribución invariante. Dicho esto, supongamos que existe una distribución invariante $\omega = (\omega_j)_{j\in E}$ diferente de $\pi^{(i)}$. Al ser invariante cumple que

$$\omega = \omega\mathcal{M}^n,$$

o equivalentemente que $\forall j \in E$

$$\omega_j = \sum_{i\in E} \omega_i p_{ij}^{(n)}.$$

Si hacemos tender n a infinito y gracias a que E es finito:

$$\omega_j = \sum_{i\in E} \omega_i \lim_{n\to\infty} p_{ij}^{(n)} = \sum_{i\in E} \omega_i \pi^{(i)} = \pi_j^{(i)}.$$

Por lo que $\pi^{(i)}$ es única.

□

Observación 5.28. Como hemos visto que $\pi^{(i)}$ es única, para cualesquiera i, j estados, $\pi^{(i)} = \pi^{(j)}$. Debido a esto pasaremos a denotarla simplemente π.

Definición 5.29. *Nombramos sucesión de transición a la sucesión de las $\mathcal{M}^n$, es decir, a $\{\mathcal{M}_n : n \geq 1\}$ verificando que:*

$$\mathcal{M}_n = \mathcal{M}^n.$$

Definición 5.30. *Diremos que $\{\mathcal{M}_n : n \geq 1\}$ ha llegado a su **estado estable** cuando $\mathcal{M}_n = \mathcal{M}_{n-1}$. En este caso $\mathcal{M}^n$ será la matriz estable y tendrá la siguiente forma:*

$$\mathcal{M}^n = \begin{pmatrix} \pi_0 & \pi_1 & \cdots & \pi_n \\ \pi_0 & \pi_1 & \cdots & \pi_n \\ \vdots & \vdots & \ddots & \vdots \\ \pi_0 & \pi_1 & \cdots & \pi_n \end{pmatrix}.$$

Cuando llegamos a un estado estable, $\mathcal{M}^n$ nos da toda la información necesaria sobre los estados de la cadena.

Teorema 5.31. *Sea $\mathcal{M}^n$ la matriz estable de una cadena de Márkov homogénea. Para todo i perteneciente al conjunto de estados, se verifica que:*

1. *Si $\pi_i > 0$, el estado es recurrente.*
2. *Si $\pi_i = 0$, el estado transitorio.*

Demostración. Evidentemente,

1. Es consecuencia directa de la Proposición 5.17.
2. Es consecuencia directa de la Proposición 5.16.

□

Observemos que gracias a este teorema es muy fácil ver si una cadena de Márkov homogénea aperiódica es recurrente, transitoria, irreducible, reducible, regular y/o ergódica. Veamos rápidamente como:

1. Si $\pi_i > 0$, para todos los estados i de la cadena, entonces la cadena es recurrente, irreducible, regular y ergódica.
2. Si $\pi_i = 0$, para todos los estados i de la cadena, entonces la cadena es transitoria. Este escenario no es posible ya que la suma de todos los π_i tiene que dar 1.
3. Si para algún j estado de la cadena se verifica que $\pi_i = 0$, entonces la cadena es reducible.

Falta ver que ocurre en el caso en que la cadena sea periódica.

Teorema 5.32. *Sea $\{X_n : n \geq 0\}$ una $CMH(\gamma, \mathcal{M})$ con conjunto de estados finito E y periodo $t \geq 2$, el límite*

$$\lim_{n \to \infty} \mathcal{M}^{nt}$$

converge a un estado estable.

Demostración. Toda cadena periódica, por la propia definición, también es recurrente. Entonces, $\mathcal{M}^t$ en realidad es la matriz de transición de una cadena recurrente y aperiódica, o equivalentemente, ergódica. Por tanto, por el Teorema 5.25 también es regular lo que como hemos visto implica que el límite

$$\lim_{n \to \infty} \mathcal{M}^{nt}$$

converge a un estado estable.

□

Por tanto, en terminos de M_n, la cadena es periódica si verifica que a partir de un cierto estado k, para cualquier $n \geq k$

$$M_n = M_m, \text{ con } m = n - t$$

donde t es el periodo del estado.

Entonces, ya podemos clasificar todos los estados de una cadena usando únicamente su matriz de transición. Veamos unos cuantos ejemplos en los cuales utilizaremos hasta 3 cifras decimales.

Ejemplo 5.33. Sea

$$\mathcal{M} = \begin{pmatrix} 0,2 & 0,5 & 0,3 \\ 0,6 & 0,2 & 0,2 \\ 0,3 & 0,4 & 0,3 \end{pmatrix}$$

la matriz de transición de una cadena de Márkov homogénea. Calcularemos la sucesión de las $\mathcal{M}_n$:

$$\mathcal{M}_2 = \begin{pmatrix} 0,43 & 0,32 & 0,25 \\ 0,3 & 0,42 & 0,28 \\ 0,39 & 0,35 & 0,26 \end{pmatrix} \qquad \mathcal{M}_3 = \begin{pmatrix} 0,353 & 0,379 & 0,268 \\ 0,396 & 0,346 & 0,258 \\ 0,366 & 0,369 & 0,265 \end{pmatrix}$$

$$\mathcal{M}_4 = \begin{pmatrix} 0,378 & 0,359 & 0,262 \\ 0,364 & 0,370 & 0,265 \\ 0,374 & 0,362 & 0,263 \end{pmatrix} \qquad \mathcal{M}_5 = \begin{pmatrix} 0,370 & 0,366 & 0,264 \\ 0,375 & 0,362 & 0,263 \\ 0,371 & 0,365 & 0,264 \end{pmatrix}$$

$$\mathcal{M}_6 = \begin{pmatrix} 0,373 & 0,364 & 0,263 \\ 0,371 & 0,365 & 0,264 \\ 0,372 & 0,364 & 0,263 \end{pmatrix} \qquad \mathcal{M}_7 = \begin{pmatrix} 0,372 & 0,365 & 0,264 \\ 0,372 & 0,364 & 0,263 \\ 0,372 & 0,364 & 0,264 \end{pmatrix}$$

$$\mathcal{M}_8 = \begin{pmatrix} 0,372 & 0,364 & 0,264 \\ 0,372 & 0,364 & 0,264 \\ 0,372 & 0,364 & 0,264 \end{pmatrix} \qquad \mathcal{M}_9 = \begin{pmatrix} 0,372 & 0,364 & 0,264 \\ 0,372 & 0,364 & 0,264 \\ 0,372 & 0,364 & 0,264 \end{pmatrix}$$

Como $M_8 = M_9$ hemos llegado al estado estable. Como todos los valores de la matriz son estrictamente positivos, concluimos que la cadena es irreducible, recurrente, regular y ergódica.

