UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

TEORÍA DEL POTENCIAL EN CADENAS DE MARKOV

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

MATEMÁTICO

P R E S E N T A :

DIEGO RAMÍREZ ARAQUE

TUTOR

DR. SERGIO IVÁN LÓPEZ ORTEGA

CIUDAD DE MÉXICO 2019

Contenido

Prefacio

Actualmente, la teoría de los procesos estocásticos se ha desarrollado prolíficamente de manera que se ha logrado conectar otras teorías matemáticas con la teoría de los procesos estocásticos hasta el punto de combinar técnicas matemáticas de la probabilidad, cálculo, álgebra lineal, teoría de conjuntos y topología como también del análisis matemático como lo es la teoría de la medida y análisis funcional. Un ejemplo de esto puede observarse en la Teoría del Potencial probabilística la cual es una rama de los procesos estocásticos que desarrolla y define conceptos de procedencia teórica potencial en términos probabilísticos. Una discusión general sobre el tema se aborda en el capítulo introductorio de este texto.

En particular, sobre la Teoría del Potencial probabilística es difícil encontrar fuentes bibliográficas sobre el tópico en el idioma español y los cursos regulares de procesos estocásticos no abordan esta teoría de manera que este estudio pretende ser una fuente de referencia introductoria mediante un recuento sistemático de los principales resultados de la teoría en un contexto adecuado para las cadenas de Markov. En consecuencia, se asumen conocimientos generales sobre la Teoría de cadenas de Markov a tiempo discreto y probabilidad. En particular, las cadenas de Markov son una clase especial de procesos estocásticos que se definen a partir de un espacio de estados discreto y suelen presentarse con parámetro de tiempo discreto o continuo. Debido a esta razón, para estudiar las propiedades potenciales de las cadenas de Markov es necesario introducir a las cadenas de Markov a tiempo continuo de manera que se dedica una sección completa al estudio de esta clase de procesos estocásticos.

La estructura de la tesis se presenta de la siguiente manera; el Capítulo 1 presenta una discusión introductoria sobre la Teoría del Potencial y su conexión con la probabilidad y los preliminares del texto, revisando las nociones básicas acerca de la teoría de la medida y probabilidad. En el Capítulo 2 revisamos con detalle la versión continua de las cadenas de Markov estableciendo las propiedades básicas y características de esta clase de procesos siguiendo los principales resultados de su análogo discreto y brindando una definición formal de ellas. En el Capítulo 3 iniciamos de manera formal el estudio de la Teoría del Potencial a través de conceptos como el potencial, la función de Green y el resolvente, en un contexto adecuado para las cadenas de Markov. En el Capítulo 4 revisamos una aplicación de la Teoría Potencial para cadenas de Markov con la intención de hacer evidente la conexión entre cadenas de Markov y la Teoría del Potencial, redefiniendo los componentes de una red eléctrica en términos probabilísticos, describiendo el voltaje y corriente en términos cuantitativos de caminatas aleatorias en gráficas no dirigidas basándonos en el problema de Dirichlet discreto. En el capítulo 5, con base a esta relación se propone un algoritmo para realizar una aproximación a a la función que soluciona el problema de Dirichlet en su versión discreta mediante técnicas de simulación estocástica basándonos en la teoría revisada para así aportar con un método numérico que resuelve el problema. Finalmente, en las Conclusiones discutimos de manera general los resultados de este estudio dirigiendo la discusión a temas avanzados en cadenas de Markov.

Para el estudio de cadenas de Markov a tiempo continuo se siguen las ideas del autor Norris en [Nor97] complementando ciertos aspectos técnicos presentados en [Doo90] por el autor Doob y por el autor Kai Lai Chung en [Chu60]. Para la introducción a la Teoría del Potencial de manera similar se sigue [Nor97] para finalmente avanzar con las ideas presentadas en las fuentes [DS00], [Dan17] y [Rus97] para la Teoría de Redes Eléctricas. El objetivo de esta tesis es el de presentar una visión general de estas ideas, aportando demostraciones rigurosas con el fin de contribuir al estudio de las cadenas de Markov y la Teoría del Potencial.

Capítulo 1

Ideas preliminares

En este capítulo se presenta una discusión introductoria con respecto a la Teoría del Potencial y las Cadenas de Markov.

1.1 Introducción

¿Qué es la Teoría del Potencial?

La Teoría del Potencial surge de la física matemática como parte del desarrollo en la Teoría de la Gravitación Universal y la Electrostática. No obstante, esta teoría es relevante para resolver una amplia gama de problemas tanto en la física como en las matemáticas. El concepto *potencial* surge de la idea que las fuerzas fundamentales de la naturaleza se derivan de potenciales que satisfacen la ecuación de Laplace y sus generalizaciones. En la actualidad, el estudio basado en propiedades de otros tipos de potenciales ha adquirido un significado independiente donde la Teoría del Potencial moderna se caracteriza por la aplicación de métodos matemáticos que combinan nociones de topología y análisis funcional acompañada del uso de métodos axiomáticos que formalizan el contexto en el que puede ser aplicado. Específicamente, la Teoría del Potencial se centra en el estudio de la ecuación de Laplace y sus generalizaciones. En matemáticas la ecuación de Laplace es una ecuación diferencial parcial relevante en varias ramas de la física que a menudo suele presentarse de dos maneras:

$$\Delta f = 0, \qquad\qquad \nabla^2 f = 0,$$

en donde

$$\Delta = \nabla \bullet \nabla = \nabla^2,$$

es el operador de Laplace, $\nabla\bullet$ es el operador divergencia, ∇ es el operador gradiente y f es una función dos veces diferenciable. Las soluciones o funciones que satisfacen está ecuación y sus generalizaciones en general son llamados potenciales y en particular las ecuaciones que resuelven a la ecuación de Laplace son llamadas funciones harmónicas. Cuando el lado derecho de la ecuación se específica una función h, como se muestra a continuación:

$$\Delta f = h,$$

es llamada la ecuación de Poisson, la cual puede ser pensada como una generalización de la ecuación de Laplace.

¿Qué es una cadena de Markov?

Un proceso estocástico es un objeto matemático empírico que modela la evolución de algún fenómeno cuya evolución se rije mediante leyes probabilísticas. En la Teoría de la Probabilidad, el término proceso estocástico se reserva para referirnos a una familia de variables aleatorias (usualmente infinitas) que se relacionan entre sí por alguna propiedad. La definición precisa del término "Cadena de Markov" del cual hacemos uso en este trabajo se precisa más adelante, sin embargo, las próximas consideraciones ayudarán a clarificar el contexto a aquellos lectores familiarizados con el tema.

Los procesos de Markov son una clase de procesos estocásticos que se distinguen por satisfacer la propiedad de Markov; las cadenas de Markov son una clase especial de procesos de Markov definidos a partir de un espacio de estados numerable finito o infinito. El parámetro del tiempo puede presentarse como el conjunto de enteros no negativos o el conjunto de reales no negativos y debido a esto se distinguen el caso de parámetro discreto y el caso de parámetro continuo. En este trabajo sólo consideramos cadenas de Markov con "probabilidades de transición homogéneas en el tiempo" de manera que la frase entrecomillada se asume al referirnos a esta clase de procesos y la cual se precisa en la siguiente sección.

Los procesos de Markov se nombran así debido a A.A. Markov quien introdujo el concepto en 1907 a través de una familia numerable de variables aleatorias indexadas por un parámetro discreto y un número finito de estados. El caso continuo es generalizado e introducido por Kolmogorov en 1936, seguido de contemporáneos cuyas contribuciones prevalecen en la Teoría de los procesos de Markov. El trabajo realizado en el caso continuo por Doob en 1942 y 1945, en conjunto con el trabajo de Paul Lévy en 1951 definieron la silueta de la teoría en el campo de manera formal. En particular, el trabajo de Doob se centra principalmente en la teoría de la probabilidad y las relaciones de esta última con la teoría del potencial [Doo84]. Entre sus estudiantes de doctorado figuran, entre otros, Paul Halmos (1938), David Blackwell (1941), J. Laurie Snell (1951) y John Walsh (1966).

¿Qué es la Teoría Potencial Probabilística?

La Teoría del Potencial y ciertos aspectos de la probabilidad se relacionan a través de los procesos estocásticos. Un ejemplo de esta conexión se debe a que en particular, las cadenas de Markov se definen a través de una matriz P que ordena las probabilidades de transición del proceso de manera que P puede ser usada para definir una función de Green de alguna teoría potencial. Debido a este hecho, existe una amplia variedad de problemas teóricos de esta índole que tienen una interpretación probabilística o pueden ser resueltos con métodos probabilísticos.

Así mismo, existe una simple y conocida correspondencia entre las caminatas aleatorias y las redes eléctricas [DS00]. El primer problema planteado en redes eléctricas es el de determinar el equilibrio de flujo y potenciales sujeto a condiciones externas. Las caminatas aleatorias en gráficas no dirigidas conforman un caso especial de cadenas de Markov; dada una gráfica finita y conectada G con ciertas conductancias asignadas a cada arista, si consideramos la caminata aleatoria que puede moverse de un vértice a otro vértice adyacente con probabilidad de transición proporcional a la conductancia asignada respectivamente, la caminata es equivalente a una cadena de Markov reversible. Las redes eléctricas proveen un lenguaje diferente para las cadenas de Markov y este punto de vista es útil por la intuición generada obtenida del estudio de las redes eléctricas. Al formular esta correspondencia entre las caminatas y las redes eléctricas podemos definir el análogo discreto de las funciones harmónicas que puede ser visto como el voltaje asociado a cada vértice. Bajo ciertas condiciones, uno resuelve el Problema de Dirichlet para encontrar este voltaje.

1.2 Preliminares

En este capítulo presentamos los fundamentos teóricos necesarios para los próximos capítulos.

Observación 1.2.0.1. Adoptamos la convención de que $\mathbb{N} = \{1, 2, ...\}$, es decir, 0 no se considera un número natural. En cambio, denotamos por $\mathbb{Z}^+ = \{0, 1, 2, ...\}$

y por $\emptyset$ al conjunto vacío. Denotamos por $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$, conocido como el conjunto de los reales extendidos; $\mathbb{R}^+ = [0, \infty)$ a los reales no negativos y por $\overline{\mathbb{R}}^+ = [0, \infty]$. De manera similar $\overline{\mathbb{Z}}^+ = \{0, 1, 2, ..., \infty\}$. Finalmente, se denota por $[a, b)$ a los intervalos de la forma $\{x : a \leq x < b \quad \text{con} \, a, b \in \overline{\mathbb{R}}\}$.

Observación 1.2.0.2. Al referirnos a un conjunto discreto, nos referimos a un conjunto numerable finito o infinito.

1.2.1 Lema de Dynkin

Definición 1.2.1.1 (π-sistema)**.** Sea Ω un conjunto. Una colección de subconjuntos $\mathscr{A}$ de Ω es un π-sistema si es cerrado ante intersecciones finitas, es decir, para toda $A_1, A_2 \in \mathscr{A}$, se tiene que

$$A_1 \cap A_2 \in \mathscr{A}.$$

Denotaremos por $\sigma(\mathscr{A})$ a la σ-álgebra generada por $\mathscr{A}$. Si se cumple que $\sigma(\mathscr{A}) = \mathscr{F}$, entonces se dice que $\mathscr{A}$ genera $\mathscr{F}$.

Definición 1.2.1.2 (λ-sistema)**.** Sea Ω un conjunto. Un λ-sistema (o d-sistema) es una colección de subconjuntos $\mathscr{N}$ de Ω que cumple las siguientes condiciones:

(i) El conjunto Ω pertenece a $\mathscr{N}$.

(ii) Para cualesquiera $A, B \in \mathscr{N}$ con $A \subseteq B$, se cumple que

$$B \setminus A \in \mathscr{N}.$$

(iii) Si $\{A_n\}_{n \in \mathbb{N}}$ con $A_n \in \mathscr{N}$ para toda $n \in \mathbb{N}$ y $A_n \subseteq A_{n+1}$, entonces

$$\bigcup_{n=1}^{\infty} A_n \in \mathscr{N}.$$

Teorema 1.2.1.3. Sea $\mathscr{K}$ una colección de subconjuntos de Ω. Entonces $\mathscr{K}$ es σ-álgebra sí y solo sí $\mathscr{K}$ es un λ-sistema y π-sistema.

Proof. Sea $\mathscr{K}$ una σ-álgebra y notemos lo siguiente:

(i) $\Omega, \emptyset \in \mathscr{K}$.

(ii) Como $\mathscr{K}$ es σ-álgebra, para cualesquiera $A, B \in \mathscr{K}$, se cumple que $A^c, B^c \in \mathscr{K}$. En particular, si $A \subseteq B$, entonces, por las propiedades de conjuntos:

$$B \setminus A = B \cap A^c \in \mathscr{K},$$

pues $\mathscr{K}$ es σ-álgebra.

(iii) En general, como $\mathscr{K}$ es σ-álgebra, la unión numerable finita o infinita de subconjuntos de Ω pertence a $\mathscr{K}$. En particular, si $\{A_n\}_{n\in\mathbb{N}}$ con $A_n \in \mathscr{N}$ para toda $n \in \mathbb{N}$ y $A_n \subseteq A_{n+1}$, entonces

$$\bigcup_{n=1}^{\infty} A_n \in \mathscr{N}.$$

Por los tres puntos anteriores, se concluye que $\mathscr{K}$ es un λ-sistema. Como σ-álgebra, en particular la intersección finita de conjuntos pertenece a $\mathscr{K}$, lo cual implica que que $\mathscr{K}$ es un π-sistema. Para el regreso de la afirmación, supongamos $\mathscr{K}$ es un λ y π sistema a la vez y consideremos lo siguiente:

(i) Como $\mathscr{K}$ es λ-sistema, se cumple que $\Omega \in \mathscr{K}$ y como Ω se contiene a sí mismo, en particular se cumple que $\Omega \setminus \Omega = \emptyset \in \mathscr{K}$, mostrando que

$$\Omega, \emptyset \in \mathscr{K}.$$

(ii) Como $\Omega \in \mathscr{K}$ y por hipótesis $\mathscr{K}$ es λ-sistema, tenemos que para toda $A \in \mathscr{K}$ se cumple que $A \subseteq \Omega$ y

$$A^c = \Omega \setminus A \in \mathscr{K}.$$

(iii) Por el inciso anterior y por las leyes DeMorgan se tiene que $(A \cup B) = S \setminus (A^c \cap B^c) \in \mathscr{K}$. Por lo tanto, para cualquier sucesión $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{K}$, de manera que si definimos la sucesión $B_n := \bigcup_{k\leq n} A_k$ satisface que $B_n \subseteq B_{n+1}$ y que

$$\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} B_n \in \mathscr{K},$$

mostrando que $\mathscr{K}$ es σ-álgebra.

$\square$

Teorema 1.2.1.4 (Lema de π-λ-Sistemas de Dynkin)**.** Sea $\mathscr{N}$ un π-sistema. Entonces cualquier λ-sistema que contenga a $\mathscr{N}$, contiene a $\sigma(\mathscr{N})$.

Proof. Definamos $\lambda(\mathscr{N})$ como la intersección de λ-sistemas que contienen a $\mathscr{N}$. De manera similiar a la demostración del Teorema 1.2.1.3 y tomando en cuenta la Definición 1.2.1.2 de λ-sistema es fácil comprobar que $\lambda(\mathscr{N})$ es un λ-sistema.

Definamos el siguiente conjunto:

$$\mathscr{D}_1 = \{A \in \lambda(\mathscr{N}) : A \cap B \in \lambda(\mathscr{N}) \text{ para toda } B \in \mathscr{N}\}.$$

Observemos que $\mathscr{D}_1$ es un λ-sistema por lo siguiente:

(i) Como $\lambda(\mathscr{N})$ es un λ-sistema, en particular se cumple que $\Omega \in \lambda(\mathscr{N})$. Notemos que para toda $B \in \mathscr{N}$, se tiene que $\Omega \cap B = B$ pues $B \subseteq \Omega$ y en particular, $\mathscr{N}$ es una colección de subconjuntos de Ω. Por lo tanto, se concluye que

$$\Omega \in \mathscr{D}_1.$$

(ii) Sean $A, B \in \mathscr{D}_1$ con $A \subseteq B$ y $C \in \mathscr{D}_1$. Por definición de $\mathscr{D}_1$, se tiene que

$$(B \cap C), (A \cap C) \in \lambda(\mathscr{N}),$$

con $(A \cap C) \subseteq (B \cap C)$ pues $A \subseteq B$. Por lo anterior, como $\lambda(\mathscr{N})$ es un λ-sistema se cumple que $(B \cap C) \setminus (A \cap C) \in \lambda(\mathscr{N})$. Con esto en mente, en combinación con las leyes de DeMorgan y la distribución de la intersección sobre la unión, observemos que

$$\begin{aligned}(B \cap C) \setminus (A \cap C) &= (B \cap C) \cap (A \cap C)^c \\ &= (B \cap C) \cap (A^c \cup C^c) \\ &= \big((B \cap C) \cap (A^c)\big) \cup \big((B \cap C) \cap (C^c)\big),\end{aligned}$$

como $\big((B \cap C) \cap (C^c)\big) = \emptyset \in \lambda(\mathscr{N})$, entonces de la última igualdad se sigue que

$$\begin{aligned}(B \cap C) \setminus (A \cap C) &= \big((B \cap C) \cap (A^c)\big) \cup \big((B \cap C) \cap (C^c)\big) \\ &= (B \cap C) \cap (A^c) \\ &= (B \cap A^c) \cap C \\ &= (B \setminus A) \cap C \in \mathscr{D}_1.\end{aligned}$$

Por lo tanto, se concluye que para cualesquiera $A, B \in \mathscr{D}_1$ con $A \subseteq B$, se cumple que

$$(B \setminus A) \in \mathscr{D}_1.$$

(iii) Sean $B \in \mathscr{N}$ y la sucesión de subconjuntos $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{D}_1$ con $A_n \subseteq A_{n+1}$. Definamos $B_n := \bigcup_{k<n}(A_n \cap B) = (\bigcup_{k<n} A_n) \cap B$. Claramente $B_n \subseteq B_{n+1}$ con $\{B_n\}_{n\in\mathbb{N}} \subseteq \mathscr{D}_1$. Por definición de $\mathscr{D}_1$, en particular se tiene que para toda $n \in \mathbb{N}$ se cumple que $A_n \cap B \in \lambda(\mathscr{N})$ e implica que que $B_n \in \lambda(\mathscr{N})$. Por lo anterior,
$$\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} A_n \cap B \ \in \lambda(\mathscr{N}).$$
Por lo tanto, si la sucesión de subconjuntos $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{D}_1$ con $A_n \subseteq A_{n+1}$, se cumple que
$$\bigcup_{n\in\mathbb{N}} A_n \ \in \mathscr{D}_1.$$

Con los incisos anteriores, se deduce que $\mathscr{D}_1$ es un λ-sistema y que $\mathscr{N} \subseteq \mathscr{D}_1$ siempre y cuando $\mathscr{N}$ sea π-sistema pues $\mathscr{N} \subseteq \lambda(\mathscr{N})$. Por definición, $\mathscr{D}_1 \subseteq \lambda(\mathscr{N})$ y debido a lo anterior, se concluye que $\lambda(\mathscr{N}) = \mathscr{D}_1$ pues $\lambda(\mathscr{N})$ es el λ-sistema más chico que contiene a $\mathscr{N}$. Ahora, definamos el siguiente conjunto:

$$\mathscr{D}_2 = \{A \in \lambda(\mathscr{N}) \,:\, A \cap B \in \lambda(\mathscr{N}) \text{ para toda } B \in \lambda(\mathscr{N})\},$$

de manera similar a como se demostró que $\mathscr{D}_1$ es un λ-sistema se puede validar que $\mathscr{D}_2$ también es un λ-sistema, e implica que $\mathscr{D}_2 = \lambda(\mathscr{N})$.

Por definición y lo anterior, $\mathscr{D}_2$ se muestra que es cerrado ante intersecciones, es decir, $\mathscr{D}_2$ es un π-sistema. Como $\lambda(\mathscr{N})$ es λ y π sistema a la vez, por el Teorema 1.2.3 se sique que $\lambda(\mathscr{N}) = \sigma(\mathscr{N})$.

Por lo tanto, si $\mathscr{N}$ es un π-sistema entonces para cualquier λ-sistema $\mathscr{D}$ que contenga a $\mathscr{N}$ se cumple que $\sigma(\mathscr{N}) \subseteq \lambda(\mathscr{N}) \subseteq \mathscr{D}$, demostrando la implicación del teorema. $\square$

Teorema 1.2.1.5. Sea $(\Omega, \mathscr{F})$, un espacio medible. Sean $\mathbb{P}_1$ y $\mathbb{P}_2$ dos medidas de probabilidad definidas en $(\Omega, \mathscr{F})$. Si $\mathbb{P}_1 = \mathbb{P}_2$ coinciden en un π-sistema $\mathscr{A}$ que genera a $\mathscr{F}$, entonces $\mathbb{P}_1 = \mathbb{P}_2$.

Proof. Definamos la siguiente colección de subconjuntos de Ω $\mathscr{D}$ como sigue:

$$\mathscr{D} = \{A \in \mathscr{F} \,:\, \mathbb{P}_1(A) = \mathbb{P}_2(A)\}.$$

Por hipótesis y la definición anterior, asumimos que $A \subseteq \mathscr{D}$. Ahora, consideremos lo siguiente:

(i) Como $\mathbb{P}_1$ y $\mathbb{P}_2$ son medidas de probabilidad, en particular tenemos que $\mathbb{P}_1(\Omega) = \mathbb{P}_2(\Omega) = 1$. Por lo tanto, como $\Omega \in \mathscr{F}$ entonces $\Omega \in \mathscr{D}$.

(ii) Sean $A, B \in \mathscr{D}$ con $A \subseteq B$, en general si $\mathbb{P}$ es una medida de probabilidad en $(\Omega, \mathscr{F})$ se sabe que

$$\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Como $A \subseteq B$, implica que $A \cap B = A$ y por lo anterior se sigue que

$$\begin{aligned} \mathbb{P}_1(B \setminus A) &= \mathbb{P}_1(B) - \mathbb{P}_1(A \cap B) \\ &= \mathbb{P}_1(B) - \mathbb{P}_1(A \cap B) \\ &= \mathbb{P}_1(B) - \mathbb{P}_1(A) \\ &= \mathbb{P}_2(B) - \mathbb{P}_2(A) \\ &= \mathbb{P}_2(B \setminus A). \end{aligned}$$

Por lo tanto, para cuales quiera $A, B \in \mathscr{D}$ con $A \subseteq B$ se cumple que $B \setminus A \in \mathscr{D}$.

(iii) Sea $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathscr{D}$ con $A_n \subseteq A_{n+1}$ y $\bigcup_{n\in\mathbb{N}} A_n = A$. En general, sin importar la elección de $\mathbb{P}$, en particular tenemos que $\mathbb{P}(\bigcup_{n\in\mathbb{N}} A_n) = \mathbb{P}(A)$. Como $P_1(A_n) = P_2(A_n)$ para toda n y $A \in \mathscr{F}$, se sigue que

$$\mathbb{P}_1(A) = \mathbb{P}_2(A).$$

Por lo tanto, $\bigcup_{n\in\mathbb{N}} A_n = A \in \mathscr{F}$.

Por lo anterior, se deduce que $\mathscr{D}$ es un λ-sistema. Como $\mathscr{A}$ es un π-sistema con $\mathscr{A} \subseteq \mathscr{D}$ y debido a que $\mathscr{F} = \sigma(\mathscr{A})$, por el Teorema 1.1.5 se concluye que $\sigma(\mathscr{A}) \subseteq \mathscr{D}$, demostrando la implicación del teorema. $\square$

La demostración del siguiente teorema se omite y se refiere a [Nor97, Teorema 6.6.3] para su demostración.

Teorema 1.2.1.6. Sea $\mathscr{A}_1$ un π-sistema que genera a $\mathscr{F}_1$ y $\mathscr{A}_2$ un π-sistema que genera a $\mathscr{F}_2$. Supongamos que

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2) \quad \text{para toda } A_1 \in \mathscr{F}_1, A_2 \in \mathscr{F}_2.$$

Entonces $\mathscr{F}_1$ y $\mathscr{F}_2$ son independientes.

1.2.2 Procesos Estocásticos y Filtraciones

Definamos al conjunto T como un conjunto arbitrario de índices de tiempo, es decir, T es una colección cuyos elementos representan el tiempo en que sucede algún evento asociado a algún fenómeno aleatorio. En particular, cabe mencionar que solo se consideran procesos estocásticos con índices de tiempo del tipo $\mathrm{T} = \overline{\mathbb{R}}^{+}$ o simplemente $\mathrm{T} = \mathbb{R}^{+}$ (a tiempo continuo) o como $\mathrm{T} = \overline{\mathbb{Z}}^{+}$ o $\mathrm{T} = \mathbb{Z}^{+}$ (a tiempo discreto) (ver Observación 1.2.0.1). Un proceso estocástico se define formalmente como sigue:

Definición 1.2.2.1 (Proceso Estocástico)**.** Sea T un conjunto de índices de tiempo y $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad. Si $(E, \mathscr{E})$ es otro espacio medible, un proceso estocástico definido en $(\Omega, \mathscr{F})$ con valores en $(E, \mathscr{E})$ es una familia de variables aleatorias $\{X_t : t \in \mathrm{T}\}$ con valores en E indexada por el parámetro $t \in \mathrm{T}$. El par $(\Omega, \mathscr{F})$ es llamado el espacio base y $(E, \mathscr{E})$ es llamado el espacio de estados del proceso $\{X_t : t \in \mathrm{T}\}$.

Observación 1.2.2.1. Un proceso $\{X_t : t \in \mathrm{T}\}$ se dice continuo si el conjunto de índices del tiempo T es continuo, de manera similar un proceso $\{X_t : t \in \mathrm{T}\}$ se dice discreto si el conjunto de índices del tiempo T es discreto. Para distinguir ambos casos se recurre a denotar de manera alternativa al proceso $\{X_t : t \in \mathrm{T}\}$ como $(X_t)_{t \geq 0}$ para el caso continuo y por $(X_n)_{n \geq 0}$ para el caso discreto.

A partir de la definición anterior, podemos pensar de manera alternativa al proceso $\{X_t : t \in \mathrm{T}\}$ como una función $X : \mathrm{T} \times \Omega \to E$ en donde

$$X(t, \omega) = X_t(\omega).$$

Cuando fijamos $t \in \mathrm{T}$, X_t denota el estado del proceso $(X_t)_{t \geq 0}$ al tiempo t. Cuando fijamos $\omega \in \Omega$ y permitimos variar a t, el conjunto $\{X_t(\omega) : t \geq 0\}$ define la trayectoria del proceso asociada a ω.

Para modelar el desarrollo de un proceso a través del tiempo, es necesario considerar una familia creciente de sub-σ-álgebras de $\mathscr{F}$.

Definición 1.2.2.2. Sea $(\Omega, \mathscr{F})$ un espacio medible, entonces una filtración $\{\mathscr{F}_t\}_{t \in \mathrm{T}}$ de $(\Omega, \mathscr{F})$ es una familia de sub-σ-álgebras de $\mathscr{F}$ tal que si $s \leq t$ en T, entonces $\mathscr{F}_s \subseteq \mathscr{F}_t$.

La familia de σ-álgebras $\{\mathscr{F}_t\}_{t \in \mathrm{T}}$ se consideran como un objeto que describe la historia del proceso y debido a ésta razón, en ocasiones, la filtración $\{\mathscr{F}_t\}_{t \in \mathrm{T}}$

es llamada como la σ-álgebra de eventos hasta el tiempo t. Un espacio de probabilidad con una filtración es llamado un espacio de probabilidad filtrado.

A partir de este punto, hemos definido la información disponible al tiempo t en términos de la σ-álgebra $\mathscr{F}_t$ y los procesos estocásticos como una colección de variables aleatorias indexadas por un parámetro t. En la siguiente definición, combinaremos estos conceptos para describir la historia del proceso $(X_t)_{t\leq 0}$ al tiempo t.

Definición 1.2.2.3. Sea $\{\mathscr{F}_t\}_{t\in\mathrm{T}}$ una filtración del espacio medible $(\Omega,\mathscr{F})$ y un proceso $(X_t)_{t\in\mathrm{T}}$ definido en $(\Omega,\mathscr{F})$ con valores en $(E,\mathscr{E})$. Entonces, se dice que $(\mathscr{F}_t)_{t\in\mathrm{T}}$ está adaptado a la filtración $\{\mathscr{F}_t\}_{t\in\mathrm{T}}$ si X_t es $\mathscr{F}_t$-medible para toda $t\in\mathrm{T}$.

Definición 1.2.2.4. Sea $(X_t)_{t\in\mathrm{T}}$ un proceso estocástico definido en un espacio de probabilidad $(\Omega,\mathscr{F},\mathbb{P})$. Se define la filtración generada al tiempo t por el proceso $(X_t)_{t\in\mathrm{T}}$ (o simplemente como la filtración natural) como la σ-álgebra en Ω generada por la variables aleatorias X_s, con $s\leq t$ dada por:

$$\mathscr{F}_t^X=\sigma\{X_s : s\leq t\}.$$

Observación 1.2.2.2. Todo proceso $(X_t)_{t\in\mathrm{T}}$ es adapatado a su filtración natural.

Dados dos procesos estocásticos, es natural preguntarse si ambos procesos modelan el mismo fenómeno.

Definición 1.2.2.5. Sean $(X_t)_{t\in\mathrm{T}}$ y $(Y_t)_{t\in\mathrm{T}}$ son dos procesos estocásticos definidos en el mismo espacio de probabilidad $(\Omega,\mathscr{F},\mathbb{P})$ con valores en $(E,\mathscr{E})$. Entonces, se dice que $(Y_t)_{t\in\mathrm{T}}$ es una modificación de $(X_t)_{t\in\mathrm{T}}$ si, para toda $t\in\mathrm{T}$, se cumple que $X_t=Y_t$ casi seguramente.

Intuitivamente, $(X_t)_{t\in\mathrm{T}}$ y $(Y_t)_{t\in\mathrm{T}}$ son modificaciones de uno u otro proceso si para cualquier tiempo t, la probabilidad de que $X_t=Y_t$ es uno. Sin embargo, dada la modificación de un proceso, no garantiza que la probabilidad de que $X_t=Y_t$ para toda t de manera simultánea sea uno.

Definición 1.2.2.6. Sean $(X_t)_{t\in\mathrm{T}}$ y $(Y_t)_{t\in\mathrm{T}}$ son dos procesos estocásticos definidos en el mismo espacio de probabilidad $(\Omega,\mathscr{F},\mathbb{P})$ con valores en $(E,\mathscr{E})$. Entonces, se dice que $(Y_t)_{t\in\mathrm{T}}$ y $(X_t)_{t\in\mathrm{T}}$ son indistinguibles si para casi toda $\omega\in\Omega$, se cumple que

$$X_t(\Omega)=Y_t(\Omega)\qquad \text{para toda } t\in\Omega.$$

Los procesos estocásticos al ser funciones medibles, se puede definir la continuidad por la derecha como sigue:

Definición 1.2.2.7. El proceso $(X_t)_{t\leq 0}$ es continuo por la derecha si para toda $\omega \in \Omega$ la trayectoria $t \to X_t(\omega)$ es continua por la derecha, es decir, para toda $\omega \in \Omega$ y $t \geq 0$ se cumple que

$$\lim_{s \to t^+} X_s(\omega) = X_t(\omega).$$

Definición 1.2.2.8 (Probabilidades de Transición Estacionarias)**.** Un proceso estocástico se dice que tiene probabilidades de transición estacionarias en el tiempo si para cada $i, j \in I$ siempre que $\mathbb{P}\left(X_s(\omega) = j\right) > 0$ y $0 \leq s < t$ la probabilidad de transición

$$\mathbb{P}(X_t = j | X_s = i) = \mathbb{P}(X_{t-s} = j | X_0 = i),$$

es decir, $\mathbb{P}(X_t = j | X_s = i)$ depende únicamente de $(t - s)$.

A esta clase de procesos también se les conoce como procesos estocásticos homogéneos en el tiempo.

1.2.3 Cadenas de Markov a Tiempo Discreto

Sea I un conjunto numerable y definamos I como nuestro espacio de estados. Diremos que $\lambda = (\lambda_i : i \in I)$ es una medida en I si $0 \leq \lambda_i < \infty$ para toda $i \in I$. Si se cumple que $\sum_{i \in I} = \lambda_i = 1$, diremos que λ es una distribución. Para propósitos de este trabajo, a partir de este punto trabajamos con espacios de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$. Se recuerda que una variable aleatoria con valores en I es una función medible $X : \Omega \to I$ y definamos

$$\lambda_i = \mathbb{P}(X = i) = \mathbb{P}(\{\omega : X(\omega) = i\}).$$

Entonces λ define la distribución de X.

Definición 1.2.3.1. Una matriz $P = (p_{ij})_{ij \in I}$ es una matriz estocástica si se satisfacen las siguientes condiciones:

(i) $p_{ij} \geq 0$ para toda $i, j \in I$;

(ii) $\sum_j p_{ij} = 1$ para toda $j \in I$.

Una cadena de Markov parámetro discreto (o cadena de Markov a tiempo discreto) es un proceso estocástico $(X_n)_{n \geq 0}$ con distribución inicial λ y matriz de transición estacionaria P si:

(i) X_0 tiene distribución λ;

(ii) para $n \leq 0$, condicionando en $X_n = i$, X_{n+1} tiene distribución $(p_{ij} : j \in I)$ y es independiente de $X_0, ..., X_{n-1}$.

Alternativamente, estas últimas condiciones se enuncian formalmente como sigue; para $n \leq 0$ e $i_0, ..., i_{n+1} \in I$, se tiene que

(i) $\mathbb{P}(X_0 = i_0) = \lambda_{i_0}$;

(ii) $\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, ..., X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n)$.

La condición (ii) en los puntos anteriores es conocida como la propiedad de Markov. Cuando se cumple que

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, ..., X_n = i_n) = \mathbb{P}(X_1 = i_{n+1} | X_0 = i_n),$$

entonces diremos que $(X_n)_{n \geq 0}$ es una cadena de Markov con probabilidades de transición homogéneas en el tiempo, de manera que para toda $p_{ij} \in I$, tenemos que

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, ..., X_n = i_n) = \mathbb{P}(X_1 = i_{n+1} | X_0 = i_n) = p_{i_n i_{n+1}},$$

definiendo las probabilidades de transición de la cadena de Markov a tiempo discreto.

En este trabajo, toda función medible (en particular distribuciones y medidas λ se consideran como vectores columna cuyos componentes están indexadas por los elementos de un conjunto numerable I, así como P es una matriz indexada por $I \times I$, de manera que el producto λP está bien definido.

Capítulo 2

Cadenas de Markov a Tiempo Continuo

Para definiciones generales, convenciones y notación, se refiere la Sección 1.2.

Las cadena de Markov a tiempo cotinuo se puede pensar como una generalización de las cadenas de Markov a tiempo discreto. Las cadenas de Markov a tiempo continuo son una clase de procesos estocásticos con valores en un conjunto discreto I que de manera similiar a las cadenas de Markov discretas satisfacen la propiedad de Markov, que diferencia de las anteriores, se definen a partir de un parámetro de tiempo continuo de manera que cada cambio de estado o transición pueden realizarse en cualquier momento. En particular, el tiempo en el que se realiza cada salto es una variable aleatoria exponencial cuyo parámetro depende de ciertas matrices ($Q-matrices$).

Un proceso estocástico con parámetro continuo, es una familia de variables aleatorias indexadas por un sólo parámetro t dado por $\{X_t : t \in \mathrm{T}\}$ definidas en el espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$. El parámetro t varía en un conjunto arbitrario lineal T, sin embargo, para objeto de nuestro estudio sólo enfocamos nuestra atención a los conjuntos $\mathrm{T} = [0, \infty)$. En este trabajo sólo contemplamos procesos en el que para toda $t \in \mathrm{T}$, X_t es una variable aleatoria discreta de manera que la unión de todos los posibles valores de X_t es un conjunto numerable finito o infinito I. Tal proceso es nombrado como un proceso continuo e I como el conjunto de espacio de estados. Observemos que a diferencia de un proceso estocástico discreto, un proceso estocástico a tiempo continuo es una familia no numerable de variables aleatorias mientras que en el caso discreto es numerable.

Desde el punto de vista de la teoría general de los procesos estocásticos, una cadena de Markov de parámetro continuo parece ser el primer caso esencial de procesos discontinuos que ha sido estudiado con cierto detalle. Es común que las las trayectorias presenten discontinuidades peores que simples saltos de manera que estas discontinuidades desempeñan un papel central en la teoría de las cadenas de Markov a tiempo continuo.

2.1 Q-matrices

En esta sección se discuten las propiedades básicas de las $Q-matrices$ y su conexión con las cadenas de Markov con mayor detalle.

Definición 2.1.1. *Sea I un conjunto numerable finito. Una $Q-matriz$ en I es una matriz cuadrada de $n \times n$ con $Q = (q_{ij} : i, j \in I)$ que satisface las siguientes condiciones:*

(i) $0 \leq -q_{ii} < \infty$ *para toda* $i \in I$*,*

(ii) $q_{ij} \geq 0$ *para toda* $i \neq j$*,*

(iii) $\sum_{j\in I} q_{ij} = 0$.

Si I es un conjunto numerable finito de cardinalidad N, la matriz antes definida $Q = (q_{ij} : i, j \in I)$ es una matriz cuadrada de $N \times N$. Para construir una $Q-matriz$ las entradas fuera de la diagonal deben ser números reales no negativos sujetos a las condición de que la suma de los elementos fuera de la diagonal de un renglón es finita, es decir:

$$q_i = \sum_{j\neq i} q_{ij} < \infty.$$

Por lo anterior, la entrada diagonal denotada por q_{ii} es igual a $-q_i$ de manera que la suma total del renglón es cero.

Definición 2.1.2. *Sea I un conjunto numerable de estados. Una matriz función de transición $P(t)$ es un arreglo numerable finito o infinito de funciones $P(t) = (p_{ij}(t))_{i,j\in I}$ definidas en el intervalo $[0,\infty)$ que satisface las siguientes condiciones: para toda $i, j \in I$ y $s,t \in [0,\infty)$;*

(i) $p_{ij}(t) \geq 0$

(ii) $\sum_{j\in I} p_{i,j}(t) = 1$

(iii) $\sum_{j\in I} p_{i,k}(s)p_{k,j}(t) = p_{ij}(s+t)$.

Observación 2.1.1. Con respecto a la definición anterior, si definimos a $P(t) = (p_{ij}(t) : i, j \in I)$, entonces las propiedades anteriores se declaran como sigue: cada entrada de la matriz $P(t)$ es una función no negativa; la suma de cada renglón es igual a uno; y la familia de matrices función $(P(t) : t \geq 0)$ es un semigrupo con respecto a la multiplicación usual de matrices. Las condiciones (i) y (ii) juntas expresan que para toda $t \geq 0$, la matriz $P(t)$ es estocástica. La condición (iii) se conoce como *la ecuación de Chapman-Kolmogorov*. En notación matricial, la ecuación anterior se escribe

$$P(s+t) = P(s)P(t).$$

Consideremos el espacio discreto $\mathbb{Z}^+$ como un espacio encajado en el espacio de parámetro continuo $[0, \infty)$. Para $p \in (0, \infty)$ una forma natural de interpolar la sucesión discreta $\{p^n : n = 0, 1, 2, ...\}$ es ocupando la función e^{tq}, de manera que $q = logp$. Si I es un conjunto finito I y la matriz $P = (p_{ij} : i, j \in I)$, ¿existe alguna manera de interpolar la sucesión de matrices $(P^n : n = 0, 1, 2, ...)$?

En general, para cualquier matriz cuadrada Q de $n \times n$, las componentes de la serie de matrices

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{(Q)^k}{k!} \quad \text{con} \quad t \in [0, \infty), \tag{2.1.1}$$

convergen uniformemente y el límite de la serie es denotado por e^{tQ}. Adicionalmente, se recuerda que para dos matrices Q_1 y Q_2 que conmutan se tiene que

$$e^{Q_1+Q_2} = e^{Q_1}e^{Q_2}.$$

Las pruebas de las afirmaciones anteriores se siguen del caso escalar (e^{tQ}) y se precisan con mayor detalle en el *Apéndice A*. Supongamos que existe una matriz Q tal que $e^Q = P$, entonces

$$e^{nQ} = (e^Q)^n = P^n,$$

de manera que $(e^{tQ} : t \geq 0)$ rellena los huecos de la sucesión discreta $(P^n : n = 0, 1, 2, ...)$.

Observación 2.1.2. Con respecto al caso escalar, si definimos la matriz función $P(t) = e^{tQ}$, esta matriz define un arreglo de funciones dada por $P(t) = (p_{ij}(t))_{ij \in I}$, sin embargo evitamos esta notación y recurrimos a denotar $P(t) = (p_{ij}(t) : t \in \mathrm{T})$. Sea Q una Q-matriz y denotemos por $q_{ij}^{(n)}$ a la (i, j) componente de la potencia n

de la matriz Q (*i.e.*, Q^n), entonces por la definición de e^{tQ}, para toda $t \in [0,\infty)$ se tiene que

$$p_{ij}(t) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{t^k q_{ij}^{(k)}}{k!}. \tag{2.1.2}$$

Observación 2.1.3. Con respecto a la matriz $P(t) = e^{tQ}$, para toda $i, j \in I$ la función $p_{ij}(t)$ es derivable en $[0,\infty)$ y su derivada es continua. Este hecho no se demuestra pero se justifica de manera informal mediante un teorema de análisis matemático real que se enuncia a continuación. Un argumento formal de esto se puede revisar en la siguiente referencia (ver Kai Lai Chung. (1960) Capítulo II, Sección 3. Markov Chains With Stationary Transition Probabilities. Springer Verlag, Berlín, Alemania).

Denotamos por I a la matriz identidad de manera que $(\mathrm{I})_{ij} = \delta_{ij}$ y

$$\delta_{ij} = \begin{cases} 1 & \text{si } j = i, \\ 0 & \text{en otro caso.} \end{cases}$$

Teorema 2.1.3. *Asumamos que la serie de potencias $\sum_{n=0}^{\infty} a_n(x-x_0)$ converge para cada x en $(x_0 - r, x_0 + r)$. Entonces la función f definida por la ecuación*

$$f(x) = \sum_{n=0}^{\infty} a_n (x-x_0)^n,$$

si $x \in (x_0 - r, x_0 + r)$ entonces f tiene derivada $f'(x)$ para toda $x \in (x_0 - r, x_0 + r)$, dada por

$$f'(x) = \sum_{n=0}^{\infty} n a_n (x-x_0)^{n-1}.$$

La demostración del teorema anterior se omite y se refiere a [Apo74, Teorema 9.23] para su demostración. Se recuerda que una matriz Q es una matriz cuadrada de $n \times n$ con n igual a la cardinalidad del conjunto I (ver Definición 2.1.1).

Teorema 2.1.4. *Sea Q una Q-matriz definida para un conjunto finito. Definamos $P(t) = e^{tQ}$ entonces $(P(t) : t \geq 0)$ posee las siguientes propiedades:*

(i) $P(s+t) = P(s)P(t)$ para toda $0 \leq s \leq t < \infty$ (Propiedad de semigrupo).

(ii) $(P(t) : t \geq 0)$ *es la única solución de la ecuación retrospectiva*

$$\frac{d}{dt}P(t) = P(t)Q \quad con \quad P(0) = \mathrm{I}.$$

(iii) $(P(t) : t \geq 0)$ *es la única solución de la ecuación prospectiva*

$$\frac{d}{dt}P(t) = QP(t) \quad con \quad P(0) = \mathrm{I}.$$

(iv) para $k = 0, 1, 2, \ldots$ *tenemos*

$$\left(\frac{d}{dt}\right)^k \Big|_{t=0} P(t) = Q^k.$$

Demostración:

(i) Definamos $P(t) = e^{tQ}$. De manera general, para toda $s,t \in [0,\infty)$ con $s \leq t$ tenemos que las matrices sQ y tQ conmutan (ver Apéndice A). Entonces se tiene que

$$P(s)P(t) = e^{sQ}e^{tQ} = e^{(s+t)Q} = P(s+t),$$

probando la propiedad de semigrupo.

(ii) En el Apéndice A, se demuestra que para toda $i, j \in I$, la serie de potencias dada por

$$\sum_{k=0}^{\infty} \frac{(tQ)_{ij}^k}{k!} \tag{2.1.3}$$

converge uniformemente y tiene radio de convergencia infinito. Definamos $p_{ij}(t) = (e^{tQ})_{ij}$ dado por la Ecuación (2.1.2) y por el Teorema 2.1.3, $p'_{ij}(t)$

existe y está dada por

$$
\begin{aligned}
p'_{ij}(t) &= \sum_{k=0}^{\infty} \frac{kt^{k-1}q_{ij}^{(k)}}{k!} \\
&= \sum_{k=0}^{\infty} \frac{t^{k-1}q_{ij}^{(k)}}{(k-1)!} \\
&= \sum_{k=0}^{\infty} \frac{t^{k-1}\sum_j q_{ij}^{(k-1)} q_{ij}}{(k-1)!} \\
&= \sum_j \left(\sum_{k=0}^{\infty} \frac{t^{k-1} q_{ij}^{(k-1)}}{(k-1)!} \right) q_{ji}, \qquad (2.1.4)
\end{aligned}
$$

pero $p_{ij}(t) = \sum_{k=0}^{\infty} \frac{t^{k-1}\sum_j q_{ij}^{(k-1)}}{(k-1)!}$ lo cual implica que la ecuación (2.1.4) es equivalente a que

$$p'_{ij}(t) = \sum_j p_{ij}(t) q_{ji}.$$

Por lo tanto, por lo anterior se sugiere que podemos diferenciar término por término la serie de matrices que define a $P(t)$ de manera que

$$
\begin{aligned}
\frac{d}{dt}P(t) &= \frac{d}{dt} \sum_{k=0}^{\infty} \frac{(tQ)^k}{k!} \\
&= \sum_{k=0}^{\infty} \frac{t^{k-1}Q^k}{(k-1)!} \\
&= \left(\sum_{k=0}^{\infty} \frac{t^{k-1}Q^{k-1}}{(k-1)!} \right) Q \\
&= P(t)Q,
\end{aligned}
$$

es decir, la derivada de la función matriz $P(t)$ existe y es igual a $P(t)Q$ demostrando que $P(t)$ satisface la ecuación retrospectiva como se requiere.