Ejemplo 5.34. Sea

$$\mathcal{M} = \begin{pmatrix} 0,2 & 0,5 & 0,3 \\ 0 & 0,6 & 0,4 \\ 0 & 0,8 & 0,2 \end{pmatrix}$$

la matriz de transición de una cadena de Márkov homogénea. Calcularemos la sucesión de las $\mathcal{M}_n$:

$$\mathcal{M}_2 = \begin{pmatrix} 0,04 & 0,64 & 0,32 \\ 0 & 0,68 & 0,32 \\ 0 & 0,64 & 0,36 \end{pmatrix} \qquad \mathcal{M}_3 = \begin{pmatrix} 0,008 & 0,66 & 0,332 \\ 0 & 0,664 & 0,336 \\ 0 & 0,672 & 0,328 \end{pmatrix}$$

$$\mathcal{M}_4 = \begin{pmatrix} 0,002 & 0,666 & 0,332 \\ 0 & 0,68 & 0,332 \\ 0 & 0,666 & 0,334 \end{pmatrix} \qquad \mathcal{M}_5 = \begin{pmatrix} 0 & 0,666 & 0,333 \\ 0 & 0,667 & 0,333 \\ 0 & 0,667 & 0,333 \end{pmatrix}$$

$$\mathcal{M}_6 = \begin{pmatrix} 0 & 0,667 & 0,333 \\ 0 & 0,667 & 0,333 \\ 0 & 0,667 & 0,333 \end{pmatrix} \qquad \mathcal{M}_7 = \begin{pmatrix} 0 & 0,667 & 0,333 \\ 0 & 0,667 & 0,333 \\ 0 & 0,667 & 0,333 \end{pmatrix}$$

Como $M_6 = M_7$ hemos llegado al estado estable. Como todos los valores de la primera columna de la matriz son 0 tenemos que el primer estado es transitorio y por tanto que la matriz es reducible.

Ejemplo 5.35. Sea

$$\mathcal{M} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

la matriz de transición de una cadena de Márkov homogénea. Calcularemos la sucesión de las $\mathcal{M}_n$:

$$\mathcal{M}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \qquad \mathcal{M}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\mathcal{M}_4 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Como $M_4 = M$ la cadena es periódica. Al ser periódica también es recurrente e irreducible.

A la hora de clasificar cadenas infinitas no es tan sencillo como mirar su matriz de transición ya que aunque esta se pueda expresar de forma coherente, esta será infinita y no converge obligatoriamente al elevarla a la n. Por tanto, para cadenas infinitas habrá que hacer un estudio directo de los estados. .Veamos un ejemplo.

Ejemplo 5.36. Supongamos que estamos en el paseo aleatorio con valores en $\mathbb{Z}$, pero esta vez la probabilidad de ir a la derecha o a la izquierda es la misma, por tanto, es 1/2. En este caso no es cierto que la cadena se vaya expandiendo en una dirección. De hecho, si suponemos que partimos de un origen O, los valores con mayor probabilidad son los cercanos a O mientras que cuando más nos alejamos del origen menor es la probabilidad de ese estado. En este escenario todos los valores siguen siendo esenciales, pero los valores cercanos al origen se visitarán infinitas veces, por lo que son recurrentes. Entonces, todos los estados serán recurrentes, ya que como vimos en el Teorema 4.30 no se puede acceder a un estado transitorio desde uno recurrente. Al estar en el caso infinito podemos estudiar el tipo de recurrencia. Sea i un estado de esta cadena:

$$\mu_i(i) = E(T_i|X_0 = i) = \sum_{n=1}^{\infty} nP(T_i = n|X_0 = i) = \sum_{n=1}^{\infty} nf_{ii}^{(n)} = \infty,$$

ya que para todo estado recurrente $\sum\limits_{n=1}^{\infty} f_{ii}^{(n)} = 1$ y $\sum\limits_{n=1}^{\infty} n = \infty$ por lo que, aplicando el criterio de comparación por paso al límite del cociente, obtenemos que la suma es divergente. Entonces, la recurrencia es nula y, por tanto, la cadena no es ergódica.

6. Ampliación de conceptos

Una vez ampliados nuestros conocimientos podemos volver a conceptos anteriores y estudiarlos con mayor profundidad. En esta sección no solo mejoraremos nuestro entendimiento de cosas ya explicadas si no que explicaremos cosas nuevas en base a estas.

6.1. El caso irreductible y recurrente

Durante esta sección trabajaremos sobre una cadena irreducible y recurrente con conjunto de estados E. Empezaremos estudiando las *excursiones* de la cadena entre dos retornos al estado i:

$$\mathcal{E}_k = (X_{T_i^k}, X_{T_i^k+1}, \cdots, X_{T_i^{k+1}}), \quad k \geq 0.$$

Estas excursiones son secuencias aleatorias con una dimensión aleatoria, finita y mayor que 2, compuestas por elementos de $E \setminus \{i\}$ salvo a exepción del elemento inicial y final los cuales son identicamente i. Denotaremos como U el conjunto de secuencias

$$u = (i, i_1, ..., i_n, i),$$

con $n \geq 0$, $i_l \neq i$, $1 \leq l \leq n$. U es numerable y el conjunto de todas las posibles excursiones. Entonces, estas variables aleatorias toman valores numerables y su distribución de probabilidades esta caracterizada por

$$P(\mathcal{E}_k = u), \quad u \in U.$$

Notación 6.1. *Llamaremos P_i a la ley de X condicionada a que $X_0 = i$. Es decir, $P_i(A)$ es la probabilidad de que ocurra el suceso A si sabemos que la cadena se inicia en el estado i.*

Definición 6.2. *Para toda $n \geq 0$, $\mathcal{F}_n$ es la σ-álgebra de todos los eventos determinados por $X_0, X_1, ..., X_n$, es decir:*

$$\mathcal{F}_n = \left\{ \{w : (X_0(w), ..., X_n(w)) \in B_n\} : B_n \in \mathcal{P}(E^{n+1}) \right\},$$

donde $\mathcal{P}(G)$ denota la colección de todos los subconjuntos de G.