Una manera simple para mostar que e^{tQ} es invertible se debe a que como Q y $-Q$ conmutan tenemos que $e^{tQ}e^{-tQ} = \mathrm{I}$, donde I es la matriz identidad y mostrando que $(e^{tQ})^{-1} = e^{-tQ}$. En particular, como el radio de convergencia de la serie dada por la Ecuación (2.1.1) es infinito, en particular, se tiene que e^{-tQ} está bien definida y existe. Ahora supongamos que existe $M(t)$ y que también satisface la ecuación retrospectiva con la condición $M(0) = \mathrm{I}$. Consideremos lo siguiente:

$$\begin{aligned}\frac{d}{dt}(M(t)e^{-tQ}) &= \frac{d}{dt}M(t)e^{-tQ} + M(t)\frac{d}{dt}(e^{-tQ})\\ &= M(t)Qe^{-tQ} - M(t)Qe^{-tQ}\\ &= \mathrm{O},\end{aligned}$$

donde O denota a la matriz cuyos elementos son todos ceros. Por lo anterior y debido a que e^{tQ} es una matriz invertible, se sigue que

$$M(t)e^{-tQ} = (c)\mathrm{I}$$

de manera que si multiplicamos por e^{tQ} por la derecha de ambos lados de la igualdad obtenemos que

$$\begin{aligned}M(t) &= (c)e^{tQ}\\ &= (c)P(t).\end{aligned}$$

Como $M(t)$ es otra solución a la ecuación retrospectiva con condición inicial $M(0) = \mathrm{I}$, se tiene que $c = 1$ para así obtener que $M(t) = P(t)$, es decir, $P(t)$ es la única solución a la ecuación retrospectiva.

(iii) Demostración análoga al punto (ii).

(iv) Observemos que $P(0) = I = Q^0$. Procederemos con una demostración por inducción para $k = 1, 2, \ldots$; caso base $k = 1$, del punto (ii) sabemos que $\frac{d}{dt}P(t) = P(t)Q$ y $\frac{d}{dt}P(t)|_{t=0} = Q$, por lo tanto la proposición es válida para el caso base. Ahora, supongamos que $\left(\frac{d}{dt}\right)^n P(t) = P(t)Q^n$, se procede a demostrar que

$$\left(\frac{d}{dt}\right)^{n+1} P(t)\Big|_{t=0} = Q^{n+1}.$$

Si calculamos $\left(d/dt\right)^{n+1}P(t) = d/dt(P(t)Q^n)$:

$$\frac{d}{dt}\left(\left(\frac{d}{dt}\right)^n P(t)\right) = \frac{d}{dt}\left(P(t)Q^n\right) \tag{2.1.5}$$

$$= \left(\frac{d}{dt}P(t)\right)Q^n \tag{2.1.6}$$

$$= P(t)QQ^n \tag{2.1.7}$$

$$= P(t)Q^{n+1};$$

(2.1.5) se debe por la hipótesis de inducción, por la regla de la cadena obtenemos (2.1.6) de manera que por el inciso anterior se sigue (2.1.7), demostrando que $\left(\frac{d}{dt}\right)\Big|_{t=0} P(t) = Q^k$ para $k = 0,1,2,...$ como se requiere.

□

Se recuerda la definición de una matriz función estocástica dada por 2.1.2. Recurrimos a la notación $f(t) = o(t)$ implica que $f(t)/t \to 0$ cuando $t \to 0$. El siguiente teorema nos permitirá darle una interpretación a la matriz $P(t)$:

Teorema 2.1.5. *Una matriz Q definida en un conjunto finito I es una $Q-matriz$ si y solo si $P(t) = e^{tQ}$ es una matriz estocástica para toda $t \geq 0$.*

Demostración:
Sea I la matriz identidad de dimensión $n \times n$ y sea Q una $Q-matriz$ de dimensión $n \times n$, definamos

$$B = Q + m\mathrm{I} \qquad \text{con} \quad m = sup_{i,j}|q_{ij}|. \tag{2.1.8}$$

Como Q está definida para un conjunto finito I por definición de una $Q-matriz$ (ver Definición 2.1.1) existe una entrada (h,k) de la matriz Q con $m = q_{hk} < \infty$ para alguna $h,k \in I$ mostrando que la matriz B existe y está bien definida. En particular, por la ecuación (2.1.8), las entradas b_{ij} de la matriz B son iguales a:

$$b_{ij} = q_{ij} + m\,\delta_{ij}. \tag{2.1.9}$$

Ahora, si Q es una $Q-matriz$ para toda $i \neq j$ en I tenemos que $b_{ij} = q_{ij} + m\delta_{ij} > q_{ij} \geq 0$. Para $i = j$, dada una Q-matriz cada entrada de la diagonal se define como $q_{ii} = -\sum_{i\neq j} q_{ij}$, debido a este hecho, existe $i \in I$ tal que m se alcanza en la diagonal de Q, es decir, $m = sup_i|q_i| \geq 0$ de manera que $b_{ii} = q_{ii} + |q_{ii}| \geq 0$. Como todas las entradas de la matriz $B = (b_{ij})_{i,j\in I}$ son positivas, las entradas de la matriz exponencial de B denotada por e^B satisfacen que

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{b_{ij}^{(k)}}{k!} = (e^B)_{ij} \geq 0,$$

donde $b_{ij}^{(k)}$ denota la (i,j) entrada del matriz B^k. Debido a esto $(e^B)_{ij} \geq 0$ pues dada una matriz de entradas positivas B se cumple B^k también tiene entradas positivas. La matriz exponencial de la matriz identidad es igual

$$e^{\mathrm{I}} = \sum_{k=0}^{\infty} \frac{I^k}{k!} = \sum_{k=0}^{\infty} \frac{\mathrm{I}}{k!} = \left(\sum_{k=0}^{\infty} \frac{1}{k!} \right) \mathrm{I} = e\,\mathrm{I},$$

con e la contante de Napier. Sea la matriz exponencial de la matriz Q. Por la igualdad 2.1.8 y 2.1.9 se tiene que $tQ = t(B - m\mathrm{I})$ para toda $t \geq 0$. Como $-tmB\mathrm{I} = -tm\mathrm{I}B$ pues $-tm$ es una escalar y la matriz identidad conmuta con cualquier matriz en el producto de matrices, Tenemos que la matriz

$$\begin{aligned} P(t) = e^{tQ} &= e^{t(B-m\mathrm{I})} \\ &= e^{tB}\, e^{-tm\mathrm{I}} \\ &= e^{tB}\, e^{-tm}\, \mathrm{I} \\ &= e^{tB} e^{-tm}. \end{aligned}$$

Por lo anterior, e^{tB} es una matriz de entradas no negativas siempre y cuando $t \geq 0$ y e^{-tm} es una escalar no negativa para concluir que $p_{ij}(t) \geq 0$ para toda $i, j \in I$. Finalmente, como Q es $Q - matriz$ se tiene que para todo natural n los renglones de Q^n suman cero pues

$$\sum_{k\in I} q_{ik}^{(n)} = \sum_{k\in I} \sum_{j\in I} q_{ij}^{(n-1)} q_{jk} = \sum_{j\in I} q_{ij}^{(n-1)} \sum_{k\in I} q_{jk} = 0,$$

donde $q_{ij}^{(n)} = (Q^n)_{ij}$. De manera que

$$\sum_{j\in I} p_{ij}(t) = 1 + \sum_{k=1}^{\infty} \frac{t^n}{n!} \sum_{j\in I} q_{ij}^{(n)} = 1,$$

demostrando que $P(t)$ es una matriz estocástica para toda $t \geq 0$.
Para probar la suficiencia de la afirmación, sabemos ahora $\sum_{j\in I} p_{ij}(t) = 1$ para toda $t \geq 0$. Esto implica

$$\sum_{j\in I} q_{ij} = \frac{d}{dt} \sum_{j\in I} p_{ij}(t)\Big|_{t=0} = \frac{d}{dt} 1\Big|_{t=0} = 0.$$

Cuando $t \to 0$, se tiene que

$$P(t) = I + tQ + 0(t^2), \tag{2.1.10}$$

de manera $q_{ij} \geq 0$ para $i \neq j$ si y sólo si $p_{ij}(t) \geq 0$ para toda i, j y $t \geq 0$ suficientemente pequeña, este hecho se demuestra de manera precisa en la Sección 2.4. Por la propiedad de semigrupo $P(t) = P(t/n)^n$ y se deduce por (2.1.10) $q_{ij} < \infty$ si $p_{ij}(t/n) < \infty$. Por lo tanto, se sigue que $q_{ij} \geq 0$ para toda $i \neq j$ si y sólo si $p_{ij}(t) \geq 0$ para toda $i, j \in I$ y toda $t \geq 0$. Por lo tanto, si $P(t)$ es estocástica para toda $t \geq 0$ entonces Q es $Q - matriz$. □

Observación 2.1.4. Por el inciso (i) del Teorema 2.1.4 en conjunto con el Teorema 2.1.5, la matriz e^{tQ} define una matriz función de transición siempre y cuando Q sea una $Q - matriz$. Adicionalmente, por los incisos $(ii), (iii)$ y (iv) del Teorema 2.1.4, se intuye que

$$q_i = -p'_{ii}(0) \quad \text{para } i = j,$$
$$q_{ij} = p'_{ij}(0) \quad \text{para } i \neq j.$$

Para terminar esta sección, pensemos ahora en el espacio de parámetros discreto $0, 1, 2, ...$ como un subespacio encajado en $[0, \infty)$. Una forma natural de interpolar la suceción discreta $(p^n : n = 0, 1, ...)$ es ocupando la función $(e^{tq} : t \geq 0)$ donde $q = log(p)$. Por otro lado, dado un proceso de Markov a tiempo discreto y un conjunto finito I, sabemos que una manera de calcular la probabilidad de saltar del estado i al estado j es $\mathbb{P}_{ij}(n) = p^{(n)}_{ij}$ donde P es una matriz estocástica. Si tratamos de tomar esta idea para dar una versión continua de los procesos Markov, suponiendo que Q es una $Q - matriz$ donde $e^Q = P$ y si definimos la sucesión $(P^n : n = 0, 1, ...)$ donde $P^n = (e^Q)^n = e^{nQ}$, el conjunto de matrices

$$(e^{tQ} : t \geq 0)$$

rellenará los huecos de la sucesión discreta de matrices antes mencionada.

Definamos $P(t) = e^{tQ}$, en la Sección 2.4 se muestra que una cadena de Markov a tiempo continuo $(X_t)_{t\geq 0}$ con probabilidades de transición estacionarias generadas por una Q-matriz Q satisface que

$$\mathbb{P}_i(X_{t_{n+1}} = i_{n+1} \,|\, X_{t_0} = i_0, \ldots, X_{t_n} = i_n) = p_{i_n\, i_{n+1}}(t_{n+1} - t_n)$$

para toda $n \in \mathbb{Z}^+$, todo tiempo $0 \leq t_0 \leq \ldots \leq t_{n+1}$ y todo estado $i_0, \ldots, i_n$, donde p_{ij} es la ij-ésima entrada de la matrix exponencial e^{tQ}. En particular, la función de probabilidad de transición de i a j al tiempo t está dada por

$$\mathbb{P}_i(X_t = j) := \mathbb{P}(X_t = j | X_0 = i) = p_{ij}(t).$$

2.2 Especificaciones Técnicas

Sea I un conjunto numerable. Un proceso continuo por la derecha

$$(X_t)_{t\geq 0} = \{X_t \,:\, 0 \leq t < \infty\}$$

con valores en I es una familia de variables aleatorias $X_t : \Omega \to I$. Se consideran las formas para especificar el comportamiento probabilístico (o ley) para un proceso $(X_t)_{t\geq 0}$. Esto nos permitirá encontrar, al menos en principio, cualquier probabilidad asociada al proceso, tales como $\mathbb{P}(X_t = i)$ o $\mathbb{P}(X_{t_0} = i_0, \ldots, X_{tn} = i_n)$, o $\mathbb{P}(X_t = i \text{ para alguna } t \geq 0)$. Una de las dificultades que uno presenta al estudiar esta clase de procesos es que a diferencia del caso discreto, dada la unión numerable disjunta de eventos se cumple que

$$\mathbb{P}\left(\bigcup_n A_n\right) = \sum_n \mathbb{P}(A_n),$$

mientras que para la unión no numerable $\bigcup_{t\geq 0} A_t$ no se cumple semejante regla. Para evitar estas dificultades sólo enfocamos nuestra atención a los procesos $(X_t)_{t\geq 0}$ continuos por la derecha con espacio de estados numerable I de manera que asumimos esta propiedad. En este contexto, si I es un conjunto discreto, si un proceso estocástico es continuo a la derecha se cumple que para toda $\omega \in \Omega$ y $t \geq 0$, entonces existe $\varepsilon > 0$ tal que

$$X_s(\omega) = X_t(\omega) \quad \text{para toda } t \leq s \leq t + \varepsilon. \tag{2.2.1}$$

Por resultados de la teoría de procesos estocásticos, la probabilidad de que cualquier evento asociado a un proceso estocástico continuo por la derecha puede ser determinada por la distribución finito dimensional del proceso; para toda $n \geq 0$, $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$ e $i_1, i_2, \ldots, i_n \in I$ entonces

$$\mathbb{P}(X_{t_0}(\omega) = i_0, X_{t_1}(\omega) = i_1, \ldots, X_{t_n}(\omega) = i_n).$$

Este hecho se debe a que la probabilidad de los eventos de la forma

$$\{X_{t_0} = i_0, \ldots, X_{t_n} = i_n\}$$

para toda $n \geq 0$, determinan las probabilidad de los eventos que dependen de $(X_t)_{t\geq 0}$. En particular, los eventos de la forma $\{X_{t_0} = i_0, \ldots, X_{t_n} = i_n\}$ forman un π-sistema que genera la σ-álgebra $\sigma(X_t : t \geq 0)$. El Teorema 1.2.1.5 justifica de manera precisa esta afirmación. En particular, sin la suposición de la continuidad por la derecha la probabilidad:

$$\mathbb{P}\big(X_t = i \text{ para alguna } t \in [0,\infty)\big) = \lim_{n\to\infty} \mathbb{P}(X_{q_1} = i, \ldots, X_{q_n} = i),$$

donde $q_1, q_2, \ldots$ es una enumeración de racionales sería inconcebible sin este supuesto.

Sea I un espacio numerable de estados. Definamos $(X_t)_{t\geq 0}$ un proceso continuo por la derecha, por la propiedad (2.1.1) toda trayectoria $t \to X_t(\omega)$ de un proceso continuo por la derecha implica que se mantiene constante en los valores $i \in I$ que toma el proceso X_t, es decir, dado $X_0 = i_0$ el proceso se mantiene en el estado i_0 digamos un tiempo S_1 y luego salta al tiempo J_1 al estado i_1, es decir, $X_{J_1} = i_1$ y se mantiene constante un intervalo de tiempo S_2, y así sucesivamente. Para referirnos al tiempo de salto nos referimos al instante en que el proceso salta o cambia de estado *(i.e.,realiza una transición)*.

En particular, si consideramos $(\Omega, \mathscr{F}, \mathbb{P})$ con $\mathscr{F} = \mathscr{F}^X$ *(ver Sección 1.2 Definición 1.2.2.4)* la σ-álgebra generada por el proceso $(X_t)_{t\geq 0}$, los tiempos de salto J_n y de espera S_n del proceso $(X_t)_{t\geq 0}$ se obtienen definiendo;

Definición 2.2.1. *Para $n = 0, 1, \ldots$, los tiempos de espera $J_n(\omega)$ del proceso $(X_t)_{t\geq 0}$ son variables aleatorias que se definen como*

$$J_0(\omega) = 0, \qquad J_{n+1}(\omega) = \inf\{t \geq J_n(\omega) : X_t(\omega) \neq X_{J_n}(\omega)\},$$

donde $\inf \emptyset = \infty$.

Definición 2.2.2. *De manera similar, los tiempos de espera $S_n(\omega)$ del proceso $(X_t)_{t\geq 0}$ con $n = 1, 2, \ldots$ son variables aleatorias se definen como:*

$$S_n(\omega) = \begin{cases} J_n(\omega) - J_{n-1}(\omega) & \text{si} \quad J_{n-1}(\omega) < \infty, \\ \infty & \text{en cualquier otro caso}, \end{cases}$$

de manera que

$$\{\omega : t \in S_n(\omega)\} = \{\omega : X_t(\omega) = X_{J_{n-1}}(\omega), J_{n-1}(\omega) \leq t < J_n(\omega)\}.$$

Tomando estas consideraciones en cuenta, toda trayectoria $t \to X_t$ de un proceso estocástico continuo por la derecha $(X_t)_{t\geq 0}$ con espacio de estados numerable finito en el intervalo $[0,\infty)$ puede comportase como se describe a continuación:

(a) Para cualquier intervalo cerrado de la forma $[s,t]$ el proceso realiza un número finito de saltos.

(b) El proceso realiza un número finito de saltos y es atrapado por algún estado j de manera que nunca sale de ahí.

(c) El proceso realiza un número infinito de saltos en un intervalo finito. En este caso, diremos que el proceso explota en el tiempo ζ haciendo referencia al hecho de que el modelo pueda llegar a presentar una inconsistencia al instante ζ. Después de ζ el proceso reinicia de nuevo; puede explotar una o varias veces, o puede no explotar.

Observemos que debido a la continuidad por la derecha de $(X_t)_{t\geq 0}$, en cualquiera de los casos anteriores de manera forzosa $S_n > 0$ para toda $n = 1,2,\ldots$. En el caso (b), se puede decir que $J_{n+1} = \infty$ para alguna n (y en particular para alguna $j \in I$) de manera que definimos $X_\infty = X_{J_n} = j$, el último valor que tomo el proceso de $(X_t)_{t\geq 0}$ de forma que X_∞ pueda ser especificado. En el caso (c) puede suceder que los tiempos de espera S sean cada vez más pequeños de tal forma que $\sum_{n=1}^{\infty} S_n < \infty$, es decir, existe la posibilidad de que el proceso efectúe un número infinito de saltos durante un intervalo de tiempo acotado de manera que la sucesión de variables aleatorias S_n converge a 0.

Definición 2.2.3. *El tiempo de explosión del proceso $(X_t)_{t\geq 0}$ se define como*

$$\zeta = \sup_{0\leq n} J_n = \sum_{n=1}^{\infty} S_n.$$

Definición 2.2.4. *El proceso discreto $(Y_n)_{n\geq 0}$ inducido por la sucesión de saltos dada por $Y_n = X_{J_n}$ es llamado el proceso de saltos de $(X_t)_{t\geq 0}$.*

Si $(Y_n)_{n\geq 0}$ cumple ser es una cadena de Markov a tiempo discreto entonces $(X_t)_{t\geq 0}$ es una cadena de Markov a tiempo continuo *(ver Sección 2.4)*, de manera que al referirnos a una cadena de saltos nos referimos al proceso de saltos asociado a la cadena.

Un hecho que se asume es que el conjunto de espacio de estados I es un conjunto numerable finito o infinito, otros casos se pueden pensar como generalizaciones del caso numerable, sin embargo, no abordamos esta discusión y nos enfocamos a estudiar el caso más simple; un conjunto discreto I dotado con la topología discreta. Para el caso en que I es infinito, se sugiere compactificar el espacio (ver [Chu60, Capítulo II, Sección 4]) como sigue; adjuntamos el estado ∞ a I definiendo $\bar{I} = I \cup \{\infty\}$ y postulando que todo subconjunto infinito de I es cerrado si y solo sí contiene a ∞. Al conjunto $\bar{I}$ lo llamaremos como el *conjunto minimal de estados* Se define $p_{i\infty}$ como la probabilidad de que la cadena se encuentre en ∞ como

$$p_{i\infty} := 1 - \sum_{j \in I} p_{ij}, \tag{2.2.2}$$

de manera que si $p_{i\infty} > 0$ la cadena se encuentra en ∞ con probabilidad $p_{i\infty}$.

Definición 2.2.5. *Sea I un espacio de estados minimal. Definamos ζ como la primera explosión del proceso $(X_t)_{t\geq 0}$. El proceso minimal $(Y_t)_{t\geq 0}$ asociado a $(X_t)_{t\geq 0}$ se define como*

$$Y_t(\omega) = \begin{cases} X_t(\omega) & si \quad t < \zeta(\omega), \\ \infty & si \quad \zeta(\omega) < t. \end{cases}$$

Observación 2.2.1. Con respecto a la definición anterior, claramente el proceso $(Y_t)_{t\geq 0}$ es igual al proceso $(Y_t)_{t\geq 0}$ en el intervalo $[0, \zeta)$ y cambia fuera de el. Si el proceso no explota, esta modificación es igual al proceso inicial, sin embargo, la definición anterior no hace alusión al estado del proceso sino al intervalo de tiempo en el que el proceso se mantiene activo.

Para objetos de nuestro estudio no consideramos lo que sucede después del tiempo de explosión y sólo consideramos el tiempo en el que el proceso se mantiene activo, es decir, lo que sucede en el intervalo $[0, \zeta)$. Con esto en mente, es conveniente adjuntar al espacio de estados el estado ∞. Si $\zeta < \infty$, definimos la igualdad $X_t = \infty$ para toda $t > \zeta$. Al referirnos a un proceso minimal asumimos que se considera la "versión" minimal del proceso de manera que todo proceso enunciado a partir de este momento se considera minimal. Notemos que un proceso continuo por la derecha $(X_t)_{t\geq 0}$ minimal, se puede reconstruir con el proceso de saltos y los tiempos de espera del proceso especificando la distribución conjunta de $S_1, S_2, \ldots$ y $(Y_n)_{n\geq 0}$, mostrando otra especificación numerable del comportamiento probabilístico de $(X_t)_{t\geq 0}$. Un ejemplo de este hecho es que podemos calcular la probabilidad de $X_t = i$ en términos de las probabilidades de salto del proceso $(Y_n)_{n\geq 0}$ como se muestra a continuación:

$$\mathbb{P}(X_t = i) = \sum_{n=0} \mathbb{P}(Y_n = i, J_n \leq t < J_{n+1}).$$

2.3 Cadena de saltos

Específicamente, bajo ciertas condiciones las trayectorias que define una cadena de Markov a tiempo continuo son funciones escalonadas, este resultado se debe gracias a Doob (ver [Doo90, Capítulo 4, Sección 1, Teorema 1.4]. En particular, debido a este resultado, toda cadena de Markov a tiempo continuo con parámetro de tiempo $t \in \mathrm{T}$ con $\mathrm{T} = \mathbb{R}^+ = [0,\infty)$ o $\mathrm{T} = \overline{\mathbb{R}}^+ = [0,\infty]$ cumple esta propiedad de manera que asumimos este resultado. De forma precisa, dada una distribución inicial λ y una matriz función de transición $P(t)$, existen distintas versiones de un proceso de Markov $(X_t)_{t\geq 0}$ de manera que podemos especificar con libertad una versión o modificación particular del proceso (X_t) cuyas trayectorias son iguales con probabilidad uno y poseen las mismas propiedades deseadas que el proceso inicial $(X_t)_{t\geq 0}$. En la siguiente sección detallamos una construcción de una versión continua por la derecha del proceso que cumple la propiedad de Markov.

En esta sección se procede a construir las cadenas de Markov a tiempo continuo, en la sección anterior vimos que un proceso estocástico minimal continuo por la derecha se puede reconstruir a través de el proceso de saltos y tiempos de espera asociados a un proceso.

Sea I un conjunto numerable, la información básica para una cadena de Markov a tiempo continuo en I está dada en la forma de una $Q-matriz$. Definamos $q(i) = q_i = -q_{ii}$.

Definición 2.3.1. *La matriz de salto* $\Pi = (\pi_{ij} : i, j \in I)$ *asociada a una Q-matriz Q de dimensión* $n\times n$*, es una matriz cuadrada de dimensión* $n\times n$ *cuyas entradas* *se definen como sigue*

$$\pi_{ij} = \begin{cases} \frac{q_{ij}}{q_i} & si \quad i \neq j \; y \; q_i \neq 0 \\ 0 & si \quad i \neq j \; y \; q_i = 0, \end{cases}$$

$$\pi_{ii} = \begin{cases} 0 & si \quad q_i \neq 0 \\ 1 & si \quad q_i = 0, \end{cases}$$

con $q_{i,j}, q_{ii} \in (q_{ij})_{i,j} \in I$ *y* $q_i = -q_{ii}$.

Sea un proceso $(X_t)_{t\geq 0}$ continuo por la derecha, sean $J_0, J_1, J_2, ...$ y $S_1, S_2, ...$ los tiempos salto y de espera respectivamente del proceso $(X_t)_{t\geq 0}$ como se definen en las definiciones 2.2.1 y 2.2.2 respectivamente. Se recuerda que asumimos $(X_t)_{t\geq 0}$ minimal (ver Definición 2.2.5) y se considera el proceso de saltos asociado a $(X_t)_{t\geq 0}$ (ver Definición 2.2.4) denotado por $(Y_n)_{n\geq 0}$. Adicionalmente, se denota por $(Y_n)_{n\geq 0} \sim MarkovTd(\lambda, P)$ a una cadena de Markov a tiempo discreto con distribución inicial λ y matriz estocástica P. Una definición alternativa para las cadenas de Markov a tiempo continuo en términos de sus tiempos de espera y proceso de saltos se da a continuación:

Definición 2.3.2. *Sea* $(X_t)_{t\geq 0}$ *un proceso minimal a tiempo continuo con valores en* I. *El proceso* $(X_t)_{t\geq 0}$ *es una cadena de Markov con distribución inicial* λ *y matriz generadora* Q *si la cadena de saltos asociada* $(Y_n)_{n\geq 0}$ *es* $MarkovTc(\lambda, \Pi)$ *y los tiempos de espera* $S_1, ..., S_n$ *dados* $Y_0, Y_1, ..., Y_{n-1}$ *son variables aleatorias independientes con distribución exponencial de parámetros* $q(Y_0), ..., q(Y_{n-1})$ *respectivamente. De manera resumida diremos que* $(X_t)_{t\geq 0}$ *es* $Markov(\lambda, Q)$.

En este capítulo al referirnos a un proceso $(X_t)_{t\geq 0} \sim Markov(\lambda, Q)$ asumimos que $(X_t)_{t\geq 0}$ es una cadena de Markov a tiempo continuo, sin embargo, en las secciones posteriores para distinguir el caso continuo del caso discreto denotaremos por $(X_t)_{t\geq 0} \sim MarkovTc(\lambda, Q)$ al caso continuo y por $(Y_n)_{t\geq 0} \sim MarkovTd(\lambda, P)$ para el caso discreto. Dada esta definición podemos ser capaces de reconstruir una cadena de Markov $(X_t)_{t\geq 0}$ con distribución inicial λ y matriz generadora Q. A continuación se proponen las siguientes construcciones de una cadena de Markov a tiempo continuo:

Construcción 2.3.1. Sea $(Y_n)_{n\geq 0}$ tal que es $MarkovTd(\lambda, \Pi)$ y $T_1, T_2, ...$ una sucesión de variables aleatorias independientes exponenciales de parámetro 1, cada una independiente del proceso $(Y_n)_{n\geq 0}$. Definamos $S_n = \frac{T_n}{q(Y_{n-1})}$, $J_n = S_1 + ... + S_n$ y

$$X_t = \begin{cases} Y_n & \text{si } \ J_n \leq t < J_{n+1} \text{ para alguna n,} \\ \infty & \text{en otro caso.} \end{cases}$$

Por definición $(Y_n)_{n\geq 0}$ es $MarkovTd(\lambda, \Pi)$ y como para toda $n = 0, 1, 2, ...$, $T_n \sim exp(1)$ implica que $S_n \sim exp(q(Y_{n-1}))$, por lo tanto $(X_t)_{t\geq 0}$ es $Markov(\lambda, Q)$ donde

$$q_{ij} = \pi_{ij} q_i \quad \text{para toda } i \neq j$$

tal que $i = Y_k$ para alguna $k = 0, 1, 2, ...$.

A continuación se describirá una segunda construcción, pero antes es necesario enunciar el siguiente teorema:

Teorema 2.3.3. *Sea I un conjunto numerable. Para cada $k \in I$ definamos a a la variable aleatoria T_k donde $T_k \sim exp(q_k)$ y $0 < q := \sum_{k\in I} q_k < \infty$. Sea $T = inf_k T_k$. Entonces el ínfimo se alcanza en un valor aleatorio único K de k, con probabilidad 1. Además, T y K son independientes con $T \sim exp(q)$ y $\mathbb{P}(K = k) = q_k/q$.*

Demostración:
Sea $K = k$ si $T_k < T_j$ para toda $j \neq k$, en otro caso K no está definida, entonces:

$$\begin{aligned}
\mathbb{P}(K = k, T \geq t) &= \mathbb{P}(T_k \geq t, T_j > T_k \quad \text{para toda } j \neq k) \\
&= \mathbb{P}(T_k \geq t)\mathbb{P}(T_j > T_k \quad \text{para toda } j \neq k | T_k \geq t) \\
&= \int_t^\infty q_k e^{-q_k s} \mathbb{P}(T_j > s \quad \text{para toda } j \neq k) ds \\
&= \int_t^\infty q_k e^{-q_k s} \prod_{j\neq k} e^{-q_j s} ds \\
&= \int_t^\infty q_k e^{-qs} ds \\
&= \frac{q_k}{q} e^{-qt}.
\end{aligned}$$

Así, la densidad conjunta es el producto de las densidades

$$\mathbb{P}(K = k, T = t) = q_k e^{-qt} = \left(\frac{q_k}{q}\right)(q e^{-qt})$$

y

$$\mathbb{P}(K = k \text{ para alguna } k \in I) = \sum_{k\in I} \frac{q_k}{q} \int_0^\infty q e^{-qt} dt = 1.$$

Por lo tanto T y K son independientes con $T \sim exp(q)$ y $\mathbb{P}(K=k)=\frac{q_k}{q}$. $\square$

Construcción 2.3.2. Está construcción hace uso directo de las entradas de una $Q-matriz$ dada. Empezemos con un estado inicial $X_0 = Y_0$ con distribución λ y sea una colección $(T_n^j :\geq 1, j \in I)$ de variables exponenciales independientes de parámetro 1. De manera inductiva para $n = 0,1,...$, si $Y_n = i$ definamos:

$$\begin{aligned}
S_{n+1}^j &= T_{n+1}^j / q_{ij} \quad \text{para } i \neq j, \\
S_{n+1} &= inf_{j\neq i} S_{n+1}^j, \\
Y_{n+1} &= \begin{cases} j & \text{si } S_{n+1}^j = S_{n+1} \\ i & \text{si } S_{n+1} = \infty. \end{cases}
\end{aligned}$$

Entonces, si $Y_n = i$, las variables aleatorias S_{n+1}^j son variables aleatorias exponenciales independientes de parámetros q_{ij} para toda $i \neq j$. Así que, condicionando sobre $Y_n = i$, por el teorema anterior, S_{n+1} es exponencial de parámetro $q_i = \sum_{i\neq j} q_{ij}$, la distribución de Y_{n+1} está dada por $(\pi_{ij} : j \in I)$. Las variables aleatorias S_{n+1} y Y_{n+1} son independientes entre sí e independientes de $Y_0,...,Y_n$ y $S_1,...,S_n$, según lo requerido. Por lo que si definimos a $(J_n)_{n=0}$ donde $J_0 = 0$ y $J_m = \sum_{n=1}^m S_n$ entonces tenemos que:

$$X_t = \begin{cases} Y_n & \text{si } \; J_n \leq t < J_{n+1} \text{ para alguna n}, \\ \infty & \text{en otro caso.} \end{cases}$$

2.4 Ecuaciones prospectivas y retrospectivas

Consideremos un proceso $(X_t)_{t\geq 0}$ continuo por la derecha con valores en un conjunto numerable I. Sea $\mathscr{F}_t$ la σ-álgebra generada por $\{X_s : s \leq t\}$, es decir, generada por los eventos $\{X_s = i\}$ donde $s \leq t$ e $i \in I$.

Definición 2.4.1. *Sea $(\Omega, \mathscr{F})$ un espacio medible con una filtración $(\mathscr{F}_t)_{t\geq 0}$ como se define en 1.2.2.2. Una variable aleatoria $T : \Omega \to [0,\infty]$ es un tiempo de paro con respecto a la filtración $(\mathscr{F}_t)_{t\geq 0}$ si*

$$\{T \leq t\} = \{\omega : T(\omega) \leq t\} \in \mathscr{F}_t \quad \textit{para toda} \quad t \in [0,\infty).$$

Intuitivamente, un tiempo de paro es un tiempo aleatoria que depende únicamente de la historia de algún proceso al tiempo t y no de lo que ocurre después de ese instante. Se recuerda que

$$\delta_{ij} = \begin{cases} 1 & \text{si } j = i, \\ 0 & \text{en otro caso,} \end{cases}$$

donde $(\mathrm{I})_{ij} = \delta_{ij}$ e I es la matriz identidad. Para objeto de las siguientes demostraciones se enuncia el siguiente teorema el cual se demuestra en el Teorema 2.5.5.

Teorema 2.4.2. *Sea $(X_t)_{t\geq 0}$ Markov(λ, Q) y T un tiempo de paro de $(X_t)_{t\geq 0}$. Entonces, condicionalmente a $\{T < \infty\}$ y $\{X_T = i\}$, $(X_{T+t})_{t\geq 0}$ es Markov(δ_i, Q) e independiente de $(X_s : s \leq T)$.*

En la sección anterior, nos enfocamos en construir las cadenas de Markov a tiempo continuo basándonos en las propiedades de la cadena de saltos y los tiempos de espera asociados a algún proceso $(X_t)_{t\geq 0}$ que cumple la propiedad de Markov. En particular, estas construcciones se basan en métodos y concepctos ligados a las cadenas de Markov a tiempo discreto aportando una idea clara de la evolución del proceso. Sea $(X_t)_{t\geq 0}$ un proceso continuo por la derecha con valores en un conjunto numerable I. A continuación presentamos la primera caracterización de las cadenas de Markov a tiempo continuo en el que consideraremos el caso finito del espacio de estados I. Para la demostración de este teorema se adopta la notación de $f(t) = o(t)$ implica que $f(t)/t \to 0$ cuando $t \to 0$.

Teorema 2.4.3. *Sea un proceso minimal continuo por la derecha $(X_t)_{t\geq 0}$ un proceso continuo por la derecha con valores en un conjunto finito I. Sea Q una $Q-matriz$ definida en I con matriz de saltos Π. Entonces las siguientes proposiciones son equivalentes:*

(a) (Definición de Cadena de saltos y tiempos de espera) El proceso $(X_t)_{t\geq 0}$ es una cadena de Markov con distribución inicial λ y matriz generadora Q si la cadena de saltos asociada $(Y_n)_{n\geq 0}$ es $MarkovTc(\lambda, \Pi)$ y los tiempos de espera $S_1, \ldots, S_n$ dados $Y_0, Y_1, \ldots, Y_{n-1}$ son variables aleatorias independientes con distribución exponencial de parámetros $q(Y_0), \ldots, q(Y_{n-1})$ respectivamente. De manera resumida diremos que $(X_t)_{t\geq 0}$ es Markov(λ, Q).

(b) (definición infinitesimal) para toda $t, h \geq 0$ condicionando en $X_t = i$, X_{t+h} es independiente de $(X_s : s \leq t)$ y cuando $h \to 0$, uniformemente en t, para toda j tenemos que

$$\mathbb{P}(X_{t+h} = j | X_t = i) = \delta_{ij} + q_{ij}h + o(h);$$

(c) (definición de probabilidad de transición) para toda $n = 0, 1, 2, ...$*, todos los tiempos* $0 \leq t_0 \leq t_1 \leq ... \leq t_{n+1}$ *y todos los estados* $i_0, ..., i_{n+1}$ *tenemos*

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, ..., X_{t_n} = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n)$$

donde $(p_{ij}(t) : i, j \in I, t \geq 0)$ *es la solución de la ecuación prospectiva*

$$P'(t) = P(t)Q, \qquad\qquad P(0) = \mathrm{I}.$$

Si $(X_t)_{t \geq 0}$ satisface cualquiera de estas tres condiciones entonces se dice que $(X_t)_{t \geq 0}$ es una *cadena de Markov con matriz generadora* Q y para referirnos a la desitribución finito dimensional de la cadena diremos que $(X_t)_{t \geq 0}$ es *Markov*(λ, Q), donde λ es la distribución de X_0.

Demostración:
(a) $\Rightarrow$ (b)
Supongamos la proposición (a). Sea $\mathscr{F}$ la σ-álgebra generada por $\{X_s : s \leq t\}$. Consideremos el evento $\{X_h = j\}$ y notemos que

$$\{X_h = j\} = \bigcup_{m=0}^{\infty} \{Y_m = j, J_{m-1} \leq h < J_m\},$$

de manera que para toda j en I los intervalos $[J_{m-1}, J_m)$ son todos ajenos entre sí. Dada esta observación, si calculamos la probabilidad $\mathbb{P}(X_h = j | X_0 = i) = \mathbb{P}_i(X_h = j)$ obtenemos que para toda j en I se cumple la siguiente igualdad

$$\mathbb{P}(X_h = j | X_0 = i) = \sum_{m=1}^{\infty} \mathbb{P}(Y_{m-1} = j, J_{m-1} \leq h < J_m | Y_0 = i). \tag{2.4.1}$$

De la ecuación anterior, si consideramos el caso $j = i$ y calculamos la probabilidad del evento $\{X_h = i\}$ dado $X_0 = i$ obtenemos que;

$$\begin{aligned}
\mathbb{P}(X_h = i | X_0 = i) &= \sum_{m=1}^{\infty} \mathbb{P}(Y_{m-1} = i, J_{m-1} \leq h < J_m | Y_0 = i) \\
&= \mathbb{P}_i(Y_0 = i, J_0 \leq h < J_1) \\
&\quad + \mathbb{P}_i(Y_1 = i, J_1 \leq h < J_2) \\
&\quad\quad + \sum_{m=3}^{\infty} \mathbb{P}_i(Y_{m-1} = i, J_{m-1} \leq h < J_m) \\
&= \mathbb{P}_i(h < J_1) \qquad\qquad (2.4.2) \\
&\quad + \sum_{m=3}^{\infty} \mathbb{P}_i(Y_{m-1} = i, J_{m-1} \leq h < J_m),
\end{aligned}$$

pues $\mathbb{P}_i(Y_0 = i) = 1$ y $\mathbb{P}_i(Y_1 = i) = 0$. Notemos que sin importar la elección de j la unión de los siguientes eventos cumple que

$$\begin{aligned}\bigcup_{m=3}^{\infty}\{Y_{m-1} = j, J_{m-1} \leq h < J_m\} &\subseteq \{S_1 + S_2 \leq h\} \\ &\subseteq \{S_1 \leq h, S_2 \leq h\},\end{aligned} \tag{2.4.3}$$

donde $S_1 = J_1$ y S_2 son variables aleatorias exponenciales independientes de parámetros $q(Y_0)$ y $q(Y_1)$ respectivamente. Si consideremos la matriz $Q = (q_{ij} : i, j \in I)$ los valores q_{ij} están acotados por $-q_{ii}$ respectivamente pues $-q_{ii} = \sum_{i \neq j} q_{ij}$. Si definimos a $q = \sup_{j \in I}(q_{jj} : q_{jj} \in Q)$, como I es finito tenemos que q existe e implica que $q = -q_k$ para alguna k en I y que

$$\begin{aligned}\mathbb{P}_i(h < J_1) + &\sum_{m=3}^{\infty} \mathbb{P}_i(Y_{m-1} = i, J_{m-1} \leq h < J_m) \\ &\leq \mathbb{P}_i(h < J_1) + \mathbb{P}_i(S_1 \leq h, S_2 \leq h) \\ &= e^{-q_i h} + (1 - e^{-q_i h})(1 - e^{-q(Y_1)h}) \\ &\leq e^{-q_i h} + (1 - e^{qh})(1 - e^{qh}).\end{aligned} \tag{2.4.4}$$

Si consideramos la expansión de Taylor de la ecuación (2.4.4) cuando $h \to 0$ se tiene que

$$\begin{aligned}\mathbb{P}_i(X_h = i) &\leq \mathbb{P}_i(h < J_1) + \sum_{m=3}^{\infty} \mathbb{P}_i(Y_{m-1} = i, J_{m-1} \leq h < J_m) \\ &\leq 1 + q_{ii}h + o(h) + (1 - (1 + qh + o(h))(1 - (1 + qh + o(h)) \\ &= 1 + q_{ii}h + o(h) + (-qh + o(h))(-qh + o(h)) \\ &= 1 + q_{ii}h + o(h) + q^2h^2 + o(h) \\ &= 1 + q_{ii}h + o(h),\end{aligned}$$

pues $q^2h^2 = o(h)$ cuando $h \to 0$. Por otra parte, de la expresión (2.4.2) se sigue que

$$\mathbb{P}_i(X_h = i) \geq \mathbb{P}_i(h < J_1),$$

de manera que si consideramos la expansión de Taylor y hacemos $h \to 0$ obtenemos la siguiente desigualdad

$$\mathbb{P}_i(X_h = i) \geq 1 + q_{ii}h + o(h),$$

por lo tanto,

$$\mathbb{P}_i(X_h = i) = 1 + q_{ii}h + o(h). \tag{2.4.5}$$

Sea $i \neq j$ entonces la probabilidad $\mathbb{P}(X_h = j|X_0 = i)$ por la igualdad (2.4.2) se sigue que

$$\begin{aligned}
\mathbb{P}(X_h = j|X_0 = i) &= \sum_{m=1}^{\infty} \mathbb{P}_i(Y_{m-1} = j, J_{m-1} \leq h < J_m) \\
&= \mathbb{P}_i(Y_0 = j, J_0 \leq h < J_1) \\
&\quad + \mathbb{P}_i(Y_1 = j, J_1 \leq h < J_2) \\
&\qquad + \sum_{m=3}^{\infty} \mathbb{P}_i(Y_{m-1} = j, J_{m-1} \leq h < J_m),
\end{aligned}$$

como $\mathbb{P}_i(Y_0 = j) = 0$ y $\{Y_1 = j, J_1 \leq h < J_2\} = \{J_1 \leq h, Y_1 = j, S_2 > h\}$ entonces de manera similar que (2.4.4),

$$\begin{aligned}
&\mathbb{P}(X_h = j|X_0 = i) \\
&\quad = \mathbb{P}_i(J_1 \leq h, Y_1 = j, S_2 > h) \qquad (2.4.6) \\
&\qquad + \sum_{m=3}^{\infty} \mathbb{P}_i(Y_{m-1} = j, J_{m-1} \leq h < J_m) \qquad (2.4.7) \\
&\quad \leq \mathbb{P}_i(J_1 \leq h, Y_1 = j, S_2 > h) + \mathbb{P}_i(S_1 < h, S_2 < h) \\
&\quad \leq \pi_{ij}(1 - e^{-q_i h})e^{-q_j h} + (1 - e^{qh})(1 - e^{qh}),
\end{aligned}$$

pues las variables aleatorias J_1, S_2 y Y_1 son independientes entre sí. Ahora, si recordamos que $\pi_{ij} = q_{ij}/q_i$, dada la expansión de Taylor cuando $h \to 0$ obtenemos que

$$\begin{aligned}
\mathbb{P}(X_h = j|X_0 = i) &\leq \frac{q_{ij}}{q_i}(q_i h + o(h))(1 - q_j h + o(h)) + o(h) \\
&= q_{ij}h + o(h).
\end{aligned}$$

Por otra parte, por las igualdades (2.4.2) y por (2.4.6), en particular cuando $h \to 0$, se cumple que

$$\begin{aligned}
\mathbb{P}(X_h = j|X_0 = i) &= \sum_{m=1}^{\infty} \mathbb{P}(Y_{m-1} = j, J_{m-1} \geq h < J_m|Y_0 = i) \\
&\geq \mathbb{P}_i(J_1 \leq h, Y_1 = j, S_2 > h) \\
&= \pi_{ij}(1 - e^{-q_i h})e^{-q_j h} \\
&= q_{ij}h + o(h),
\end{aligned}$$

para concluir que:

$$\mathbb{P}(X_h = j|X_0 = i) = q_{ij}h + o(h). \tag{2.4.8}$$

Por lo tanto, de (2.4.5) y (2.4.8) se sigue que para toda $i, j \in I$

$$\mathbb{P}(X_h = j|X_0 = i) = \delta_{ij} + q_{ij}h + o(h).$$

Ahora, por la propiedad fuerte de Markov, para toda t, $h \geq 0$ condicionando a $X_t = i$, X_{t+h} es independiente de $(X_s : s \leq t)$ de tal manera que $(X_{t+h})_{t\geq 0}$ es $Markov(\delta_i, Q)$ y, mientras $h \to 0$, uniformemente en t, tenemos que

$$\begin{aligned}\mathbb{P}(X_{t+h} = j|X_t = i) &= \mathbb{P}(X_h = j|X_0 = i)\\ &= \delta_{ij} + q_{ij}h + o(h).\end{aligned}$$

(b) $\Rightarrow$ (c)
Definamos a $p_{ij}(t) = \mathbb{P}(X_t = j|X_0 = i) = \mathbb{P}_i(X_t = j)$. Si suponemos (b), mientras $h \to 0$, tenemos

$$\begin{aligned}p_{ij}(t+h) &= \sum_{k\in I} \mathbb{P}_i(X_t = k)\mathbb{P}(X_{t+h} = j|X_t = k)\\ &= \sum_{k\in I} p_{ik}(t)(\delta_{kj} + q_{kj}h + o(h))\\ &= p_{ij}(t)(1 + q_{jj}h + o(h)) + \sum_{k\neq j} p_{ik}(t)(q_{kj}h + o(h))\\ &= p_{ij}(t) + p_{ij}(t)q_{jj}h + \sum_{k\neq j} p_{ik}(t)q_{kj}h + o(h),\end{aligned}$$

la última igualdad usando la convención de que para toda $c \neq 0$ constante, $c * o(h) = o(h)$ ya que $\sum_{k\neq j} p_{ik}(t)$ es una suma finita de términos no negativos. Tomando lo último deducimos que

$$\begin{aligned}\frac{p_{ij}(t+h) - p_{ij}(t)}{h} &= \frac{p_{ij}(t)q_{jj}h + \sum_{k\neq j} p_{ik}(t)q_{kj}h + o(h)}{h}\\ &= \frac{\sum_{k\in I} p_{ik}(t)q_{kj}h + o(h)}{h}\\ &= \sum_{k\in I} p_{ik}(t)q_{kj} + \frac{o(h)}{h},\end{aligned}$$

lo que implica

$$\lim_{h\to 0}\frac{p_{ij}(t+h)-p_{ij}(t)}{h}=\sum_{k\in I}p_{ik}(t)q_{kj},$$

es decir,

$$p'_{ij}(t)=\sum_{k\in I}p_{ik}(t)q_{kj} \qquad\qquad \text{con} \qquad\qquad p_{ij}(0)=\delta_{ij}.$$

Ahora, como I es finito, por el Teorema 2.1.4, $P(t)=(p_{ij}(t):i,j\in I,t\geq 0)=e^{tQ}$ es la única solución a la ecuación prospectiva

$$P'(t)=P(t)Q, \qquad P(0)=I.$$

También, si suponemos (b), tenemos que para toda $t,h\geq 0$, condicionando a $X_{t_n}=i_n$ tenemos que $X_{t_{n+1}}$ es independiente de $(X_s:s\leq t_n)$ lo que implica que

$$\mathbb{P}(X_{t_{n+1}}=i_{n+1}|X_{t_0}=i_0,...,X_{t_n}=i_n)=\mathbb{P}(X_{t_{n+1}}=i_{n+1}|X_{t_n}=i_n),$$

donde por la propiedad fuerte de Markov (Ver Sección 2.5), tenemos que el proceso $(X_{t_n+t})_{t\geq 0}$ condicionado a estar en i al tiempo t_n se distribuye $Markov(\delta_{ij},Q)$, con lo que concluimos

$$\mathbb{P}(X_{t_{n+1}}=i_{n+1}|X_{t_n}=i_n)=p_{ij}(t_{n+1}-t_n),$$

demostrando (c).