Proposición 6.3. *Bajo la probabilidad P_i, la secuencia de las excursiones es independiente e idénticamente distribuida, o equivalentemente, existe una probabilidad $\{p_u : u \in U\}$ con U verificando que, para toda $k > 0$, $u_0, ..., u_k \in U$,*

$$P_i(\mathcal{E}_0 = u_0, \mathcal{E}_1 = u_1, \cdots, \mathcal{E}_k = u_k) = \prod_{l=0}^{k} p_{u_l}.$$

Demostración. Es una consecuencia de la propiedad fuerte de Márkov. Evidentemente, $\{\mathcal{E}_0 = u_0\} \in \mathcal{F}_{T_i}$, y el evento

$$\{\mathcal{E}_1 = u_1, \cdots, \mathcal{E}_k = u_k\}$$

es de la forma:

$$\{X_{T_x+1} = i_1, \cdots, X_{T_x+p} = i_p\},$$

con $p > 0$ y $x_1, \cdots, x_p \in E$. Entonces,

$$
\begin{aligned}
P_i(\mathcal{E}_0 = u_0, \mathcal{E}_1 = u_1, \cdots, \mathcal{E}_k = u_k) = &\\
&= P_i(\{\mathcal{E}_0 = u_0\} \cap \{X_{T_x+1} = i_1, \cdots, X_{T_x+p} = i_p\} | T_x < \infty)\\
&= P_i(\mathcal{E}_0 = u_0) P_i(X_1 = i_1, \cdots, X_p = i_p)\\
&= P_i(\mathcal{E}_0 = u_0) P_i(\mathcal{E}_0 = u_1, \cdots, \mathcal{E}_{k-1} = u_k)\\
&= P_i(\mathcal{E}_0 = u_0) P_i(\mathcal{E}_0 = u_1) \times \cdots \times \mathcal{P}_i(E_0 = u_k)\\
&= p_{u_0} p_{u_1} \cdots p_{u_k},
\end{aligned}
$$

donde $\{p_u : u \in U\}$ es la ley de $\mathcal{E}_0$ en P_i.

□

Teorema 6.4. *Sea $\{X_n : n \geq 0\}$ una cadena de Márkov irreducible y recurrente con matriz de transición $\mathcal{M}$, existe una medida invariante estrictamente positiva γ, la cual es única salvo por la multiplicidad por una constante.*

Demostración. Para probar la existencia, llamaremos γ_y^x al numero medio de visitas al estado y durante la excursion $\mathcal{E}_0$ iniciada en x, es decir,

$$
\begin{aligned}
\gamma_y^x = E_x \sum_{n=1}^{T_x} \mathbb{1}_{\{X_n = y\}} &= \sum_{n=1}^{\infty} P_x(X_n = y, n \leq T_x)\\
&= \sum_{z \in E} \sum_{n=1}^{\infty} P_x(\{X_{n-1} = z, n-1 < T_x\} \cap \{X_n = y\})\\
&= \sum_{z \in E} \left(\sum_{n=2}^{\infty} P_x(X_{n-1} = z, n-1 \leq T_x) \right) \mathcal{M}_{zy}\\
&= (\gamma^x \mathcal{M})_y.
\end{aligned}
$$

En la penúltima igualdad hemos usado que la cadena es recurrente. Como esta también es irreducible, existen n,m tales que $(\mathcal{M}^n)_{xy} > 0$, $(\mathcal{M}^m)_{yx} > 0$. Entonces, como $\gamma_x^x = 1$:

$$0 < (\mathcal{M}^n)_{xy} = \gamma_x^x (\mathcal{M}^n)_{xy} \leq (\gamma^x \mathcal{M}^n)_y = \gamma_y^x,$$

$$\gamma_y^x (\mathcal{M}^m)_{yx} \leq \gamma^x \mathcal{M}^m)_x = \gamma_x^x = 1.$$

En consecuencia, γ^x es una medida estrictamente positiva cumpliendo que $\gamma_x^x = 1$. Veamos ahora la unicidad. Sea λ una medida invariante tal que $\lambda_x = 1$. Entonces,

$$
\begin{aligned}
\lambda_y = \mathcal{M}_{xy} + \sum_{z_1 \neq x} \mathcal{M}_{z_1 y} = \mathcal{M}_{xy} + \sum_{z_1 \neq x} \mathcal{M}_{x z_1} \mathcal{M}_{z_1 y} + \sum_{z_1, z_2 \neq x} \lambda_{z_2} \mathcal{M}_{z_2 z_1} \mathcal{M}_{z_1 y} \geq\\
\geq \sum_{n=0}^{\infty} \sum_{z_1, \ldots, z_n \neq x} \mathcal{M}_{x z_n} \mathcal{M}_{z_n z_{n-1}} \mathcal{M}_{z_1 y} = \sum_{n=0}^{\infty} P_x(X_{n+1} = y, T_x \geq n+1) = \gamma_y^x.
\end{aligned}
$$

Como $\tau = \lambda - \gamma^x$ también es una medida invariante, con $\tau_x = 0$. Sean $y \in E$ y n tal que $(\mathcal{M}^n)_{yx} > 0$. Entonces,

$$0 = \tau_x = \sum_{z \in E} \tau_z (\mathcal{M}^n)_{zx} \geq \tau_y (\mathcal{M}^n)_{yx}.$$

Entonces $\tau_y = 0$ para todo $y \in E$, lo que demuestra la unicidad.

$\square$

Teorema 6.5. *Sea $\{X_n : n \geq 0\}$ una cadena de Márkov irreducible. Existe una distribución de probabilidad invariante π si, y solo si, la cadena es recurrente positiva (todos sus estados son recurrentes positivos). También se verifica que un estado i es recurrente positivo si y solo si todos los estados de la cadena lo son.*

Demostración. Observemos que

$$\mu_i(i) = \sum_{j \in E} \gamma_j^i,$$

recordemos que $\mu_i(i)$ es el tiempo medio de recurrencia al estado i.

Suponiendo que i es recurrente positivo, la probabilidad

$$\pi = \left(\pi_j = \frac{\gamma_j^i}{\mu_i(i)}, j \in E \right)$$

es invariante lo que implica que π es estrictamente positivo. Entonces si i es un estado cualquiera, $\lambda = (\lambda_j = \dfrac{\pi_j}{\pi_i}, j \in E)$ es una medida invariante cumpliendo que $\lambda_x = 1$. Por la irreducibilidad y la demostración de unicidad del teorema anterior:

$$\mu_i(i) = \sum_{j \in E} \gamma_j^i = \sum_{j \in E} \frac{\pi_j}{\pi_i} = \frac{1}{\pi_i} < \infty.$$

$\square$

A continuación veremos el *Teorema ergódico*, el cual es una generalización de la ley de los grandes numeros.

Teorema 6.6. *Sea $\{X_n : n \geq 0\}$ una cadena de Márkov con conjunto de estados finito E, irreducible y recurrente. Sea π su única probabilidad invariante. Si $f : E \longrightarrow \mathbb{R}$ es acotada, entonces casi seguramente,*

$$\frac{1}{n} \sum_{k=1}^{n} f(X_k) \to \sum_{i \in E} \pi_i f(i),$$

cuando n tiende a infinito.

Demostración. Como f es acotada, existe k tal que $|f(i)| \leq k$, para todo $i \in E$. Sea

$$N_i(n) = \sum_{1 \leq k \leq n} \mathbb{1}_{\{X_k = i\}}$$

el número de retornos al estado i con tiempo anterior a n. Estudiaremos el limite cuando $n \to \infty$ de

$$\frac{N_i(n)}{n}$$

Por la propia definición de las S_i

$$S_i^0 + \cdots + S_i^{N_i(n)-1} \leq n \leq S_i^0 + \cdots S_i^{N_i(n)}$$

Entonces

$$\frac{S_i^0 + \cdots + S_i^{N_i(n)-1}}{N_i(n)} \leq \frac{n}{N_i(n)} \leq \frac{S_i^0 + \cdots S_i^{N_i(n)}}{N_i(n)}$$

Pero como las S_i son independientes idénticamente distribuidas, si n tiende a infinito obtenemos que, casi seguramente,

$$\frac{S_i^0 + \cdots S_i^{N_i(n)}}{N_i(n)} \to E_i(T_i) := m_i,$$

ya que $N_i(n) \to \infty$ casi seguramente. Por la ley de los grandes números, casi seguramente:

$$\frac{n}{N_i(n)} \to m_i$$

si $m_i \neq 0$, es equivalente

$$\frac{N_i(n)}{n} \to \frac{1}{m_i}.$$

Sea ahora $F \subset E$. Definimos $\tilde{f} = \sum_{i \in E} \pi_i f(i)$, $c = \sup_i |f(i)|$. Entonces

$$\begin{aligned}
\left| \frac{1}{n} \sum_{k=1}^{n} f(X_k) - \tilde{f} \right| &= \left| \sum_{i \in E} \left(\frac{N_i(n)}{n} - \pi_i \right) f(x) \right| \\
&\leq c \sum_{i \in F} \left| \frac{N_i(n)}{n} - \pi_i \right| + c \sum_{i \notin F} \left(\frac{N_i(n)}{n} - \pi_i \right) \\
&= c \sum_{i \in F} \left| \frac{N_i(n)}{n} - \pi_i \right| + c \sum_{i \in F} \left(\pi_i - \frac{N_i(n)}{n} \right) + 2c \sum_{i \notin F} \pi_i \\
&\leq 2c \sum_{i \in F} \left| \frac{N_i(n)}{n} - \pi_i \right| + 2c \sum_{i \notin F} \pi_i .
\end{aligned}$$

Escogiendo F tal que $\sum_{i \notin F} \pi_i \leq \frac{\epsilon}{4c}$, y $N(w)$ tal que para toda n, $n \geq N(w)$,entonces

$$\sum_{i \in F} \left| \frac{N_i(n)}{n} - \pi_i \right| \leq \frac{\epsilon}{4c},$$

lo que prueba el teorema.

□

6.2. Aperiodicidad

Ya hemos visto que un estado es aperiódico siempre que no sea periódico. En esta sección daremos una definición formal y veremos algunas de sus propiedades.

Notación 6.7. *Sea A una matriz, A_{ij} representa el valor de la matriz en la fila i, columna j.*

Definición 6.8. *Un estado i se dice que es aperiódico si existe N tal que*

$$(\mathcal{M}^n)_{ii} > 0, \quad \forall n \geq N.$$

Lema 6.9. *Las cadenas irreducibles con un estado aperiódico son aperiódicas.*

Demostración. Sea $\mathcal{M}$ la matriz de transición de la cadena. Como la cadena es irreducible existen $r, s \in \mathbb{N}$ verificando que $(\mathcal{M}^r)_{ij} > 0$, $(\mathcal{M}^s)_{jk} > 0$ para ciertos i, j y k. De hecho,

$$(\mathcal{M}^{r+n+s})_{ik} \geq (\mathcal{M}^r)_{ij}(\mathcal{M}^n)_{jj}(\mathcal{M}^s)_{jk} > 0,$$

con $n \geq N$. Por tanto la aperiodicidad queda demostrada ya que se verifica la propiedad para $n \geq N + r + s$

□

Supongamos que tenemos una cadena irreducible y recurrente positiva. Sea π verificando que $\pi_i > 0$ para todo estado i. Entonces el hecho de que exista N tal que, para toda $n \geq N$, $(\mathcal{M}^n)_{ji} > 0$ es una condición necesaria para que se verifique la convergencia $(\mathcal{M}^n)_{ji} \to \pi_i$. Veremos a continuación que esta es también una condición suficiente.