(c) $\Rightarrow$ (a)

Similar a la prueba (b) $\Rightarrow$ (a) de Teorema 2.4.5. □

Por el Teorema 2.1.4 sabemos que si I es finito entonces las ecuaciones prospectiva y retrospectiva tienen la misma solución. Por lo que en la condición (c) del resultado anterior podemos reemplazar la ecuación prospectiva por la retrospectiva.

En general, si $(X_t)_{t\leq 0}$ es un proceso de Markov con valores en I (*i.e.,espacio de estados igual a I*), dada una matriz función de transición $P(s,t)$, siempre existe un proceso de Markov $(X_t)_{t\geq 0}$ escogiendo una distribución inicial $\lambda=\mathbb{P}\big(X_0(\omega)=j\big)$, $i_n\in I$ con $n=1,2,...,N$ y $0=t_0<t_1<...<t_n$ y definiendo

$$\begin{aligned}\mathbb{P}(X_{t_0}(\omega)=i_0,X_{t_1}(\omega)=i_1,...,X_{t_n}(\omega)=i_n)=&\\ =\mathbb{P}(X_0(\omega)=i_0)p(0,t_1)\,p(t_2,t_1)...p(t_{n-1},t_n).&\end{aligned}$$

Si $0 \leq s < t < \infty$ y $\mathbb{P}(X_s = i) > 0$, definamos $p_{ij}(s,t) = \mathbb{P}(X_t = j | X_s = i)$ y la matriz función $P(s,t) = (p_{ij}(s,t) : i, j \in I)$. Dada una cadena de Markov a tiempo continuo $(X_t)_{t \geq 0}$ con probabilidades de transición estacionarias satisface que

$$\mathbb{P}(X_t(\omega) = j | X_s(\omega) = i) = \mathbb{P}(X_{(t-s)}(\omega) = j | X_0(\omega) = i),$$

mostrando que $p(s,t)$ depende únicamente del incremento $(t-s)$. De este hecho se desprende que al referirnos a una cadena de Markov continua asumimos que es homogénea en el tiempo. Observemos que de este hecho se sigue que

$$\begin{aligned}
\mathbb{P}(X_{t_0}(\omega) = i_0, X_{t_1}(\omega) = i_1, ..., X_{t_n}(\omega) = i_n) = \\
&= \mathbb{P}(X_0(\omega) = i_0) p(0,t_1)\, p(t_1,t_2) ... p(t_{n-1},t_n) \\
&= \mathbb{P}\,(X_0(\omega) = i_0)\, p(t_1)\, p(t_2 - t_1) ... p(t_n - t_{n-1}).
\end{aligned}$$

Las caracterizaciones de una cadena de Markov dependerán de la finitud de un conjunto numerable I. Por otro lado, si I es un conjunto infinito la ecuación retrospectiva también será escrita de la misma forma

$$P'(t) = QP(t), \qquad P(0) = I,$$

con la diferencia de que ahora tenemos un sistema infinito de ecuaciones diferenciales

$$p'_{ij}(t) = \sum_{k \in I} q_{ik} p_{kj}(t), \qquad p_{ij}(0) = \delta_{ij}$$

tales que los resultados de matrices exponenciales no nos serán útiles. Una solución de la ecuación retrospectiva es una matriz $(p_{ij}(t) : i, j \in I, t \geq 0)$ de funciones diferenciales que satisfacen el sistema de ecuaciones diferenciales.

Teorema 2.4.4. *Sea Q una $Q-matriz$. Entonces la ecuación retrospectiva*

$$P'(t) = QP(t), \qquad P(0) = I$$

tiene una solución mínima no negativa $(P(t) : t \geq 0)$. Esta solución forma un semigrupo de matrices

$$P(s)P(t) = P(s+t) \quad \textit{para toda} \quad s,t \geq 0.$$

Este resultado será probado junto con el siguiente teorema. Notemos que si I es finito tenemos que $P(t) = e^{tQ}$ por el Teorema 2.1.4. Llamaremos a $(P(t) : t \geq 0)$ como el *semigrupo mínimo no negativo* asociado a Q, o simplemente el *semigrupo* de Q. El siguiente resultado caracteriza a las cadenas de Markov a tiempo continuo con espacio de estados I infinito.

Teorema 2.4.5. *Sea $(X_t)_{t\geq 0}$ un proceso minimal continuo por la derecha con valores en I. Sea Q una $Q-matriz$ en I con matriz de salto Π y semigrupo $(P(t):t\geq 0)$. Entonces las siguientes condiciones son equivalentes:*

(a) (Definición de Cadena de saltos/tiempos de espera) Condicionando en $X_0 = i$, la cadena de saltos $(Y_n)_{n\geq 0}$ asociada a $(X_t)_{t\geq 0}$ es $MarkovTc(\delta_i,\Pi)$ y para toda $n\geq 1$, condicionando en $Y_0,...,Y_{n-1}$, los tiempos de espera $S_1,...,S_n$ son variables aleatorias exponenciales independientes de parámetros $q(Y_0),...,q(Y_{n-1})$ respectivamente;

(b) (definición de probabilidad de transición) para toda $n = 0,1,2,...$, todos los tiempos $0\leq t_0\leq t_1\leq ...\leq t_{n+1}$ y todos los estados $i_0,...,i_{n+1}$ tenemos

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1}|X_{t_0} = i_0,...,X_{t_n} = i_n) = p_{i_n i_{n+1}}(t_{n+1}-t_n).$$

Si $(X_t)_{t\geq 0}$ satisface cualquiera de estas dos condiciones entonces se dice que $(X_t)_{t\geq 0}$ es una *cadena de Markov con matriz generadora Q*, de forma resumida diremos que $(X_t)_{t\geq 0}$ es $Markov(\lambda,Q)$, donde λ es la distribución de X_0.

Demostración Teoremas 2.4.4 y 2.4.5 :
Sea $(X_t)_{t\geq 0}$ un proceso que satisface la condición (a). Definamos a $P(t)$ por

$$p_{ij}(t) = \mathbb{P}_i(X_t = j).$$

Si condicionamos en $X_0 = i$, tenemos que $J_1\sim exp(q_i)$ y $X_{J_1}\sim(\pi_{ij}:k\in I)$ ya que $\mathbb{P}_i(Y_1 = j) = \mathbb{P}_i(X_{J_1} = j)$. Por la Propiedad Fuertede de Markov (ver Teorema 2.4.2, si condicionamos a $J_1 = s$ y $X_{J_1} = k$, se tiene que $(X_{s+t})_{t\geq 0}\sim Markov(\delta_k,Q)$, ya que J_1 es un tiempo de paro (ver Lema 2.5.3). Dicho esto, observemos que

$$\mathbb{P}_i(X_t = j, t < J_1) = e^{-q_i t}\delta_{ij}$$

y

$$\mathbb{P}_i(J_1\leq t, X_{J_1} = k, X_t = j) = \int_0^t q_i e^{-q_i s}\pi_{ik}p_{kj}(t-s)ds.$$

Por lo tanto

$$\begin{aligned} p_{ij}(t) &= \mathbb{P}_i(X_t = j, t < J_1) + \sum_{k\neq i}\mathbb{P}_i(J_1\leq t, X_{J_1} = k, X_t = j) \\ &= e^{-q_i t}\delta_{ij} + \sum_{k\neq i}\int_0^t q_i e^{-q_i s}\pi_{ik}p_{kj}(t-s)ds. \end{aligned} \tag{2.4.9}$$

Consideremos ahora el cambio de variable $u = t - s$, por el teorema de convergencia monótona (ver [Nor97, Teorema 6.4.1]) obtenemos

$$p_{ij}(t) = e^{-q_i t}\delta_{ij} + \int_0^t \sum_{k \neq i} q_i e^{-q_i(t-u)} \pi_{ik} p_{kj}(u) ds.$$

Definamos a $(R_n)_{n=1}^{\infty}$ como

$$R_n = \sum_{m=1}^{n} q_i e^{-q_i(t-u)} \pi_{ik_m} p_{k_m j}(u)$$

y notemos que $\lim_{n\to\infty} R_n = \sum_{k\neq i} q_i e^{-q_i(t-u)} \pi_{ik} p_{kj}(u)$. Observemos que,

$$\begin{aligned} q_i e^{-q_i(t-u)} \pi_{ik_m} p_{k_m j}(u) &= q_i e^{-q_i(t-u)} \frac{q_{ik_m}}{q_i} p_{k_m j}(u) \\ &= e^{-q_i(t-u)} q_{ik_m} p_{k_m j}(u) \\ &\leq q_{ik_m}, \end{aligned}$$

pues $e^{-q_i(t-u)} p_{k_m j}(u) \leq 1$. Si definimos a $M_n = \sum_{m=1}^{n} q_{ik_m}$ tenemos que $R_n \leq M_n$ para toda $n \geq 1$, donde $\lim_{n\to\infty} M_n$ existe y es q_i ya que Q es una $Q-matriz$. De la ecuación 2.4.9 se sigue que $p_{ij}(t)$ es continua pues es la suma de funciones continuas. Por la prueba M de Weierstrass (ver [Car04, Lema 10.9]) se tiene que R_n converge uniformemente a una función continua , pues cada término de R_n es un producto de funciones continuas. Dicho lo anterior, podemos concluir que p(t) es diferenciable, por lo que si multiplicamos a la expresión (2.4.9) por $e^{q_i t}$ obtenemos

$$e^{q_i t} p_{ij}(t) = \delta_{ij} + \int_0^t \sum_{k \neq i} q_i e^{q_i u} \pi_{ik} p_{kj}(u) du, \tag{2.4.10}$$

con lo que si consideramos a $\frac{d}{dt}(e^{q_i t} p_{ij}(t))$ obtenemos que

$$\frac{d}{dt}(e^{q_i t} p_{ij}(t)) = e^{q_i t}(q_i p_{ij}(t) + p'_{ij}(t))$$

por lo que si derivamos el lado izquierdo de la expresión (2.4.10) se sigue que

$$e^{q_i t}(q_i p_{ij}(t) + p'_{ij}(t)) = \sum_{k \neq i} q_i e^{q_i t} \pi_{ik} p_{kj}(t)$$

tomando en cuenta que $q_i = -q_{ii}$ y $q_{ik} = q_i \pi ik$, obtenemos

$$\begin{aligned}
p'_{ij}(t) &= \sum_{k \neq i} q_i \frac{q_{ik}}{q_i} p_{kj}(t) - q_i p_{ij}(t) \\
&= \sum_{k \neq i} q_{ik} p_{kj}(t) + q_{ii} p_{ij}(t) \\
&= \sum_{k \in I} q_{ik} p_{kj}(t),
\end{aligned}$$

con lo que hemos demostrado que $P(t)$ satisface la ecuación restrospectiva.

Ahora, consideremos la siguiente probabilidad

$$\begin{aligned}
&\mathbb{P}_i(X_t = j, t < J_{n+1}) \\
&\qquad = \mathbb{P}_i(X_t = j, t < J_1) + \sum_{k \neq i} \mathbb{P}_i(J_1 \leq t < J_{n+1}, X_{J_1} = k, X_t = j) \\
&\qquad = e^{-q_i t} \delta_{ij} \\
&\qquad\quad + \sum_{k \neq i} \int_0^t q_i e^{-q_i s} \pi_{ik} \mathbb{P}_i(X_{t-s} = j, t - s < J_n) ds.
\end{aligned}$$

Sea $\tilde{P}(t)$ otra solución no negativa de la ecuación retrospectiva y observemos que también debe de satisfacer la forma integral de la misma, es decir

$$\tilde{p}_{ij}(t) = e^{-q_i t} \delta_{ij} + \sum_{k \neq i} \int_0^t q_i e^{-q_i s} \pi_{ik} \tilde{p}_{kj}(t - s) ds. \tag{2.4.11}$$

Si $\tilde{P}(t) \geq 0$, entonces,

$$\mathbb{P}_i(X_t = j, t < J_0) = 0 \leq \tilde{p}_{ij}(t) \qquad \text{para toda } i, j \text{ y } t.$$

Supongamos inductivamente que

$$\mathbb{P}_i(X_t = j, t < J_n) \leq \tilde{p}_{ij}(t) \qquad \text{para toda } i, j \text{ y } t,$$

entonces por ((2.4.11)) y por hipótesis de inducción

$$\begin{aligned}
&\mathbb{P}_i(X_t = j, t < J_{n+1}) \\
&\qquad = e^{-q_i t} \delta_{ij} + \sum_{k \neq i} \int_0^t q_i e^{-q_i s} \pi_{ik} \mathbb{P}_i(X_{t-s} = j, t - s < J_n) ds \\
&\qquad \leq e^{-q_i t} \delta_{ij} + \sum_{k \neq i} \int_0^t q_i e^{-q_i s} \pi_{ik} \tilde{p}_{kj}(t - s) ds.
\end{aligned}$$

Por lo tanto,

$$\begin{aligned} p_{ij}(t) &= \lim_{n\to\infty} \sum_{m=1}^{n} \mathbb{P}(Y_{m-1} = j, J_{m-1} \leq h < J_m | Y_0 = i) \\ &= \lim_{n\to\infty} \mathbb{P}_i(X_t = j, t < J_n) \\ &\leq \tilde{p}_{ij}(t) \end{aligned}$$

para toda i, j en I y $t \geq 0$. Ahora, sean $i, j \in I$ y $0 \leq s \leq t$

$$\begin{aligned} p_{ij}(t+s) = \mathbb{P}_i(X_{t+s} = j) &= \sum_{k\in I} \mathbb{P}_i(X_{t+s} = j | X_t = k)\mathbb{P}_i(X_t = k) \\ &= \sum_{k\in I} \mathbb{P}_k(X_s = j)\mathbb{P}_i(X_t = k) = \sum_{k\in I} p_{kj}(s) p_{ik}(t), \end{aligned}$$

por la propiedad de Markov. Si $p_{i\infty}(s) > 0$ (ver ecuación (2.2.2)) considerando que ∞ es un estado absorbente por lo anterior tenemos que

$$p_{i\infty}(t+s) = 1 - \sum_{j\in I} p_{ij}(t+s) \qquad = 1 - \sum_{j\in I}\sum_{k\in I} p_{ik}(s) p_{ki}(s) = 1,$$

pues $p_{ij}(t) = p_{ij}(s) = 0$, mostrando que $p_{i\infty}(t+s) = p_{i\infty}(s)p_{i\infty}(s)$. Esto demuestra que $(P(t) : t \geq 0)$ es un semigrupo de matrices como las requeridas en el Teorema 2.4.4.

Ahora, como $(X_t)_{t\geq 0}$ satisface (a) y por la propiedad de Markov

$$\begin{aligned} &\mathbb{P}_i(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, ..., X_{t_n} = i_n) \\ &\qquad\qquad = \mathbb{P}_{i_n}(X_{t_{n+1}} = i_{n+1}) = p_{i_n i_{n+1}}(t_{n+1} - t_n) \end{aligned}$$

para concluir que $(X_t)_{t\geq 0}$ satisface (b).

(b) $\Rightarrow$ (a)
Supongamos (b) y asumamos que $X_0 = i$. Para $s \geq 0$ el evento $\{S_1 > s\}$ es equivalente al evento $\{X_u = i \text{ para } 0 \leq u \leq s\}$, de manera similar, para $s, t \geq 0$, evento $\{S_1 > s + t\}$ es equivalente a $\{X_u = i \text{ para } 0 \leq u \leq s + t\}$, entonces

$$\begin{aligned} &\mathbb{P}_i(S_1 > s + t) \\ &\quad = \mathbb{P}(X(u) = i \text{ para } 0 \leq u \leq s + t | X(u) = i \text{ para } 0 \leq u \leq s) \\ &\quad = \mathbb{P}(X(u) = i \text{ para } s \leq u \leq s + t | X(u) = i \text{ para } 0 \leq u \leq s), \end{aligned}$$

donde por la propiedad de Markov:

$$\begin{aligned}\mathbb{P}(X(u) = i \text{ para } s \leq u \leq s+t | X(u) = i \text{ para } 0 \leq u \leq s) \\ = \mathbb{P}(X(u) = i \text{ para } 0 \leq u \leq t | X(0) = i) \\ = \mathbb{P}_i(S_1 > t),\end{aligned}$$

mostrando que $S_1 \sim exp(\lambda_i)$ para alguna $\lambda_i \geq 0$. Por otra parte, sabemos que $P(t)$ satisface la ecuación retrospectiva donde

$$P'(0) = QP(0) = Q,$$

de manera que $p'_{ii}(0) = -q_i$ donde q_i es la razón de dejar el estado i y $p'_{ij}(0) = q_{ij}$ es la razón de llegar al estado j desde i, lo que nos conduce a

$$\begin{aligned}\mathbb{P}_i(X_{J_1} = j) &= \mathbb{E}_i\left[\int_0^{J_1} q_{ij} dt\right] \\ &= q_{ij} E_i[J_1] \\ &= \frac{q_{ij}}{\lambda_i}\end{aligned}$$

y

$$\mathbb{P}_i(X_{J_1} = i) = 0;$$

lo último debido a que por definición $J_1 = S_1$ es el tiempo en el que el proceso $(X_t)_{t\geq 0}$ deja al estado inicial. Ahora, observemos que $P(t)$ debe de satisfacer que para toda $t \geq 0$

$$\sum_{k\in I} p_{ik}(t) = 1,$$

en particular

$$\begin{aligned}\sum_{k\in I} p_{ik}(J_1) &= \sum_{k\in I} \frac{q_{ik}}{\lambda_i} \\ &= \sum_{k\neq i} \frac{q_{ik}}{\lambda_i} \\ &= 1,\end{aligned}$$

lo que implica que

$$\begin{aligned}\sum_{k\neq i} q_{ij} &= \lambda_i\\ &= q_i.\end{aligned}$$

Definamos al proceso $(Y_n)_{n\geq 0}$ como $Y_n = X_{J_n}$, es decir la cadena de saltos asociada al proceso $(X_t)_{t\geq 0}$ y observemos que por la propiedad de Markov tenemos que

$$\begin{aligned}\mathbb{P}_i(Y_{n+1} = i_{n+1}|Y_0 = i_0,\ldots,Y_n = i_n)&\\ &= \mathbb{P}_i(X_{J_{n+1}} = i_{n+1}|X_{J_0} = i_0,\ldots,X_{J_n} = i_n)\\ &= \mathbb{P}_{i_n}(X_{J_{n+1}} = i_{n+1})\\ &= \mathbb{P}_{i_n}(Y_{n+1} = i_{n+1})\\ &= \frac{q_{ij}}{q_i},\end{aligned}$$

donde $\pi_{ij} = q_{ij}/q_i$. Por lo tanto, condicionando a $Y_0 = i$, la cadena de saltos $(Y_n)_{n\geq 0}$ es $Markovdis(\delta_{ij},\Pi)$. Ahora, por lo anterior

$$\begin{aligned}\mathbb{P}_i(S_{n+1} > s\,|\,Y_n = j) &= \mathbb{P}(S_{n+1} > s\,|\,Y_n = j, Y_0 = i)\\ &= \mathbb{P}(S_{n+1} > s\,|\,Y_n = j)\\ &= \mathbb{P}(S_1 > s\,|\,Y_0 = j),\end{aligned}$$

por lo tanto, si condicionamos en $Y_0,\ldots,Y_{n-1}$, los tiempos de espera $S_1,\ldots,S_n$ son variables aleatorias exponenciales independientes de parámetros $q(Y_0)$, $\ldots,q(Y_{n-1})$ respectivamente, concluyendo la demostración del teorema. □

Notemos que a lo largo de la demostración del teorema anterior no mencionamos a la ecuación prospectiva para el caso en el que el espacio de estados es infinito, el siguiente teorema muestra que el semigrupo $(P(t) : t \geq 0)$ de Q, también satisface la ecuación prospectiva:

Teorema 2.4.6. *La solución mínima no negativa $(P(t) : t \geq 0)$ de la ecuación retrospectiva también es la solución mínima no negativa de la ecuación prospectiva*

$$P'(t) = QP(t), \qquad P(0) = I.$$

La demostración del teorema anterior se omitirá; puede ser consultada en [Nor97, p. 100].

2.5 Tiempos de paro y la Propiedad Fuerte de Markov

La idea intuitiva detrás la propiedad fuerte de Markov es que dada una cadena de Markov a tiempo continuo, lo que pase en el futuro dado que "paramos" el proceso será independiente de lo que sucedió en el pasado. La propiedad fuerte de Makov no puede se estudiada propiamente sin la teoría de la medida. El problema reside propiamente en la noción de como medir los eventos que "dependen" de algún fenómeno. Esta noción se precisa con la teoría de la medida, específicamente, con conceptos como σ-álgebra, filtraciones y medida (ver Capítulo 1 Preliminares). En teoría de la medida las filtraciones (ver Definición 1.2.2.2) representan la información disponible hasta el tiempo t de algún fenómeno en algún espacio medible en términos de la σ-álgebra $\mathscr{F}_t$. Por otra parte, los procesos estocásticos se han definido como una colección de variables aleatorias indexadas por t. Podemos combinar estos dos conceptos de manera que para algún proceso estocástico podamos decir que el valor que toma X_t se sabe hasta el tiempo t. En esta sección nos enfocamos a estudiar la Propiedad Fuerte de Markov bajo este enfoque.

Sea $(X_t)_{t\geq 0}$ un proceso estocástico continuo por la derecha con valores en un conjunto numerable I. Sea $\mathscr{F}_t$ la σ-álgebra generada por $\{X_s : s \leq t\}$, es decir, generada por los eventos $\{X_s = i\}$ donde $s \leq t$ e $i \in I$ (ver Definición 1.2.2.4). Sea un tiempo de paro T con respecto a la filtración canónica de $(X_t)_{t\geq 0}$ (ver Definición 2.4.1), para referirnos a T con estas especificaciones de manera simple diremos que T es un tiempo de paro de $(X_t)_{t\geq 0}$, es decir, un tiempo de paro asociado a la filtración canónica de $(X_t)_{t\geq 0}$.

Dado un tiempo de paro T, definimos lo siguiente:

Definición 2.5.1. *Sea T un tiempo de paro con respecto a la filtración $\mathscr{F}_t$, se define como*

$$\mathscr{F}_T := \{A \in \mathscr{F} : A \cap \{T \leq t\} \in \mathscr{F}_t \quad \textit{para toda } t \geq 0\},$$

la σ-álgebra asociada al tiempo de paro T.

Es fácil demostrar que en efecto $\mathscr{F}_T$ es una σ-álgebra por lo cual se omite su demostración.