Teorema 6.10. *Sea $\{X_n : n \geq 0\}$ una $CMH(\mu, \mathcal{M})$ irreducible, recurrente positiva y aperiódica. Sea π la única distribución de probabilidad invariante, para todo $i \in E$,*

$$P(X_n = i) \longrightarrow \pi_i,$$

con $n \to \infty$. O equivelentemente,

$$(\mu\mathcal{M}^n)_i \longrightarrow \pi_i,$$

para cualquier distribución de probabilidades inicial μ. En particular, para todos i, j estados,

$$(\mathcal{M}^n)_{ji} \longrightarrow \pi_i.$$

Demostración. Sea $\{Y_n : n \geq 0\}$ una $CMH(\pi, \mathcal{M})$, independiente de $\{X_n : n \geq 0\}$, y i un estado cualquiera. Definimos

$$T = \text{ínf}\{n \geq 0 : X_n = Y_n = i\}.$$

Demostraremos que $P(T < \infty) = 1$. Sea $\{W_n = (X_n, Y_n) : n \geq 0\}$ una $CMH(\lambda, \mathcal{N})$ donde $\lambda_{(i,k)} = \mu_i\pi_k$ y $\mathcal{N}_{(i,k)(j,q)} = \mathcal{M}_{ij}\mathcal{M}_{kq}$. Como $\mathcal{M}$ es aperiódica, para cualesquiera estados i,j,k,q y para toda n lo suficientemente grande, se verifica que

$$\mathcal{N}^n_{(i,k)(j,q)} = \mathcal{M}^n_{ij}\mathcal{M}^n_{kq} > 0.$$

entonces $\mathcal{N}$ es irreducible. Además, $\mathcal{N}$ tiene una distribución invariante

$$\tau_{(i,k)} = \pi_i\pi_k.$$

Por el Teorema 6.5 esto implica que $\mathcal{N}$ es recurrente positiva. Sea T el primer tiempo de paso de la cadena Wn por el punto (i,i), es casi seguramente finito.

Definimos

$$Z_n = \begin{cases} X_n, & \text{si } n \leq T. \\ Y_n, & \text{si } n > T. \end{cases}$$

Por la propiedad fuerte de Márkov, los dos procesos $\{X_{T+n} : n \geq 0\}$ y $\{Y_{T+n} : n \geq 0\}$ son cadenas de Márkov homogéneas independientes de $(X_0, ..., X_T)$. Entonces, $\{Z_n : n \geq 0\}$ es, como $\{X_n\}$, una $CMH(\mu, \mathcal{M})$.

Con todo lo que hemos visto tenemos las tres identidades:

$$\begin{aligned}
&P(Z_n = j) = P(X_n = j),\\
&P(Y_n = j) = \pi_j,\\
&P(Z_n = j) = P(X_n = j, n \leq T) + P(Y_n = j, n > T).
\end{aligned}$$

Entonces,

$$|P(X_n = j) - \pi_j| = |P(Z_n = j) - P(Y_n = j)| \leq P(n < T) \longrightarrow 0.$$

cuando $n \to \infty$.

□

6.2.1. Condición de Doeblin

Diremos que una cadena de Márkov verifica la condición de Doeblin si existen $n \in \mathbb{N}$, $\beta > 0$ y una probabilidad ν en el conjunto de estados E verificando que

$$(\mathcal{M}^{n_0})_{ij} \geq \beta \nu_j, \quad \forall i, j \in E.$$

donde $\mathcal{M}$ es la matriz de transición de la cadena.

En esta sección estudiaremos equivalencias e implicaciones de esta condición.

Observación 6.11. La condición de Doeblin es equivalente a que se verifique que existen $i \in E$ y $n_0 \geq 1$ tales que

$$\inf_{j \in E} (\mathcal{M}^{n_0})_{ji} > 0.$$

Esto evidentemente implica que el estado i es aperiódico.

Lema 6.12. *Las cadenas de Márkov irreducibles, aperiódicas y con conjunto de estados E finito verifican la condición de Doeblin.*

Demostración. Sea i un estado de E. Para todo $j \in E$ existe n_j tal que

$$n \geq n_j \Longrightarrow (\mathcal{M}^n)_{ji} > 0.$$

Sean

$$\tilde{n} = \sup_{j \in E} n_j, \qquad \alpha = \inf_{j} (\mathcal{M}^{\tilde{n}})_{ji}$$

Entonces $\alpha > 0$, y para todo $i \in E$,

$$(\mathcal{M}^{\tilde{n}})_{ji} \geq \alpha.$$

Entonces se verifica la condición de Doeblin para $n_0 = \tilde{n}$, $\beta = \alpha$ y $\nu = \delta_i$.

□

Definición 6.13. *El acoplamiento de dos probabilidades p y q en E es cualquier pareja (X, Y) de variables aleatorias con valores en E tales que p es la ley de X y q es la ley de Y.*

Notación 6.14. *Aclararemos dos notaciones que usaremos en el siguiente lema.*

1. *Escribiremos* $(X,Y) = AC(p,q)$ *para abreviar que* (X,Y) *es el acoplamiento de* p *y* q.

2. *Denotaremos el mínimo de* p *y* q *como* $p \wedge q$, *es decir,* $p \wedge q = \text{mín}(p,q)$.

Sea $||p-q||_1 = \sum\limits_{x\in E} |p_x - q_x|$:

Lema 6.15. *Sean* p *y* q *dos probabilidades en* E. *Se verifica que*

$$||p-q||_1 = 2 \inf_{(X,Y)=AC(p,q)} P(X \neq Y).$$

Demostración. Para cualquier (X.Y) acoplamiento de p y q se verifica que

$$P(X=Y) = \sum_{x\in E} P(X=Y=x) \leq \sum_{x\in E} p_x \wedge q_x,$$

por tanto,

$$P(X \neq Y) \geq 1 - \sum_{x\in E} p_x \wedge q_x = \sum_{x\in E} (p_x - q_x)^+.$$

Lo que implica que

$$||p-q||_1 = \sum_{x\in E} |p_x - q_x| \leq 2P(X \neq Y).$$

Por otro lado, definimos $\alpha = \sum\limits_{x\in E} p_x \wedge q_x$. Si $\mathcal{E}$, U, V y W son variables aleatorias independientes entre si cumpliendo que $P(\mathcal{E}=1) = 1 - P(\mathcal{E}=0) = \alpha$, la ley de U es r definida por $r_x = \alpha^{-1} p_x \wedge q_x$, la ley de V es $\tilde{p}$ definida por $\tilde{p}_x = (1-\alpha)^{-1}(p_x - q_x)^+$, y por último la ley de W es $\tilde{q}_x = (1-\alpha)^{-1}(q_x - p_x)^+$, entonces

$$X = \mathcal{E}U + (1-\mathcal{E})V,$$

$$Y = \mathcal{E}U + (1-\mathcal{E})W$$

es un acoplamiento de p y q verificando que $2P(X \neq Y) = ||p-q||_1$.

□

Gracias a lo que hemos visto sobre acoplamientos podremos demostrar el siguiente teorema.

Teorema 6.16. *Sea* $\{X_n : n \geq 0\}$ *una cadena de Márkov irreducible, con matriz de transición* $\mathcal{M}$, *que cumple la condición de Doeblin. Entonces la cadena es aperiódica, recurrente positiva y, si* π *es su distribución invariante,*

$$\sum_{j\in E} |(\mathcal{M}^n)_{ij} - \pi_j| \leq 2(1-\beta)^{[\frac{n}{n_0}]}, \quad \forall i \in E, \quad n \in \mathbb{N},$$

donde $[\frac{n}{n_0}]$ *es la parte entera de* $\frac{n}{n_0}$.

Demostración. Una cadena irreducible que cumple la condición de Doeblin evidentemente es aperiódica. Probemos el resto.

Primero veremos que para dos probabilidades λ y ν en E,

$$||\lambda\mathcal{M}^n - \nu\mathcal{M}^n||_1 \leq 2(1-\beta)^{[\frac{n}{n_0}]}. \tag{6.1}$$

Para probarlo, por el Lema 6.15, es suficiente con construir un acoplamiento (X_n, Y_n) de $\lambda\mathcal{M}^n$ y $\nu\mathcal{M}^n$ verificando que

$$P(X_n \neq Y_n) \leq (1-\beta)^{[\frac{n}{n_0}]}.$$

Supongamos que $n = kn_0 + m$, con $m < n_0$. Dado (X_0, Y_0) con ley $\lambda \times \nu$ en $E \times E$, para $l = 0, 1, ..., k-1$, definimos $(X_{(l+1)n_0}, Y_{(l+1)n_0})$ en terminos de (X_{ln_0}, Y_{ln_0}) de la siguiente manera. Sea $\{\mathcal{E}_l, U_l, V_l : l \geq 0\}$ una secuencia de variables aleatorias independientes con $\mathcal{E}_l$ siendo una distribución de Bernoulli verificando que $P(\mathcal{E}_l = 1) = \beta = P(\mathcal{E}_l = 0)$, la ley de U_l siendo $\tilde{m} = \beta^{-1}m$ y V_l siendo uniforme en [0,1]. Definimos

$$Q_{ij} = (1-\beta)^{-1}((P_{ij}^{n_0}) - m_y)$$

y $f : E \times [0,1] \longrightarrow E$ verificando que, para todos $i, j \in E$, $\{u : f(i,u) = j\}$ es un subconjunto de Borel de [0,1] y, como V es uniforme en [0,1], la ley de $f(x, V)$ es $Q_{i\hat{A}}$. Sean

$$X_{(l+1)n_0} = \mathcal{E}_l U_l + (1-\mathcal{E}) f(X_{ln_0}, V_l),$$
$$Y_{(l+1)n_0} = \mathcal{E}_l U_l + (1-\mathcal{E}) f(Y_{ln_0}, V_l).$$

Hemos construido un acoplamiento (X_{ln_0}, Y_{ln_0}) de $\lambda\mathcal{M}^{ln_0}$ y $\nu\mathcal{M}^{ln_0}$, para $l = 0, ..., k$, el cual verifica que

$$P(X_{ln_0} \neq Y_{ln_0}) \leq P\big(\bigcap_{m=0}^{l} \mathcal{E}_m = 0\big) = (1-\beta)^l.$$

Entonces como $\{X_n \neq Y_n\} \subset \{X_{kn_0} \neq Y_{kn_0}\}$ se puede construir fácilmente el acoplamiento de (X_n, Y_n). Como con que exista el acoplamiento se verifica la Desigualdad 6.1 y este era el objetivo no la buscaremos en profundidad.

Veamos ahora que para cualquier probabilidad λ en E, $\{\lambda\mathcal{M}^n : n \geq 0\}$ es una secuencia de Cauchy en el espacio de Banach $l^1(E)$. Si $\nu = \lambda\mathcal{M}^m$, gracias a la Desigualdad 6.1 se verifica que

$$||\lambda\mathcal{M}^{n+m} - \lambda\mathcal{M}^n||_1 = ||\nu\mathcal{M}^n - \lambda\mathcal{M}^n||_1 \leq 2c^{n-n_0},$$

donde $c = (1-\beta)^{\frac{1}{n_0}}$. Lo que claramente implica que $\{\lambda\mathcal{M}^n : n \geq 0\}$ es una secuencia de Cauchy en el espacio de Banach $l^1(E)$.

Entonces, como $\{\lambda\mathcal{M}^n : n \geq 0\}$ es una secuencia de Cauchy en el espacio de Banach $l^1(E)$ también es convergente en este mismo espacio. Sea π la probabilidad a la que convergue la secuencia:

$$\pi\mathcal{M} = \lim_{n\to\infty} \lambda\mathcal{M}^{n+1} = \pi.$$

Por tanto, π es invariante y la cadena es recurrente positiva. En consecuencia, y volviendo a usar la Desigualdad 6.1, para toda probabilidad λ de E:

$$||\lambda\mathcal{M}^n - \pi||_1 \leq 2(1-\beta)^{[\frac{n}{n_0}]},$$

lo que establece la tasa de convergencia supuesta en el teorema. De hecho esto también prueba la aperiódicidad.

$\square$

6.3. Cadenas de Márkov reversibles

En esta sección trabajaremos con una cadena irreducible y recurrente positiva. La propiedad de Márkov nos dice que cuando $\{X_n : n \geq 0\}$ es una cadena de Márkov, se deduce que, para toda N, $\{\mathcal{X}_n^N = X_{N-n} : N \geq n \geq 0\}$ es también una cadena de Márkov. En general, la cadena revertida no es homogénea, a no ser que $\{X_n\}$ se inialice con su distribución invariante π.