Lema 2.5.2. *Sean S y T tiempos de paro de $(X_t)_{t\geq 0}$ entonces X_T y $\{S \leq T\}$ son $\mathscr{F}_T$-medibles.*

Demostración:
Como $(X_t)_{t\geq 0}$ es un proceso continuo por la derecha, si pensamos en el evento $\{T < t\}$, existe un natural $n \geq 0$ tal que para toda $m \geq n$, para alguna $k \geq 1$ se tiene

$$(k-1)2^{-m} \leq T < k2^{-m} \leq t \quad \text{donde} \quad X_{k2^{-m}} = X_T.$$

Así, para toda i en I tenemos

$$\begin{aligned}\{X_T = i\} \cap \{T \leq t\} = (\{X_t = i\} \cap \{T = t\})& \\ \cup \left(\bigcup_{n=0}^{\infty} \bigcap_{m=n}^{\infty} \bigcup_{k=1}^{[2^m t]} \{X_{k2^{-m}} = i\} \cap \{(k-1)2^{-m} \leq T < k2^{-m}\} \right)&.\end{aligned}$$

Por una parte

$$\{T = t\} = \{T \leq t\} \setminus \{T < t\} \quad \text{y} \quad \{T < t\} = \bigcup_{n=1}^{\infty} \{T \leq t - \frac{1}{n}\}$$

donde

$$\{T \leq t - \frac{1}{n}\} \in \mathscr{F}_{t-\frac{1}{n}} \subseteq \mathscr{F}_t \quad \text{para toda } n = 1, 2, ...$$

así, como $\mathscr{F}_t$ es σ-álgebra tenemos que que $(\Omega \setminus \{T < t\})$ también está en $\mathscr{F}_t$, entonces

$$\begin{aligned}\{T = t\} &= \{T \leq t\} \setminus \{T < t\} \\ &= \{T \leq t\} \cap (\Omega \setminus \{T < t\}) \in \mathscr{F}_t\end{aligned}$$

para obtener que

$$\{X_t = i\} \cap \{T = t\} \in \mathscr{F}_t$$

pues $\{X_t = i\}$ pertenece a $\mathscr{F}_t$. Por otro lado

$$\{(k-1)2^{-m} \leq T < k2^{-m}\} \in \mathscr{F}_{k2^{-m}} \subseteq \mathscr{F}_t$$

para toda $m \geq n$ y $1 \leq k < 2^m t$ por lo que

$$\left(\bigcup_{n=0}^{\infty} \bigcap_{m=n}^{\infty} \bigcup_{k=1}^{[2^m t]} \{X_{k2^{-m}} = i\} \cap \{(k-1)2^{-m} \leq T < k2^{-m}\} \right) \in \mathscr{F}_t$$

ya que está en términos de uniones e intersecciones numerables e implica que X_T es $\mathscr{F}_T$-medible. Consideremos ahora

$$\{S > T\} \cap \{T \leq t\} = \bigcup_{s \in \mathbb{Q}, s \leq t} \big(\{T < s\} \cap \{S > s\}\big)$$

pertenece a $\mathscr{F}_t$ pues cualquier subconjunto de $\mathbb{Q}$ es numerable. Por lo tanto, $\{S > T\}$ está en $\mathscr{F}_T$ lo que implica que $\{S < T\}$ también lo está pues $\mathscr{F}_T$ es σ-álgebra. $\square$

Observación 2.5.1. Supongamos S y T son tiempos de paro. Entonces $S \wedge T = min\{S,T\}$ y $S \vee T = max\{S,T\}$ son tiempos de paro. Si $\{T_n\}_{n\in\mathbb{N}}$ es una sucesión de tiempos de paro, entonces $sup_n\{T_n\}$ es un tiempo de paro. Este hecho no se demuestra sin embargo para su demostración se refiere [Sam15, Lema 3.1.9].

Lema 2.5.3. *Para toda $m \geq 0$, el tiempo de salto J_m es un tiempo de paro de $(X_t)_{t\geq 0}$.*

Demostración: Demostración por inducción. Para $m = 0$ es claro que J_0 es tiempo de paro pues para toda $t \geq 0$ tenemos que $\{0 \leq t\} = \Omega$, el cual pertenece a $\mathscr{F}_t$. Ahora, supongamos inductivamente que J_m es un tiempo de paro, entonces

$$\{J_{m+1} \leq t\} = \bigcup_{s \in \mathbb{Q}, s \leq t} \Big(\{J_m \leq s\} \cap \{X_s \neq X_{J_m}\}\Big).$$

Por hipótesis de inducción J_m es un tiempo de paro y esto implica que $\{J_m \leq s\} \in \mathscr{F}_t$ para toda $t \geq 0$ y el evento $\{X_s \neq X_{J_m}\} \in \mathscr{F}_s$ para toda $s \leq t$ pues si definimos a $X = \bigcup_{i\in I}\{X_{J_m} = i\}$ entonces $\{X_s \neq X_{J_m}\} = X \setminus \{X_s = X_{J_m}\}$. De esto se sigue que $\{J_{m+1} \leq t\}$ también está en $\mathscr{F}_t$ pues es una unión numerable de conjuntos en $\mathscr{F}_t$. Por lo tanto, para toda $m \geq 0$ el tiempo de salto J_m es un tiempo de paro de $(X_t)_{t\geq 0}$. $\square$

Sea $(X_t)_{t\geq 0}$ un proceso minimal continuo por la derecha y consideremos la construcción de $(X_t)_{t\geq 0}$ por medio de su proceso de saltos $(Y_n)_{n\geq 0}$ y sus tiempos de espera $(S_n)_{n\geq 1}$(ver Sección 2.2). Sea $\mathscr{F} = \sigma(X_t : t \geq 0)$, la σ-álgebra generada por $(X_t)_{t\geq 0}$. Denotemos por $\mathscr{G}_m$ a la σ-álgebra generada por $Y_0, \ldots, Y_m$ y $S_1, \ldots, S_{m-1}$ dada por

$$\mathscr{G}_m = \sigma\Big((Y_n)_{n\geq 0}^{m}, (S_n)_{n\geq 1}^{m-1}\Big), \tag{2.5.1}$$

con $\mathscr{G}_0 = \Omega$. Adicionalmente, definamos

$$\mathscr{G} = \sigma\Big((Y_n)_{n\geq 0}, (S_n)_{n\geq 1}\Big), \tag{2.5.2}$$

de manera que

$$\mathscr{G} = \sigma\Big(\bigcup_{m=1}^{\infty} \mathscr{G}_m\Big). \tag{2.5.3}$$

Por los Lemas 2.5.2 y 2.5.3, ambos $(Y_n)_{n\geq 0}$ $(S_n)_{n\geq 1}$ son $\mathscr{F}$-medibles y esto implica que $\mathscr{G} \subseteq \mathscr{F}$. Por otro lado, para toda $i \in I$

$$\{X_t = i\} = \bigcup_{n\geq 0} \{J_n \leq t < J_{n+1}\} \cap \{Y_n = i\} \in \mathscr{G},$$

mostrando que $\mathscr{F} \subseteq \mathscr{G}$. Un π-sistema útil para generar a $\mathscr{G}$ está dado por los conjuntos de la forma

$$B = \{Y_0 = i_0, ..., Y_n = i_n, S_1 > s_1, ..., S_n > s_n\}.$$

La definición vía proceso de saltos y tiempos de espera de una cadena de Markov $(X_t)_{t\geq 0}$ con distribución inicial λ y matriz generadora Q se declar como sigue para los eventos de la forma B

$$\mathbb{P}(B) = \lambda_{i_0}\pi_{i_0 i_1}...\pi_{i_{n-1}i_n}e^{-q_{i_0}s_1}...e^{-q_{i_{n-1}}s_n}.$$

Por el Teorema 1.2.1.5 y lo anterior $\mathbb{P}$ está determinada por el π-sistema que genera a $\mathscr{G}$ y por ende a $\mathscr{F}$.

Lema 2.5.4. *Sea T un tiempo de paro de $(X_t)_{t\geq 0}$ y sea $A \in \mathscr{F}_T$. Entonces para toda $m \geq 0$ existe una variable aleatoria T_m y un conjunto $A_m \in \mathscr{G}_m$, ambos medibles con respecto a $\mathscr{G}_m$ tal que $T = T_m$ y $\mathbb{1}_A = \mathbb{1}_{A_m}$ en $\{T < J_{m+1}\}$.*

Demostración: Sea $t \geq 0$ y consideremos a

$$\mathscr{A}_t = \Big\{A \in \mathscr{F}_t : A \cap \{t < J_{m+1}\} = A_m \cap \{t < J_{m+1}\} \text{ para algún } A_m \in \mathscr{G}_m\Big\}.$$

Sea $X = \mathscr{P}(I)$, notemos que $\mathscr{A}_t$ es σ-álgebra pues

1. Como $\mathscr{G}_m$ es σ-álgebra tenemos que $\emptyset \in \mathscr{G}_m$ para toda $m \geq 0$. Por lo tanto, $\emptyset \in \mathscr{A}_t$ pues $\emptyset \cap \{t < J_{m+1}\} = \emptyset \in \mathscr{G}_m$ para toda $m \geq 0$.

2. Sea $B \in \mathscr{A}_t$, por definición de $\mathscr{A}_t$ se sigue que $B \cap \{t < J_{m+1}\} = B_m \cap \{t < J_{m+1}\}$ para algún $B_m \in \mathscr{G}_m$. Por un lado, como $\mathscr{F}_t$ y $\mathscr{G}_m$ son σ-álgebras respectivamente, existen $B^c \in \mathscr{F}_t$ y $B_m^c \in \mathscr{G}_m$. Supongamos $B^c \cap \{t < J_{m+1}\} \neq$

$B_m^c \cap \{t < J_{m+1}\}$ y supongamos que $B^c \cap \{t < J_{m+1}\} \neq \emptyset \neq B_m^c \cap \{t < J_{m+1}\}$ respectivamente, entonces existe $x \in B^c \cap \{t < J_{m+1}\}$ tal que $x \notin B_m^c \cap \{t < J_{m+1}\}$ o existe $x_m \in B_m^c \cap \{t < J_{m+1}\}$ tal que $x_m \notin B^c \cap \{t < J_{m+1}\}$. Por lo anterior, sin pérdida de generalidad supongamos existe $x \in B^c \cap \{t < J_{m+1}\}$ tal que $x \notin B_m^c \cap \{t < J_{m+1}\}$, esto implica que $x \in B^c$ y $x \in \{t < J_{m+1}\}$. Como $x \notin B_m^c \cap \{t < J_{m+1}\}$ se sigue que $x \in B_m \cap \{t < J_{m+1}\}$ y $x \notin B$, esto implica que $B \cap \{t < J_{m+1}\} \subsetneq B_m \cap \{t < J_{m+1}\}$, lo cual implica una contradicción pues $B \cap \{t < J_{m+1} = B_m \cap \{t < J_{m+1}\}$. Por lo tanto, $B^c \cap \{t < J_{m+1}\} = B_m^c \cap \{t < J_{m+1}\}$ para algún $B_m \in \mathscr{G}_m$ y lo cual implica que $B^c \in \mathscr{A}_t$.

3. Sea $B_1, B_2, ...$ una sucesión de elementos en $\mathscr{A}_t$, por definición de $\mathscr{A}_t$, para toda $n = 1, 2, ...$, existen $B_{m_1}, B_{m_2}, ...$ tales que

$$B_n \cap \{t < J_{m+1}\} = B_{m_n} \cap \{t < J_{m+1}\}$$

con $B_{m_n} \in \mathscr{G}_m$ para algún $m \geq 0$. Definamos $B_m = \bigcup_{n=1}^{\infty} B_{m_n}$ y $B = \bigcup_{n=1}^{\infty} B_n$, en particular como $\mathscr{G}_m$ es σ-álgebra tenemos que $B_m \in \mathscr{G}_m$. Entonces, por distribución de la unión sobre la intersección

$$\begin{aligned} B \cap \{t < J_{m+1}\} &= \left(\bigcup_{n=1}^{\infty} B_i \right) \cap \{t < J_{m+1}\} \\ &= \bigcup_{n=1}^{\infty} (B_i \cap \{t < J_{m+1}\}) \\ &= \bigcup_{n=1}^{\infty} (B_{m_i} \cap \{t < J_{m+1}\}) \\ &= \left(\bigcup_{n=1}^{\infty} B_{m_i} \right) \cap \{t < J_{m+1}\} \\ &= B_m \cap \{t < J_{m+1}\}, \end{aligned}$$

como $B_m \in \mathscr{G}_m$ para alguna m se sigue que $B \in \mathscr{A}_t$ mostrando que $\mathscr{A}_t$ es σ-álgebra.

Por definición de $\mathscr{A}_t$ tenemos que $\mathscr{A}_t \subseteq \mathscr{F}_t$. Observemos que para $s \leq t$, se tiene que

$$\begin{aligned} &\{X_s = i\} \cap \{t < J_{m+1}\} \\ &\qquad = \left(\bigcup_{k=0}^{m-1} \{Y_k = i, J_k \leq s < J_{k+1}\} \cup \{Y_m = i, J_m \leq s\} \right) \cap \{t < J_{m+1}\}, \end{aligned}$$

donde

$$\left(\bigcup_{k=0}^{m-1}\{Y_k = i, J_k \leq s < J_{k+1}\} \cup \{Y_m = i, J_m \leq s\}\right) \cap \{t < J_{m+1}\},$$

mostrando que $\{X_s = i\} \in \mathscr{A}_t$. Como los eventos de la forma $\{X_s = i\}$ para $s \leq t$ generan a la σ-álgebra $\mathscr{F}_t$, se sigue que $\mathscr{F}_t \subseteq \mathscr{A}_t$ demostrando que

$$\mathscr{A}_t = \mathscr{F}_t \quad \text{para toda } t \geq 0.$$

Sea T un tiempo de paro y $A \in \mathscr{F}_T$, definamos

$$\begin{aligned} B(t) &:= \{T \leq t\} \\ A(t) &:= A \cap \{T \leq t\}, \end{aligned}$$

como T es un tiempo de paro y $A \in \mathscr{F}_T$ se cumple que $B(t) = \{T \leq t\} \in \mathscr{F}_t$ y $A(t) = A \cap \{T \leq t\} \in \mathscr{F}_t$ para toda $t \geq 0$. Como $\mathscr{A}_t = \mathscr{F}_t$, entonces por lo anterior, para algún $m \geq 0$ existen $B_m(t), A_m(t) \in \mathscr{G}_m$ tales que

$$\begin{aligned} B(t) \cap \{T < J_{m+1}\} &= B_m(t) \cap \{T < J_{m+1}\} \\ A(t) \cap \{T < J_{m+1}\} &= A_m \cap \{T < J_{m+1}\}. \end{aligned}$$

Definamos

$$T_m = \sup_{t \in \mathbb{Q}} t \mathbb{1}_{B_m(t)} \quad , \quad A_m = \bigcup_{t \in \mathbb{Q}} A_m(t).$$

Como $B_m(t) \in \mathscr{G}_m$, se tiene que $\mathbb{1}_{B_m(t)}$ es $\mathscr{G}_m$-medible de manera que a consecuencia de la Observación 2.5.1 se sigue que $\sup_{t\in\mathbb{Q}} t\mathbb{1}_{B_m(t)}$ pues $t\mathbb{1}_{B_m(t)}$ define una sucesión de tiempos de paro en $\mathscr{G}_m$ de manera que $\sup_{t\in\mathbb{Q}} t\mathbb{1}_{B_m(t)}$ es también un tiempo de paro $\mathscr{G}_m$. Adicionalmente, como $\mathscr{G}_m$ es σ-álgebra se sigue que $A_m = \bigcup_{t\in\mathbb{Q}} A_m(t) \in \mathscr{G}_m$. Entonces,

$$(\sup_{t\in\mathbb{Q}} t\mathbb{1}_{B_m(t)})\mathbb{1}_{\{T<J_{m+1}\}} = \sup_{t\in\mathbb{Q}} t\mathbb{1}_{B_m(t)\cap\{T<J_{m+1}\}}, \tag{2.5.4}$$

por lo siguiente;

1. Si $T \leq t$ y $T < J_{m+1}$, entonces

$$\begin{aligned} (\sup_{t\in\mathbb{Q}} t\mathbb{1}_{B_m(t)})\mathbb{1}_{\{T<J_{m+1}\}} &= (\sup_{t\in\mathbb{Q}} t\mathbb{1}_{B_m(t)}) \\ &= T_m. \end{aligned}$$

Si $T < t < J_{m+1}$, entonces

$$\begin{aligned}\sup_{t\in\mathbb{Q}} t\,\mathbb{1}_{B_m(t)\cap\{T<J_{m+1}\}} &= \sup_{t\in\mathbb{Q}} t\,\mathbb{1}_{B_m(t)}\\ &= T_m,\end{aligned}$$

y si $T < J_{m+1} < t$, entonces

$$\begin{aligned}\sup_{t\in\mathbb{Q}} t\,\mathbb{1}_{B_m(t)\cap\{T<J_{m+1}\}} &= \sup_{t\in\mathbb{Q}} t\,\mathbb{1}_{B_m(t)}\\ &= T_m.\end{aligned}$$

2. Si $T \leq t$ y $T > J_{m+1}$,

$$\begin{aligned}(\sup_{t\in\mathbb{Q}} t\,\mathbb{1}_{B_m(t)})\mathbb{1}_{\{T<J_{m+1}\}} &= (\sup_{t\in\mathbb{Q}} t\,\mathbb{1}_{B_m(t)})\,0\\ &= 0,\end{aligned}$$

   y de manera similar

$$\sup_{t\in\mathbb{Q}} t\,\mathbb{1}_{B_m(t)\cap\{T<J_{m+1}\}} = 0,$$

   pues $\{T \leq t\}$ y $\{T < J_{m+1}\}$ no se cumple.

3. Si $T > t$ y $T > J_{m+1}$,

$$\begin{aligned}(\sup_{t\in\mathbb{Q}} t\,\mathbb{1}_{B_m(t)})\mathbb{1}_{\{T<J_{m+1}\}} &= (\sup_{t\in\mathbb{Q}} t\,0)\,0\\ &= 0,\end{aligned}$$

   y de manera similar

$$\sup_{t\in\mathbb{Q}} t\,\mathbb{1}_{B_m(t)\cap\{T<J_{m+1}\}} = 0,$$

   pues $\{T \leq t\}$ y $\{T < J_{m+1}\}$ no se cumple.

4. Finalmente, Si $T > t$ y $T < J_{m+1}$,

$$\begin{aligned}(\sup_{t\in\mathbb{Q}} t\,\mathbb{1}_{B_m(t)})\mathbb{1}_{\{T<J_{m+1}\}} &= (\sup_{t\in\mathbb{Q}} t\,0)\,1\\ &= 0,\end{aligned}$$

y de manera similar

$$\sup_{t\in\mathbb{Q}} t\mathbb{1}_{B_m(t)\cap\{T<J_{m+1}\}} = 0,$$

pues $\{T \leq t\}$ y $\{T < J_{m+1}\}$ no se cumple.

Demostrando que si $T \leq t$ y $T < J_{m+1}$, entonces

$$(\sup_{t\in\mathbb{Q}} t\mathbb{1}_{B_m(t)})\mathbb{1}_{\{T<J_{m+1}\}} = \sup_{t\in\mathbb{Q}} t\mathbb{1}_{B_m(t)\cap\{T<J_{m+1}\}}.$$

Entonces,

$$\begin{aligned} T_m\mathbb{1}_{\{T<J_{m+1}\}} &= (\sup_{t\in\mathbb{Q}} t\mathbb{1}_{B_m(t)})\mathbb{1}_{\{T<J_{m+1}\}} \\ &= \sup_{t\in\mathbb{Q}} t\mathbb{1}_{B_m(t)\cap\{T<J_{m+1}\}}, \end{aligned}$$

como $B(t)\cap\{T < J_{m+1}\} = B_m(t)\cap\{T < J_{m+1}\}$ en $\mathscr{G}_m$ para algún $m \geq 0$,

$$\begin{aligned} &= \sup_{t\in\mathbb{Q}} t\mathbb{1}_{B(t)\cap\{T<J_{m+1}\}} \\ &= \sup_{t\in\mathbb{Q}}(t\mathbb{1}_{\{T\leq t\}\cap\{T<J_{m+1}\}}) \qquad (2.5.5) \\ &= \sup_{t\in\mathbb{Q}} t\mathbb{1}_{\{T\leq t\}}\mathbb{1}_{\{T<J_{m+1}\}} \qquad (2.5.6) \\ &= T\mathbb{1}_{\{T<J_{m+1}\}}, \end{aligned}$$

de manera que las Igualdades (2.5.5) y (2.5.6) se obtienen de manera similiar que la Igualdad (2.5.4). Para A_m tenemos que

$$\begin{aligned} A_m\cap\{T < J_{m+1}\} &= \left(\bigcup_{t\in\mathbb{Q}} A_m(t)\right)\cap\{T < J_{m+1}\} \\ &= \bigcup_{t\in\mathbb{Q}}\big(A_m(t)\cap\{T < J_{m+1}\}\big) \\ &= \bigcup_{t\in\mathbb{Q}}\big(A(t)\cap\{T \leq t\}\cap\{T < J_{m+1}\}\big) \\ &= \left(\bigcup_{t\in\mathbb{Q}} A(t)\cap\{T \leq t\}\right)\cap\{T < J_{m+1}\} \\ &= A\cap\{T < J_{m+1}\}. \end{aligned}$$

Por lo tanto $T = T_m$ y $\mathbb{1}_A = \mathbb{1}_{A_m}$ en $\{T < J_{m+1}\}$. $\square$

Teorema 2.5.5 (Propiedad Fuerte de Markov.). *Sea $(X_t)_{t\geq 0}$ un proceso Markov(λ, Q) a tiempo continuo y sea T un tiempo de paro de $(X_t)_{t\geq 0}$. Entonces, condicionalmente a $\{T < \zeta\}$ y $\{X_T = i\}$, el proceso $(X_{T+t})_{t\geq 0}$ es Markov(δ_{ij}, Q) e independiente de $\mathscr{F}_T$ donde*

$$\delta_{ij} = \begin{cases} 1 & \text{si } j = i, \\ 0 & \text{en otro caso.} \end{cases}$$

Demostración:

Denotemos por $(Y_n)_{n\geq 0}$ a la cadena de saltos asociada al proceso $(X_t)_{t\geq 0}$. Por hipótesis, como $(X_t)_{t\geq 0}$ es una cadena de markov continua, en particular es un proceso minimal, si condicionamos $T < \zeta$, quiere decir que el tiempo de paro T sucede antes de que la cadena explote y en particular se cumple $T < \zeta \leq \infty$.