Proposición 6.17. *Sea $\{X_n : n \geq 0\}$ una $CMH(\pi, \mathcal{M})$ irreducible y π su probabilidad invariante. Entonces, la cadena revertida $\{\mathcal{X}_n^N = X_{N-n} : N \geq n \geq 0\}$ es una $CMH(\pi, \mathcal{N})$ con*

$$\pi_j \mathcal{N}_{ji} = \pi_i \mathcal{M}_{ij}, \quad \forall i, j \in E.$$

Demostración. Evidentemente,

$$P(\mathcal{X}_{n+1} = i | \mathcal{X}_n = j) = P(X_n = i | X_{n+1} = j) = P(X_{n+1} = j | X_n = i) \times \frac{P(X_n = i)}{P(X_{n+1} = j)}.$$

Por lo que claramente la cadena revertida es una cadena de Márkov con la misma ley.

□

Diremos que la cadena $\{X_n : n \geq 0\}$ es **reversible** si $\mathcal{N} = \mathcal{M}$, lo que ocurre si y solo si se cumple que

$$\pi_i \mathcal{M}_{ij} = \pi_j \mathcal{M}_{ji}, \quad \forall i, j \in E.$$

donde π es la distribución invariante de la cadena. De hecho cualquier π verificando esta igualdad es invariante para $\mathcal{M}$. Esta condición todo y ser suficiente no es necesaria.

6.4. Estadística de las cadenas

A continuación introduciremos las nociones básicas para la estimación de los parámetros de una cadena de Márkov. Como vimos en la Proposición 3.20, para toda $n > 0$, la ley del vector aleatorio $(X_0, ..., X_n)$ depende únicamente de la ley inicial λ y de la matriz de transición $\mathcal{M}$. Entonces, lo que nos interesa es bajo que condiciones podemos estimar la pareja $(\lambda, \mathcal{M})$, dada la observación de $(X_0, ..., X_n)$ de forma que el error de la observación tienda a zero para $n \to \infty$.

Empecemos viendo la estimación de λ. Para todo estado i, consideramos la media empírica

$$\Lambda_i^n = \frac{1}{n+1} \sum_{l=0}^{n} \mathbb{1}_{\{X_l = i\}}$$

la cual es un estimador consistente de λ_i. La siguiente proposición que es consequencia immediata del Teorema ergódico (6.6) lo justifica.

Proposición 6.18. *Para cualquier estado i, $\Lambda_i^n \longrightarrow \lambda_i$ casi seguramente, cuando $n \to \infty$.*

Pasamos a ver la estimación de $\mathcal{M}_{ij}$ para cualesquiera i, j estados. Escogemos el estimador

$$\mathbb{M}_{ij}^n = \frac{\sum_{l=0}^{n-1} \mathbb{1}_{\{X_l = i, X_{l+1} = j\}}}{\sum_{l=0}^{n-1} \mathbb{1}_{\{X_l = i\}}}$$

El cual es el cociente empírico del número de veces que se visita el estado j desde el estado i entre el número de veces que se visita el estado i, por tanto, es el promedio de veces que desde el estado i se accede al j. Entonces se verifica la siguiente proposición.

Proposición 6.19. *Para cualesquiera i, j estados, $\mathbb{M}_{ij}^n \longrightarrow \mathcal{M}_{ij}$, casi seguramente, cuando $n \longrightarrow \infty$.*

Demostración. Evidentemente,

$$\mathbb{M}_{ij}^n = \left(\frac{1}{n} \sum_{l=0}^{n-1} \mathbb{1}_{\{X_l = i\}} \right)^{-1} \frac{1}{n} \sum_{l=0}^{n-1} \mathbb{1}_{\{X_l = i, X_{l+1} = j\}}. \tag{6.2}$$

Sabemos, por la Proposición 6.18, que

$$\frac{1}{n} \sum_{l=0}^{n-1} \mathbb{1}_{\{X_l = i\}} \longrightarrow \lambda_i.$$

Para $n \geq 0$, definimos $\mathbb{X}_n = (X_n, X_{n+1})$. Es fácilmente verificable que $\{\mathbb{X}_n; n \geq 0\}$ es una cadena de Márkov irreducible, por tanto recurrente positiva, con conjunto de estados $\mathcal{E} = \{(x, y) \in E \times E : \mathcal{M}_{ij} > 0\}$ y matriz de transición $\mathcal{N}_{(i,j)(q,k)} = \delta_{jq} \mathcal{M}_{qk}$. Esta además admite la distribución invariante $\tau_{(i,j)} = \lambda_i \mathcal{M}_{ij}$. Entonces, aplicando el Teorema ergódico (6.6) a la cadena $\{\mathbb{X}_n\}$, obtenemos que casi seguramente, cuando $n \longrightarrow \infty$:

$$\frac{1}{n} \sum_{l=0}^{n-1} \mathbb{1}_{\{X_l = i, X_{l+1} = j\}} \longrightarrow \lambda_i \mathcal{M}_{ij}.$$

Volviendo a la Igualdad 6.2 obtenemos que $\mathbb{M}_{ij}^n \longrightarrow \mathcal{M}_{ij}$.

□

A continuación veremos un ejemplo sencillo en el que con la observación del comportamiento de una cadena de Márkov aproximaremos su distribución inicial y su matriz de transición.

Ejemplo 6.20. Sea $\{X_n : n \geq 0\}$ una cadena de Márkov con distribución inicial λ y matriz de transición $\mathcal{M}$ desconocidas. Supongamos que tiene 5 posibles estados $i_0, ..., i_4$. Supongamos que hemos experimentado con la cadena y hemos obtenido que en una muestra de 500 resultados hemos obtenido que el estado i_0 se ha obtenido 125 veces, i_1 se ha obtenido 60 veces, i_2 se ha obtenido 85 veces, i_3 se ha obtenido 97 veces y i_4 se ha obtenido 133 veces. Entonces usaremos las proposiciones introducidas en esta sección para aproximar λ y $\mathcal{M}$.