Debido a esto, para

Para el evento $\{T < \zeta\}$, definamos $\tilde{X}_t = X_{T+t}$ y denotemos por $(\tilde{Y}_n)_{n\geq 0}$ a la cadena de saltos asociada a $(\tilde{X}_t)_{t\geq 0}$ y como $(\tilde{S}_n)_{n\geq 1}$ a los tiempos de espera de $(\tilde{X}_t)_{t\geq 0}$. Condicionando en $X_T = i$, existe $m \geq 0$ tal que $\{J_m \leq T < J_{m+1}\}$ y por los lemas anteriores sabemos que para toda $m \geq 0$ tenemos que J_m es un tiempo de paro de $(X_t)_{t\geq 0}$ lo que implica que

$$\{J_m \leq T\} \cap \{X_T = i\} \in \mathscr{F}_T.$$

Sin pérdida de generalidad, consideremos $A \subseteq \{J_m \leq T\} \cap \{X_T = i\}$. Por el lema anterior podemos tomar $T = T_m$ y $\mathbb{1}_A = \mathbb{1}_{A_m}$ en $\{T < J_{m+1}\}$ donde T_m y A_m son $\mathscr{G}_m$-medibles.
En $\{J_m \leq T < J_{m+1}\}$ tenemos

$$\tilde{Y}_n = Y_{m+n} \ , \ \tilde{S}_1 = S_{m+1} - (T - J_m) \ , \ \tilde{S}_n = S_{m+n}.$$

De esta manera, condicionando en $Y_m = i$, S_{m+1} es independiente de $\mathscr{G}_m$ pues está generada por las variables aleatorias $Y_0, ..., Y_m$ y $S_1, ..., S_m$ y por otro lado $J_m = \sum_{k=0}^{m} S_k$ por lo que S_{m+1} también es independiente de $T_m - J_m$. Así, por la propiedad de pérdida de memoria de la distribución exponencial, tenemos

$$\begin{aligned} \mathbb{P}(S_{m+1} > s_1 + (T_m - J_m) | Y_m = i, S_{m+1} > T_m - J_m) &= \mathbb{P}(S_{m+1} > s_1 | Y_m = i) \\ &= e^{-q_i s_1} \\ &= \mathbb{P}(S_1 > s_1 \,|\, Y_m = i). \end{aligned}$$

Por lo tanto, por la propiedad de Markov de la cadena de saltos tenemos

$$
\begin{aligned}
&\mathbb{P}(\{\tilde{Y}_0 = i_0, \ldots, \tilde{Y}_n = i_n,\\
&\quad \tilde{S}_1 > s_1, \ldots, \tilde{S}_n > s_n\} \cap A \cap \{J_m \leq T < J_{m+1}\} \cap \{X_T = i\})\\
&= \mathbb{P}(\{Y_m = i_0, \ldots, Y_{m+n} = i_n, S_{m+1} > s_1 + (T_m - J_m),\\
&\quad S_{m+2} > s_2, \ldots, S_{m+n} > s_n\} \cap A_m \cap \{S_{m+1} > T - J_m\})\\
&= \mathbb{P}_i(\{Y_0 = i_0, \ldots, Y_n = i_n,\\
&\quad S_1 > s_1, \ldots, S_n > s_n)\mathbb{P}(A \cap \{J_m \leq T < J_{m+1}\} \cap \{X_T = i\}).
\end{aligned}
$$

Concluimos que $(X_{T+t})_{t\geq 0}$ es $Markov(\delta_i, Q)$ e independiente de $\mathscr{F}_T$. □

2.6 Recurrencia y transitoriedad

Consideremos a $(X_t)_{t\geq 0}$ una cadena de Markov minimal, con matriz generadora Q. Diremos que un estado i es *recurrente* si

$$\mathbb{P}_i(\{t \geq 0 : X_t = i\} \text{ no es acotado}) = 1.$$

Por otra parte, diremos que i es *transitorio* si

$$\mathbb{P}_i(\{t \geq 0 : X_t = i\} \text{ no es acotado}) = 0.$$

Consideremos $(X_t)_{t\geq 0}$ condicionado a $X_0 = i$, si el proceso explota en algún momento, el estado i no será recurrente, pues el estado ∞ se considera absorbente. El próximo teorema nos mostrará que las propiedades de recurrencia y transitoriedad están determinadas por la cadena de saltos asociada al proceso.

Teorema 2.6.1. *Sea $(X_t)_{t\geq 0}$ una cadena de Markov con matriz generadora Q. Entonces, tenemos las siguientes propiedades:*

(i) Si i es un estado recurrente de la cadena de saltos $(Y_n)_{n\geq 0}$, entonces i es recurrente para $(X_t)_{t\geq 0}$;

(ii) Si i es un estado transitorio de la cadena de saltos, entonces i es transitorio para $(X_t)_{t\geq 0}$;

(iii) todo estado es recurrente o transitorio;

(iv) recurrencia y transitoriedad son propiedades de clase.

Demostración:

(i) Supongamos que i es recurrente para $(Y_n)_{n\geq 0}$. Recordemos que si i es recurrente para una cadena de Markov a tiempo discreto, entonces

$$\mathbb{P}_i(Y_n = i \text{ para un número infinito de } n) = 1.$$

Condicionemos a $X_0 = i$ y observemos que el proceso $(X_t)_{t\geq 0}$ no puede explotar, pues de lo contrario i no sería recurrente. Por ello, la sucesión de tiempos de saltos cumple que

$$\lim_{n\to\infty} J_n = \infty.$$

Por otra parte, se sabe que

$$X_{J_n} = Y_n = i$$

para un número infinito de n, mostrando que el conjunto $\{t \geq 0 : X_t = i\}$ no es acotado y tiene probabilidad 1.

(ii) Supongamos ahora que i es transitorio para la cadena de saltos $(Y_n)_{n\geq 0}$, lo que implica

$$\mathbb{P}_i(Y_n = i \text{ para un número infinito de } n) = 0.$$

Así, si condicionamos $X_0 = i$ tenemos que con probabilidad 1 existe N tal que

$$N(\omega) = \sup\{n \geq 0 : Y_n(\omega) = i\} < \infty$$

con lo que obtenemos que para toda t en $\{t \geq 0 : X_t(\omega) = i\}$ se tiene que $t \leq J_{N(\omega)+1}$ para concluir que

$$\mathbb{P}_i(\{t \geq 0 : X_t = i\} \text{ no está acotado}) = 0.$$

(iii) De resultados básicos de cadenas de Markov a tiempo discreto, sabemos que todo estado es recurrente o transitorio, en particular para $(Y_n)_{n\geq 0}$ esto se cumple. Por otro lado, de la Sección 2.3 sabemos que podemos construir a una cadena de Markov a tiempo continuo a partir de su cadena de saltos asociada donde

$$X_t(\omega) = \begin{cases} Y_n(\omega) & \text{si } \ J_n(\omega) \leq t < J_{n+1}(\omega) \text{ para alguna n} \\ \infty & \text{en otro caso,} \end{cases}$$

por lo que podemos concluir que todo estado es recurrente o transitorio.

(iv) Por el mismo argumento del inciso anterior, recurrencia y transitoriedad son clases de equivalencia o comunicación de las cadenas de Markov a tiempo discreto, por lo que en particular lo es para la cadena de saltos y el proceso $(X_t)_{t\geq 0}$ se mueve a los mismos estados de $(Y_n)_{n\geq 0}$, por lo tanto, podemos concluir que recurrencia y transitoriedad son propiedades de clase.

$\square$

Definamos a T_i como el primer tiempo de llegada de $(X_t)_{t\geq 0}$ al estado i, definido por

$$T_i(\omega) = \inf\{t \geq J_1(\omega) : X_t(\omega) = i\}.$$

Para demostrar el siguiente teorema recurriremos a enunciar un resultado análogo para cadenas de Markov a tiempo discreto, recordemos que el primer tiempo de llegada de una cadena de Markov a tiempo discreto está definido como

$$N_i(\omega) = \inf\{n \geq 1 : Y_n(\omega) = i\}.$$

Teorema 2.6.2. *Consideremos $(Y_n)_{n\geq 0}$ una cadena de Markov a tiempo discreto con matriz de transición P. Entonces se cumple la siguiente dicotomía:*

(i) Si $P_i(N_i < \infty) = 1$, entonces i es recurrente y $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$;

(ii) Si $P_i(N_i < \infty) < 1$, entonces i es transitorio y $\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$,

donde $p_{ii}^{(n)}$ es la entrada (i,i) de la matriz P^n.

La demostración de este teorema se omitirá; puede ser consultada en [Nor97, p. 26].

Teorema 2.6.3. *Sea $(X_t)_{t\geq 0}$ una cadena de Markov con matriz generadora Q. Entonces se cumple la siguiente dicotomía:*

(i) si $q_i = 0$ o $P_i(T_i < \infty) = 1$, entonces i es recurrente y $\int_0^{\infty} p_{ii}(t)dt = \infty$;

(ii) si $q_i > 0$ y $P_i(T_i < \infty) < 1$, entonces i es transitorio y $\int_0^{\infty} p_{ii}(t)dt < \infty$.

Demostración:
Consideremos el Teorema de Tonelli (ver [Sam15, Teorema 1.4.6] y observemos

que

$$\begin{aligned}
\int_0^\infty p_{ii}(t)dt &= \int_0^\infty \mathbb{E}_i[\mathbb{1}_{\{X_t=i\}}]dt \\
&= \mathbb{E}_i\left[\int_0^\infty \mathbb{1}_{\{X_t=i\}}dt\right] \\
&= \mathbb{E}_i\left[\sum_{n=0}^\infty S_{n+1}\mathbb{1}_{\{Y_n=i\}}\right] \\
&= \sum_{n=0}^\infty \mathbb{E}[S_{n+1}\,|\,Y_n=i]\mathbb{P}_i(Y_n=i) \\
&= \frac{1}{q_i}\sum_{n=0}\pi_{ii}^{(n)},
\end{aligned}$$

el penúltimo paso debido a que S_{n+1} y Y_n son independientes. Ahora, sea $X_0=0$ y definamos $J=\sup(J_m : J_m < T_i(\omega))$, notemos que en el caso de que $T_i(\omega)$ exista y sea finito, el primer tiempo de llegada debe ser un tiempo de salto ya que

$$T_i = J + S_{m+1} = J_{m+1},$$

pues $X_t \neq i$ para toda t en el intervalo $[J_1, T_i)$, en particular, en $[J_m, T_i) = [J_m, J_{m+1})$. De está manera podemos calcular $\mathbb{P}_i(T_i<\infty)$ en términos de los tiempos de salto J_n. Dicho lo anterior, consideremos al primer tiempo de llegada al estado i de la cadena de saltos $(Y_n)_{n\geq 0}$ denotado por N_i y notemos que

$$\begin{aligned}
\mathbb{P}_i(T_i<\infty) &= \sum_{n=1}^\infty \mathbb{P}_i(T_i=J_n) \\
&= \sum_{n=1}^\infty \mathbb{P}_i(X_{J_1}\neq i,\dots,X_{J_{n-1}}\neq i, X_{J_n}=i) \\
&= \sum_{n=1}^\infty \mathbb{P}_i(Y_1\neq i,\dots,Y_{n-1}\neq i, Y_n=i) \\
&= \sum_{n=1}^\infty \mathbb{P}_i(N_i=n) \\
&= \mathbb{P}_i(N_i<\infty). \qquad (2.6.1)
\end{aligned}$$

Supongamos que $q_i=0$, esto implica que el proceso $(X_t)_{t\geq 0}$ no puede salir de i, es decir, i es un estado absorbente, de manera que $p_{ii}(t)=1$ para toda $t\geq 0$, por lo que en particular, i es recurrente pues $S_1=\infty$ y $\int_0^\infty p_{ii}(t)dt=\infty$.
Supongamos ahora que $q_i>0$, por los Teoremas 2.6.1 y 2.6.3 se sigue que:

(i) Si $\mathbb{P}_i(T_i < \infty) = 1$, el estado i es recurrente para el proceso $(X_t)_{t\geq 0}$ y

$$\int_0^\infty p_{ii}(t)dt = \frac{1}{q_i}\sum_{n=0} \pi_{ii}^{(n)} = \infty;$$

(ii) si $\mathbb{P}_i(T_i < \infty) < 1$, el estado i es transitorio para el proceso $(X_t)_{t\geq 0}$ y

$$\int_0^\infty p_{ii}(t)dt = \frac{1}{q_i}\sum_{n=0} \pi_{ii}^{(n)} < \infty.$$

$\square$

Con este teorema concluimos el primer capítulo referente a procesos de Markov a tiempo continuo cuyos resultados serán ocupados a lo largo de los siguientes capítulos.

Capítulo 3

Teoría del Potencial

Sea I un espacio discreto. Definamos una partición cualquiera de I dada por $I = D \cap \partial D$ y definamos $c : D \to \overline{\mathbb{R}}$ y $f : \partial D \to \overline{\mathbb{R}}$, en este contexto llamaremos a ∂D como la frontera de D. Observemos que las especificaciones de c y f definen una función por partes dada por

$$\tilde{c}_i = \begin{cases} c_i & \text{si } i \in D; \\ f_i & \text{si } i \in \partial D, \end{cases}$$

donde $c_i = c(i)$ y $f_i = f(i)$. Consideremos un proceso estocástico minimal $(X_t)_{t \in \mathrm{T}}$ con valores en I y $T = \overline{\mathbb{Z}}^+$ o $T = \overline{\mathbb{R}}^+$. Se recuerda que la condición "minimal" involucra al estado $\infty \in I$ (ver Sección 2.2. Sea $\mathscr{F}$ la σ-álgebra generada por $(X_t)_{t \in \mathrm{T}}$ y definamos $\tau = \inf\{t \geq 0 : X_t(\omega) \in \partial D\}$. Ahora, definamos $Y_t = \tilde{c}(X_t)\mathbb{1}_{T<\infty}$ y consideremos

$$\int_0^\tau c(x_t)dt + \tilde{c}(X_\tau)\mathbb{1}_{T<\infty}.$$

Si es que la integral puede ser especificada. ¿Bajo que condiciones existe la integral? ¿Cuándo es finita o infinita?. A partir de este capítulo, desarrollamos la correspondencia entre estos conceptos y la teoría de la probabilidad, estudiando una versión discreta de la teoría del potencial examinando sus propiedades en un contexto adecuado para las cadenas de Markov. Empezaremos introduciendo la idea de potenciales asociados a una cadena de Markov y cómo calcularlos.

3.1 Definición del Modelo

Sean $(X_n)_{n\geq 0} \sim MarkovTd(\lambda, P)$ una cadena de Markov a tiempo discreto con matriz de probabilidades de transición P y a $(X_t)_{t\geq 0} \sim MarkovTc(\lambda, Q)$ con distribución inicial λ una cadena de Markov con matriz generadora Q y matriz función de transición $P(t)$. Consideremos una cadena de Markov con valores en un espacio de estados numerable I y una partición de I dada por D y ∂D, donde ∂D es la frontera de D en I. En este contexto ∂D será visto como el conjunto de estados absorbentes en I de $(X_t)_{t\in\mathrm{T}}$ según el caso de T.

Definición 3.1.1. *Dadas dos funciones $c : D \to \mathbb{R}^+ \cup \{0\}$ y $f : \partial D \to \mathbb{R}^+ \cup \{0\}$, se define el potencial ϕ asociado a c y f como una función $\phi : I \to \overline{\mathbb{R}}^+$ donde $\phi_i(c, f)$ se especifica a continuación:*

1. *Discreto:*

$$\begin{aligned}\phi_i &= \mathbb{E}\left[\sum_{n<\tau} c(X_n) + f(X_\tau)\mathbb{1}_{\tau<\infty} \,|\, X_0 = i\right] \\ &= \mathbb{E}_i\left[\sum_{n<\tau} c(X_n) + f(X_\tau)\mathbb{1}_{\tau<\infty}\right].\end{aligned}$$

2. *Continuo:*

$$\phi_i = \mathbb{E}_i\left[\int_0^\tau c(X_t)dt + f(X_\tau)\mathbb{1}_{\tau<\infty}\right].$$

En ambos casos la variable aleatoria $\tau = \inf\{t \in [0,\infty] : X_t \in \partial D\}$ denota el tiempo de llegada a ∂D. Para el caso en que $X_t = \infty$ para alguna $t \in [0, \tau]$ se define $c(\infty) = 0$. Para el caso en que c y f están especificados, de manera simple denotaremos ϕ para referirnos al potencial asociado a c y f, de lo contrario de manera general denotaremos por $\phi(c, f)$ cuando c y f no están especificadas.

La interpretación más obvia del potencial en términos de costo: la cadena deambula en D hasta que llega a la frontera: mientras la cadena se mantenga en D, en el estado i se incurre un costo c_i por unidad de tiempo; cuando llega a la frontera ∂D, digamos a un estado j se incide en un costo final dado por f_i. Notemos que en ningún momento se asume que la cadena llega a la frontera, o que la frontera es no vacía.

Ejemplo 3.1.2. Sea $(X_n)_{n\geq 0}^{\infty}$ una caminata aleatoria simple sobre $\mathbb{Z}$, es decir, una sucesión de variables aleatorias idénticamente distribuidas sobre los enteros; la caminata empieza en un estado $X_0 = n_0$ y sólo puede saltar al mínimo entero mayor o al máximo entero menor. Las probabilidades de transición de la caminata están dadas por

$$p_{ij} = \begin{cases} q & \text{si} \quad j = i-1, \\ p & \text{si} \quad j = i+1, \\ 0 & \text{e.o.c.}, \end{cases}$$

con $q < p$ tales que $p+q = 1$. Supongamos que la caminata empieza en 0 e imaginemos que en todo momento que la caminata se mueve, se incurre en un gasto o costo por visitar al estado i dada por la función $c : \mathbb{Z} \to \mathbb{R}^+ \cup \{0\}$ definida por $c_i = \beta^i$, con $\beta > 0$. ¿Cuál será es el costo total esperado, resultado del costo generado por la caminata aleatoria sin que pare jamás? Definamos ϕ_0 como el costo total incurrido por la caminata aleatoria que empieza en 0 hasta infinito.

Si $q < p$ tenemos que la caminata aleatoria en la recta de los enteros en particular tiende a tomar los valores positivos, es decir, es más probable que la caminata salte a la derecha que la izquierda sin importar el valor de la caminata aleatoria al tiempo n. Para validar esta afirmación basta demostrar que la caminata aleatoria diverge casi seguramente, es decir

$$\mathbb{P}_0\big(\lim_{n\to\infty} X_n = \infty\big) = 1,$$

lo cual es equivalente a demostrar que

$$\mathbb{P}_0\big(\,\forall\, K > 0 \in \mathbb{Z},\, \exists\, N > 0 \in \mathbb{N} \; \textit{tal que } X_N > K\big) = 1.$$

Por un lado, la caminata aleatoria empezada en 0 cumple que

$$\mathbb{P}_0\big(X_n = i\big) = \begin{cases} \binom{n}{\frac{1}{2}(n+i)} p^{\frac{1}{2}(n+i)} q^{\frac{1}{2}(n-i)}. & \textit{si } |i| < n \text{ y} \quad n, i \textit{ son ambos par}, \\ 0 & \text{e.o.c..} \end{cases}$$

Por otro lado, notemos que la probabilidad de que la caminata aleatoria alcanze el valor K es positiva cuando $K \leq n$, de manera que sin pérdida de generalidad podemos suponer $0 < K < n$ con n y K ambos par. Entonces

$$\begin{aligned}
\mathbb{P}_0\big(\lim_{n\to\infty} X_n > K\big) &= 1 - \mathbb{P}_0\big(\lim_{n\to\infty} X_n \leq K\big) \\
&= 1 - \lim_{n\to\infty} \sum_{\substack{i=-n \\ i \text{ par}}}^{K} \mathbb{P}_0\big(X_n = i\big) \\
&= 1 - \lim_{n\to\infty} \sum_{\substack{i=-n \\ i \text{ par}}}^{K} \binom{n}{\frac{1}{2}(n+i)} p^{\frac{1}{2}(n+i)} q^{\frac{1}{2}(n-i)}. \qquad (3.1.1)
\end{aligned}$$

Para demostrar que $\mathbb{P}_0\big(\lim_{n\to\infty} X_n > K\big) = 1$, mostraremos que la serie en la igualdad (3.1.1) converge a 0. Notemos que para toda $0 \leq k \leq n$ el coeficiente binomial $\binom{n}{k}$ alcanza su máximo cuando $k = n/2$ para así acotar la serie en cuestión de la siguiente manera:

$$\begin{aligned}
\sum_{\substack{i=-n \\ i \text{ par}}}^{K} \binom{n}{\frac{1}{2}(n+i)} p^{\frac{1}{2}(n+i)} q^{\frac{1}{2}(n-i)} &\leq \sum_{\substack{i=-n \\ i \text{ par}}}^{K} \binom{n}{\frac{n}{2}} p^{\frac{1}{2}(n+i)} q^{\frac{1}{2}(n-i)} \\
&= \sum_{\substack{i=-n \\ i \text{ par}}}^{K} \binom{n}{\frac{n}{2}} (pq)^{\frac{n}{2}} (p/q)^{i} \\
&= \binom{n}{\frac{n}{2}} \Big((pq)^n\Big)^{\frac{1}{2}} \sum_{\substack{i=-n \\ i \text{ par}}}^{K} (p/q)^{i} \\
&= \left(\frac{n!}{(\frac{n}{2}!)^2}\right) \Big((pq)^n\Big)^{\frac{1}{2}} \sum_{\substack{i=-n \\ i \text{ par}}}^{K} (p/q)^{i}. \qquad (3.1.2)
\end{aligned}$$

Por un lado, la serie en (3.1.2) describe una serie geométrica de manera que podemos expresar la serie de manera alternativa como sigue:

$$\begin{aligned}
\sum_{\substack{i=-n \\ i \text{ par}}}^{K} (p/q)^i &= \sum_{\substack{i=0 \\ i \text{ par}}}^{K} (p/q)^i + \sum_{\substack{i=1 \\ i \text{ par}}}^{n} (q/p)^i \\
&= \sum_{i=0}^{K/2} \left(\left(\frac{p}{q} \right)^2 \right)^i + \left(\frac{q}{p} \right)^2 \left(\sum_{i=0}^{\frac{n}{2}-1} \left(\left(\frac{q}{p} \right)^2 \right)^i \right) \\
&= \frac{1 - \left((p/q)^2 \right)^{\frac{K}{2}+1}}{1-(p/q)^2} + \left(\frac{q}{p} \right)^2 \left(\frac{1 - \left((q/p)^2 \right)^{\frac{n}{2}-1+1}}{1-(q/p)^2} \right) \\
&= \frac{1-(p/q)^{K+2}}{1-(p/q)^2} + \left(\frac{q}{p} \right)^2 \left(\frac{(p/q)^2}{(p/q)^2} \right) \left(\frac{1-(q/p)^n}{1-(q/p)^2} \right) \\
&= \frac{1-(p/q)^{K+2}}{1-(p/q)^2} + \left(\frac{q}{p} \right)^2 \left(\frac{(p/q)^2-(q/p)^{n-2}}{(p/q)^2-1} \right) \\
&= \frac{1-(p/q)^{K+2}}{1-(p/q)^2} - \frac{1-(q/p)^n}{1-(p/q)^2} \\
&= \frac{(q/p)^n-(p/q)^{K+2}}{1-(p/q)^2}. \qquad (3.1.3)
\end{aligned}$$

Por la Fórmula de Stirling (ver [Nor97, p. 43]) podemos aproximar $n!$; para n suficientemente grande se sabe que

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e} \right)^n \quad \textit{cuando} \quad n \to \infty,$$

donde $a_n \sim b_n$ significa que $a_n/b_n \to 1$. Con lo anterior, si sustituimos el término obtenido en (3.1.3) en la ecuación (3.1.2) y aplicamos la Fórmula de Stirling para los términos que aplican, tenemos

$$
\begin{aligned}
\left(\frac{n!}{(\frac{n}{2}!)^2}\right)\left((pq)^n\right)^{\frac{1}{2}}\sum_{\substack{i=-n\\ i\text{ par}}}^{K}(p/q)^i &= \left(\frac{\sqrt{2\pi n}\,(n/e)^n}{\left(\sqrt{2\pi\frac{n}{2}}\left((n/2)/e\right)^{\frac{n}{2}}\right)^2}\right)\left((pq)^{\frac{1}{2}}\right)^n \\
&\qquad\times\left(\frac{(q/p)^n-(p/q)^{K+2}}{1-(p/q)^2}\right) \\
&= \left(\frac{\sqrt{2}}{\sqrt{\pi n}}\frac{(n/e)^n}{(n/2e)^n}\right)\left((pq)^{\frac{1}{2}}\right)^n \\
&\qquad\times\left(\frac{(q/p)^n-(p/q)^{K+2}}{1-(p/q)^2}\right) \\
&= \left(\frac{\sqrt{2}}{\sqrt{\pi n}}\frac{(2ne)^n}{(ne)^n}\right)\left((pq)^{\frac{1}{2}}\right)^n \\
&\qquad\times\left(\frac{(q/p)^n-(p/q)^{K+2}}{1-(p/q)^2}\right) \\
&= \left(\frac{\sqrt{2}}{\sqrt{\pi n}}2^n\right)\left((pq)^{\frac{1}{2}}\right)^n\left(\frac{(q/p)^n-(p/q)^{K+2}}{1-(p/q)^2}\right) \\
&\leq \sqrt{2}\left(2pq\right)^n\frac{(q/p)^n-(p/q)^{K+2}}{1-(p/q)^2}.
\end{aligned}
$$