Sabemos que

$$\lambda_{i_n} \simeq \Lambda_{i_n}^{500} = \frac{1}{500} \sum_{l=0}^{499} \mathbb{1}_{\{X_l = i_n\}}$$

por la Proposición 6.18. Entonces,

$$\lambda_{i_0} \simeq \Lambda_{i_0}^{500} = \frac{1}{500}\sum_{l=0}^{499} \mathbb{1}_{\{X_l=i_0\}} = \frac{125}{500} = 0,25.$$

$$\lambda_{i_1} \simeq \Lambda_{i_1}^{500} = \frac{1}{500}\sum_{l=0}^{499} \mathbb{1}_{\{X_l=i_1\}} = \frac{60}{500} = 0,12.$$

$$\lambda_{i_2} \simeq \Lambda_{i_2}^{500} = \frac{1}{500}\sum_{l=0}^{499} \mathbb{1}_{\{X_l=i_2\}} = \frac{85}{500} = 0,17.$$

$$\lambda_{i_3} \simeq \Lambda_{i_3}^{500} = \frac{1}{500}\sum_{l=0}^{499} \mathbb{1}_{\{X_l=i_3\}} = \frac{97}{500} = 0,194.$$

$$\lambda_{i_4} \simeq \Lambda_{i_4}^{500} = \frac{1}{500}\sum_{l=0}^{499} \mathbb{1}_{\{X_l=i_4\}} = \frac{133}{500} = 0,266.$$

Por tanto $\lambda \simeq (0,25; 0,12; 0,17; 0,194; 0,266)$. Solo nos queda aproximar $\mathcal{M}$. Sabemos que

$$\mathcal{M}_{i_n i_m} \simeq \mathbb{M}_{i_n i_m}^{500} = \frac{\sum\limits_{l=0}^{499} \mathbb{1}_{\{X_l=i_n, X_{l+1}=i_m\}}}{\sum\limits_{l=0}^{499} \mathbb{1}_{\{X_l=i\}}}$$

Por la Proposición 6.19. Como ya tenemos calculados los

$$\frac{1}{500}\sum_{l=0}^{499} \mathbb{1}_{\{X_l=i_n\}}$$

solo necesitamos calcular

$$\omega_{i_n i_m} = \frac{1}{500}\sum_{l=0}^{499} \mathbb{1}_{\{X_l=i_n, X_{l+1}=i_m\}}$$

para cada pareja de n, m con $4 \leq n, m \leq 0$. Pero no tenemos información suficiente para calcular $\mathbb{1}_{\{X_l=i_n, X_{l+1}=i_m\}}$, ya que para esto necesitaríamos saber el orden de las observaciones. Por tanto en este ejemplo no podemos calcular una aproximación de $\mathcal{M}$.

7. Parte de programación

Una vez introducida toda la teoría sobre la clasificación de cadenas de Márkov, se puede iniciar la parte de programación. El objetivo principal es crear un programa capaz de clasificar cualquier cadena finita a partir de su matriz de transición. Para tal fin se ha usado el lenguaje VBA debido a que es el lenguaje que mejor domino y me parecía muy adecuado para la tarea.

El interés principal de esta sección es que la programación ayuda mucho a relacionar conceptos. En este caso ayuda a relacionar las matrices de transición con la clasificación de las cadenas.

El programa es capaz de clasificar cualquier cadena de Márkov homogénea finita siempre que se indique su matriz de transición. Primero detecta la dimensión de la matriz, después va calculando las iteraciones de los M_n. Cuando ha llegado a un estado estable, toma las conclusiones con el criterio ya comentado en la sección anterior. En caso de no encontrar un estado estable, este comprueba si la cadena es periódica comparando el último M_n con todos los anteriores.

Si la cadena no es periódica ni converge en un estado estable, saldrá un mensaje de error indicando que no se ha podido clasificar y preguntando si no ha habido un error a la hora de indicar la matriz.

Como anexo se encuentra el archivo Excel con el programa listo para clasificar cadenas. El propio archivo Excel tiene las indicaciones suficientes para poder usar el programa.

Me gustaría comentar que algunas versiones de Excel no funcionan demasiado bien con el uso de matrices. Yo he utilizado la versión más reciente y no he tenido ningún problema pero al probarlo en la versión de 2016 tenía problemas de funcionamiento.

8. Conclusiones

Todo y que la idea principal ha sido crear una asignatura sobre las cadenas de Márkov en realidad la mayoría del trabajo ha resultado ser un estudio de las mismas. A pesar de esto, considero que el objetivo se ha cumplido con creces, ya que el proyecto se me hace parecido a los documentos con la teoría relativa a una asignatura.

Sinceramente, al empezar no me imaginaba que hubiera tantísima información basada en estas cadenas. De hecho, incluso centralizándolo en cadenas homogéneas y finitas en tiempo discreto, se ha tenido que escoger muy bien que partes entraban, ya que al tener un límite de páginas no se han podido incluir muchas cosas realmente interesantes.

El tema ya me fascinó desde el principio, cuando mi tutor me lo recomendó, pero actualmente mi interés se ha incrementado, puesto que al investigar he visto que hay muchas aplicaciones de estas cadenas que en este trabajo en realidad se han obviado, ya que nos hemos centrado en la clasificación de las cadenas y no tanto su utilidad.

En cuanto al programa clasificador de cadenas, todo y tener sus limitaciones a causa de las propias limitaciones de las cadenas, considero que se ha logrado un buen resultado. El programa es capaz de clasificar cualquier cadena de Márkov homogénea finita, lo cual considero que con la información que se ha recopilado en este proyecto es lo máximo que se podía conseguir.

Para concluir quiero añadir que creo que el tema que he escogido ha sido perfecto para mí, realmente me ha interesado y ojalá en un futuro esta sea una optativa real de la carrera.

Referencias

[1] Julià, O.; Márquez-Carreras, D.: Un Primer Curs d'estadística. Publicacions i edicions de la Universitat de Barcelona, 2011.

[2] Modica G.; Poggiolini L.: A First Course in Probability and Markov Chains. Wiley, Universidad de Florencia, Italia, 2013.

[3] Pardoux E.: Markov Processes and Applications. Wiley series in Probability and Statistics. Wiley; Dunod, Estados Unidos, 2008.

[4] Privault N.: Understanding Markov Chains. Springer, Singapur, 2018.