Definamos la sucesión $A_n = (2pq)^n$, como $q = 1-p$ y $0 < q < p < 1$, en particular se cumple que $q < 1/2 < p < 1$, de manera que si multiplicamos por $2q$ la desigualdad anterior se tiene que $2q^2 < pq < 2pq < 2q < 1$. En consecuencia, se cumple que para toda $0 < n$, sin importar la paridad de n se tiene que $(2pq)^n < 1$, de manera que el límite cuando $n \to \infty$ de la sucesión A_n existe y es 0. Ahora, definamos B_n como:

$$
B_n = \frac{(q/p)^n-(p/q)^{K+2}}{1-(p/q)^2}.
$$

Como K es un valor fijo, en particular se cumple que $(p/q)^{K+2}$ es un valor constante con respecto a n y en particular, $(p/q)^{K+2} > 1$ y debido a que $(q/p) < 1$

pues $(p/q) > 1$ se sigue que

$$\begin{aligned}\lim_{n\to\infty} B_n &= \lim_{n\to\infty} \frac{(q/p)^n}{1-(p/q)^2} - \lim_{n\to\infty} \frac{(p/q)^{K+2}}{1-(p/q)^2}\\ &= 0 - \frac{(p/q)^{K+2}}{1-(p/q)^2}\\ &= -\frac{(p/q)^{K+2}}{1-(p/q)^2}.\end{aligned}$$

Por lo tanto, para toda $K > 0$

$$\begin{aligned}\lim_{n\to\infty} \sum_{\substack{i=-n\\ i \text{ par}}}^{K} \mathbb{P}_0\big(X_n = i\big) &= \lim_{n\to\infty} \sum_{\substack{i=-n\\ i \text{ par}}}^{K} \binom{n}{\frac{1}{2}(n+i)} p^{\frac{1}{2}(n+i)} q^{\frac{1}{2}(n-i)}\\ &\leq \lim_{n\to\infty} \sqrt{2}\, A_n B_n\\ &= \sqrt{2}\,(\lim_{n\to\infty} A_n)\,(\lim_{n\to\infty} B_n)\\ &= \sqrt{2}\;(0)\,(-\frac{(p/q)^{K+2}}{1-(p/q)^2})\\ &= 0,\end{aligned}$$

demostrando que

$$\begin{aligned}\mathbb{P}_0\big(\lim_{n\to\infty} X_n > K\big) &= 1 - \lim_{n\to\infty} \sum_{\substack{i=-n\\ i \text{ par}}}^{K} \mathbb{P}_0\big(X_n = i\big)\\ &= 1.\end{aligned}$$

Por lo tanto, la caminata aleatoria empezada en 0 con probabilidades de transición $q < p$ diverge casi seguramente a infinito.

Observemos que la serie

$$\sum_{i=0}^{\infty} c(X_i) = \sum_{i=0}^{\infty} \beta^{X_i} \qquad \text{y} \qquad \phi_0 = \mathbb{E}_0\left[\sum_{i=0}^{\infty} c(X_i)\right],$$

divergen si $\beta \geq 1$. Entonces, ¿para qué valores de β obtenemos un valor finito para ϕ_0? Antes de abordar esta pregunta, vale la pena hacer las siguientes observaciones. Sea $X_0 = 0$ y notemos que

$$
\begin{aligned}
\phi_0 &= \mathbb{E}\left[\sum_{i=0}^{\infty} c(X_i)\,|\,X_0 = 0\right] \\
&= \mathbb{E}\big[\,c(X_0) + \sum_{i=1}^{\infty} c(X_i)\,|\,X_0 = 0\big] \\
&= \mathbb{E}\big[\,\beta^{X_0} + \sum_{i=1}^{\infty} c(X_i)\,|\,X_0 = 0\big] \\
&= \mathbb{E}\big[\,\beta^{0} + \sum_{i=1}^{\infty} c(X_i)\,|\,X_0 = 0\big],
\end{aligned}
$$

donde por el Teorema de Convergencia Monótona [Nor97, Teorema 6.4.1,p. 241], se sigue que

$$
\begin{aligned}
&= 1 + \sum_{j\in\mathbb{Z}} \mathbb{P}(X_1 = j\,|\,X_0 = 0)\,\mathbb{E}\left[\sum_{i=1}^{\infty} c(X_i)\,|X_1 = j\right] \\
&= \beta^0 + p\,\mathbb{E}\left[\sum_{i=0}^{\infty} c(X_i)\,|\,X_1 = 1\right] + q\,\mathbb{E}\left[\sum_{i=0}^{\infty} c(X_i)\,|\,X_0 = -1\right] \\
&= 1 + p\,\phi_1 + q\,\phi_{-1}. \qquad (3.1.4)
\end{aligned}
$$

Por otra parte,

$$
\begin{aligned}
\phi_i &= \mathbb{E}_i\left[\sum_{i=0}^{\infty} c(X_i)\right] \\
&= \mathbb{E}\big[\,c(X_0) + c(X_1) + c(X_2) + \ldots\,|\,X_0 = i\big] \\
&= \mathbb{E}\big[\,\beta^{X_0} + \beta^{X_1} + \beta^{X_2} + \ldots\,|\,X_0 = i\big] \\
&= \mathbb{E}\big[\,\beta\left(\beta^{X_0-1} + \beta^{X_1-1} + \beta^{X_2-1} + \ldots\right)|\,X_0 = i\big] \\
&= \beta\,\mathbb{E}\big[\,\beta^{X_0-1} + \beta^{X_1-1} + \beta^{X_2-1} + \ldots\,|\,X_0 = i\big] \\
&= \beta\,\mathbb{E}\big[\,c(X_0 - 1) + c(X_1 - 1) + c(X_2 - 1) + \ldots\,|\,X_0 = i\big] \\
&= \beta\,\mathbb{E}\big[\,c(X_0) + c(X_1) + c(X_2) + \ldots\,|\,X_0 = i - 1\big] \\
&= \beta\,\phi_{i-1},
\end{aligned}
$$

para obtener que

$$\phi_1 = \beta\,\phi_0 \qquad , \qquad \phi_{-1} = \frac{\phi_0}{\beta}.$$

Si sustituimos ϕ_1 y ϕ_{-1} en la Ecuación (3.1.4), se sigue que

$$\begin{aligned} \phi_0 &= 1 + p\beta\,\phi_0 + q\,\frac{\phi_0}{\beta} \\ &= 1 + \left(p\beta + \frac{q}{\beta} \right) \phi_0, \end{aligned}$$

para obtener que

$$\phi_0 = \frac{\beta}{\beta - p\beta^2 - q}. \tag{3.1.5}$$

De esta manera logramos expresar a ϕ_0 en función de β. Recordemos que ϕ_0 es la esperanza de una suma cuyos términos son positivos y (3.1.5) toma valores positivos en el intervalo $(\frac{q}{p}, 1)$ para concluir que

$$\phi_0 = \begin{cases} \frac{\beta}{\beta - p\beta^2 - q} & \text{si}\, \beta \in (\frac{q}{p}, 1) \\ \infty & e.o.c. \end{cases}$$

Otra forma de validar este resultado es considerando el costo esperado hasta el tiempo n empezando desde 0, es decir,

$$\phi_0^{(n)} = \mathbb{E}_0 \left[\sum_{k=0}^{n} c(X_k) \right].$$

Procederemos a demostrar que $\phi_0^{(n)} = \phi_0$ cuando n tiende a infinito. Si proseguimos de manera similar a como obtuvimos la identidad (3.1.4), teniendo en cuenta que este para este caso estamos manipulando una suma finita, tenemos que

$$\phi_0^{(n)} = 1 + \phi_0^{(n-1)} \left(p\beta + \frac{q}{\beta} \right).$$

Es claro que,

$$\begin{aligned} \phi_0^{(1)} &= 1 + \phi_0^{(0)} \left(p\beta + \frac{q}{\beta} \right) \\ &= 1 + \left(p\beta + \frac{q}{\beta} \right). \end{aligned}$$

Afirmamos

$$\phi_0^{(n)} = \sum_{k=0}^{n} \left(p\beta + \frac{q}{\beta} \right)^k, \tag{3.1.6}$$

y observemos que tanto para $\phi_0^{(0)}$ y $\phi_0^{(1)}$ la afirmación es válida, por lo que si procedemos de manera inductiva suponiendo que (3.1.6) es válido, tenemos

$$\begin{aligned}
\phi_0^{(n+1)} &= 1 + \phi_0^{(n)} \left(p\beta + \frac{q}{\beta} \right) \\
&= 1 + \left(p\beta + \frac{q}{\beta} \right) \left(\sum_{k=0}^{n} \left(p\beta + \frac{q}{\beta} \right)^k \right) \\
&= 1 + \left(p\beta + \frac{q}{\beta} \right) \left(1 + \left(p\beta + \frac{q}{\beta} \right) + \ldots + \left(p\beta + \frac{q}{\beta} \right)^n \right) \\
&= \sum_{k=0}^{n+1} \left(p\beta + \frac{q}{\beta} \right)^k,
\end{aligned}$$

para concluir que (3.1.6) es válido para toda $n \geq 0$. Con esto, en este ejemplo tenemos que $\phi_0^{(0)}$ puede ser representada por una serie geométrica, la cuál sabemos converge si y solo si $|p\beta + q/\beta| < 1$. En particular sabemos $0 < p\beta + q/\beta$ de manera que podemos podemos ignorar el valor absoluto y obtener que

$$p\beta < 1 - \frac{q}{\beta} = p + q - \frac{q}{\beta} = p + q\left(\frac{\beta - 1}{\beta} \right).$$

Resolviendo la desigualdad para β ocurre que $\beta > q/p$, y concluimos que si β pertenece al complemento del intervalo $(q/p\,, 1)$ entonces ϕ_0 diverge, de lo contrario

$$\phi_0 = \lim_{n\to\infty} \sum_{k=0}^{n} \left(p\beta + \frac{q}{\beta} \right)^k$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \frac{1-\left(p\beta+\frac{q}{\beta}\right)^{n+1}}{1-\left(p\beta+\frac{q}{\beta}\right)} \\
&= \frac{1}{1-\left(p\beta+\frac{q}{\beta}\right)} \\
&= \frac{\beta}{\beta-p\beta^2-q}
\end{aligned}
$$

como ya se había mostrado anteriormente. ■

El potencial para cadenas de Markov puede variar conforme al contexto en el que se esté aplicando. En particular, esta herramienta es útil para resolver problemas con valores frontera; dadas dos funciones $c_i = c(i)$ y $f_i = f(i)$ definidas en una partición de I dada por D y ∂D respectivamente. La caminata inicia en algún estado en D y salta de un estado a otro hasta llegar a ∂D. La función evaluada c_i puede ser pensada como la carga o costo en el estado i, de manera que el potencial representa la carga o costo total hasta el momento en que la cadena llega a ∂D. En general, la interpretación sobre el potencial y las funciones c y f puede variar conforme al contexto en el que se aplique el potencial.

Teorema 3.1.3. *Supongamos que $(c_i : i \in D)$ y $(f_i : i \in \partial D)$ son dos funciones no negativas, es decir que para toda i, $c_i \geq 0$ y $f_i \geq 0$. Sea ϕ_i definida como en la Definición 3.1.1 para el caso discreto. Entonces:*

1. *El potencial $\phi = (\phi_i : i \in I)$ satisface:*

$$
\begin{cases}
\phi = P\phi + c & \text{en} \quad D, \\
\phi = f & \text{en} \quad \partial D.
\end{cases}
\tag{3.1.7}
$$

2. *Sea $\psi = (\psi_i : i \in I)$ que satisface:*

$$
\begin{cases}
\psi \geq P\psi + c & \text{en} \quad D, \\
\psi \geq f & \text{en} \quad \partial D,
\end{cases}
\tag{3.1.8}
$$

   y $\psi_i \geq 0$ para toda i en I. Entonces $\psi_i \geq \phi_i$ para toda i en I.

3. *Si $\mathbb{P}_i(\tau < \infty) = 1$ para toda i en I entonces ϕ existe y es la única solución acotada de (3.1.7).*

Demostración:

1. Sea i en ∂D, como τ es una variable discreta notemos que en este caso, se cumple que $\tau = 0 < \infty$ y en particular se tiene que $\mathbb{P}_i(\tau < \infty) = 1$ y $\mathbb{P}_i(\tau = 0) = 1$ para así obtener que

$$\begin{aligned}
\phi_i &= \mathbb{E}_i\left[\sum_{n<\tau} c(X_n) + f(X_\tau)\mathbb{1}_{\tau<\infty}\right] \\
&= \mathbb{E}_i\left[\sum_{n=1}^{\tau-1} c(X_n)\mathbb{1}_{n<\tau} + f(X_\tau)\mathbb{1}_{\tau<\infty}\right] \\
&= \mathbb{E}_i\big[f(X_\tau)\mathbb{1}_{\tau<\infty}\big] \\
&= \mathbb{E}_i\big[f(i)\mathbb{1}_{\tau<\infty}\big] \\
&= \mathbb{E}_i\big[f(i)\big] = f_i.
\end{aligned}$$

Por lo tanto, para toda i en ∂D tenemos que $\phi_i = f_i$.
Ahora, sea i en D, por el Teorema de Convergencia Monótona tenemos:

$$\begin{aligned}
\phi_i &= \mathbb{E}_i\left[\sum_{n<\tau} c(X_n) + f(X_\tau)\mathbb{1}_{\tau<\infty}\right] \\
&= \mathbb{E}\left[\sum_{n<\tau} c(X_n) + f(X_\tau)\mathbb{1}_{\tau<\infty}|X_0 = i\right] \\
&= \mathbb{E}\big[c(X_0) + \sum_{1\leq n<\tau} c(X_n) + f(X_\tau)\mathbb{1}_{\tau<\infty}|X_0 = i\big] \\
&= c(X_0) + \sum_{j\in I} p_{ij}\mathbb{E}\left[\sum_{1\leq n<\tau} c(X_n) + f(X_\tau)\mathbb{1}_{\tau<\infty}|X_1 = j\right]. \qquad (3.1.9)
\end{aligned}$$

Recordemos que p_{ij} denota la entrada (i, j) de la matriz de probabilidades de transcición P asociada al proceso. Por otra parte, por la propiedad de

Markov tenemos

$$\begin{aligned}\mathbb{E}\Big[\sum_{1\leq n<\tau} c(X_n)+f(X_\tau)\mathbb{1}_{\tau<\infty}|X_1=j\Big]\\ &=\mathbb{E}\left[\sum_{n<\tau} c(X_n)+f(X_\tau)\mathbb{1}_{\tau<\infty}|X_0=j\right]\\ &=\phi_j,\end{aligned}$$

con lo que si sustituimos la anterior en (3.1.9) obtenemos que:

$$\begin{aligned}\phi_i &= c(X_0)+\sum_{j\in I} p_{ij}\mathbb{E}\left[\sum_{n<\tau} c(X_n)+f(X_\tau)\mathbb{1}_{\tau<\infty}|X_0=j\right]\\ &= c_i+\sum_{j\in I} p_{ij}\phi_j.\end{aligned}$$

Por lo tanto, para toda i en D:

$$\phi = P\phi + c.$$

2. Definamos el costo esperado hasta el tiempo n como:

$$\phi_i^{(n)} = \mathbb{E}_i\left[\sum_{k=0}^{n} c(X_k)\mathbb{1}_{k<\tau}+f(X_\tau)\mathbb{1}_{\tau\leq n}\right]$$

Notemos que si proseguimos de manera similar al punto anterior obtenemos que

$$\begin{aligned}\phi_i^{(n)} &= \mathbb{E}_i\left[\sum_{k=0}^{n} c(X_k)\mathbb{1}_{k<\tau}+f(X_\tau)\mathbb{1}_{\tau\leq n}\right]\\ &= c(X_0)+\sum_{j\in I} p_{ij}\mathbb{E}\left[\sum_{k=1}^{n} c(X_k)\mathbb{1}_{k<\tau}+f(X_\tau)\mathbb{1}_{\tau\leq n}|X_1=j\right]\\ &= c_i+\sum_{j\in I} p_{ij}\mathbb{E}\left[\sum_{k=0}^{n-1} c(X_k)\mathbb{1}_{k<\tau}+f(X_\tau)\mathbb{1}_{\tau\leq n-1}|X_0=j\right]\\ &= c_i+\sum_{j\in I} p_{ij}\phi_j^{(n-1)},\end{aligned}$$

esto último es válido, mientras i pertenezca a D. Consideremos ahora el caso en el que i pertenece a ∂D

$$\phi_i^{(n)} = \mathbb{E}_i\left[\sum_{k=0}^{n} c(X_k)\mathbb{1}_{k<\tau} + f(X_\tau)\mathbb{1}_{\tau\leq n}\right] = \mathbb{E}_i\big[f(X_\tau)\mathbb{1}_{\tau\leq n}\big] = f_i,$$

para concluir que

$$\phi^{(n)} = \begin{cases} c + P\phi^{(n-1)} & en \quad D, \\ f & en \quad \partial D. \end{cases}$$

Definamos a $(S_n)_{n\geq 0}$ como $S_n = \sum_{k=0}^{n} c(X_k)\mathbb{1}_{k<\tau} + f(X_\tau)\mathbb{1}_{\tau\leq n}$ de manera que

$$\lim_{n\to\infty} S_n = \sum_{n<\tau} c(X_n) + f(X_\tau)\mathbb{1}_{\tau\leq\infty},$$

y dado que $(S_m)_{m\geq 0}$ es una sucesión monotona (pues c_i y f_i son funciones no negativas), por el teorema de convergencia monótona tenemos

$$\begin{aligned}
\phi_i^{(n)} = \lim_{n\to\infty}\mathbb{E}_i\big[S_n\big] = \lim_{n\to\infty}\mathbb{E}_i&\left[\sum_{k=0}^{n} c(X_k)\mathbb{1}_{k<\tau} + f(X_\tau)\mathbb{1}_{\tau\leq n}\right] \\
= \mathbb{E}_i&\left[\sum_{n<\tau} c(X_n) + f(X_\tau)\mathbb{1}_{\tau<\infty} + c(\infty)\mathbb{1}_{\tau=\infty}\right] \\
= \mathbb{E}_i&\left[\sum_{n<\tau} c(X_n) + f(X_\tau)\mathbb{1}_{\tau<\infty}\right] \\
= \phi_i&,
\end{aligned}$$

es decir, $\lim_{n\to\infty}\phi^{(n)} = \phi$. Ahora, sea ψ tal que satisface (3.1.8) y $\psi \geq 0 = \phi^{(0)}$; demostraremos por inducción que $\psi \geq \phi^{(n)}$ para toda $n \geq 1$.
En el caso base $n = 1$, tenemos que:

$$\psi \geq P\psi + c \geq P\phi^{(0)} + c = \phi^{(1)},$$

con lo que el caso base es válido. De manera inductiva, supongamos que la proposición es válida para k, es decir,

$$\psi \geq \phi^{(k)}.$$

Entonces, si multiplicamos por P a la izquierda en ambos lados y sumamos c de la misma manera, obtenemos que:

$$P\psi + c \geq P\phi^{(k)} + c = \phi^{(k+1)},$$

por lo tanto, para toda $n \geq 1$ tenemos

$$\psi \geq \phi^{(n)},$$

de manera que

$$\psi \geq \lim_{n\to\infty} \phi^{(n)} = \phi.$$

3. Sea ψ tal que satisface (3.1.8) y que para toda i en I tenemos $|\psi_i| \leq C$, definamos a $(M_n)_{n\geq 0}$ como:

$$M_n = \sum_{k=0}^{n-1} c(X_k)\mathbb{1}_{k<\tau} + f(X_\tau)\mathbb{1}_{\tau<n} + \psi(X_n)\mathbb{1}_{n\leq\tau}.$$

Ahora, consideremos a $\mathscr{F}_n$, la σ-álgebra generada por las variables aleatorias $(X_1, X_2, ..., X_n)$. Entonces

$$\begin{aligned}\mathbb{E}\left[M_{n+1} \,|\, \mathscr{F}_n\right] &= \mathbb{E}\left[\sum_{k=0}^{n} c(X_k)\mathbb{1}_{k<\tau} + f(X_\tau)\mathbb{1}_{\tau<n+1} + \psi(X_{n+1})\mathbb{1}_{n+1\leq\tau} \,|\, \mathscr{F}_n\right]\\ &= \mathbb{E}\left[\sum_{k=0}^{n-1} c(X_k)\mathbb{1}_{k<\tau} + c(X_n)\mathbb{1}_{n<\tau} + f(X_\tau)\mathbb{1}_{\tau<n}\right.\\ &\qquad\qquad \left. + f(X_\tau)\mathbb{1}_{\tau=n} + \psi(X_{n+1})\mathbb{1}_{n<\tau} \,|\, \mathscr{F}_n\right],\end{aligned}$$

donde en la última igualdad, todos los términos excepto $\psi(X_{n+1})\mathbb{1}_{n<\tau}$ son

$\mathscr{F}_n$-medibles, es decir,

$$\begin{aligned}\mathbb{E}\big[M_{n+1}|\mathscr{F}_n\big] &= \sum_{k=0}^{n-1} c(X_k)\mathbb{1}_{k<\tau} + c(X_n)\mathbb{1}_{n<\tau} + f(X_\tau)\mathbb{1}_{\tau<n} \\ &\qquad + f(X_\tau)\mathbb{1}_{\tau=n} + \mathbb{E}\big[\psi(X_{n+1})\mathbb{1}_{n<\tau}\,|\,\mathscr{F}_n\big] \\ &= \sum_{k=0}^{n-1} c(X_k)\mathbb{1}_{k<\tau} + c(X_n)\mathbb{1}_{n<\tau} + f(X_\tau)\mathbb{1}_{\tau<n} \\ &\qquad + f(X_\tau)\mathbb{1}_{\tau=n} + (P\psi)(X_n)\mathbb{1}_{n<\tau} \\ &= \sum_{k=0}^{n-1} c(X_k)\mathbb{1}_{k<\tau} + f(X_\tau)\mathbb{1}_{\tau<n} \\ &\quad + (P\psi + c)(X_n)\mathbb{1}_{n<\tau} + f(X_\tau)\mathbb{1}_{\tau=n},\end{aligned}$$

pero

$$(P\psi + c)(X_n)\mathbb{1}_{n<\tau} + f(X_\tau)\mathbb{1}_{\tau=n} \leq \psi(X_n)\mathbb{1}_{n\leq\tau},$$

lo que implica que

$$\begin{aligned}&\sum_{k=0}^{n-1} c(X_k)\mathbb{1}_{k<\tau} + f(X_\tau)\mathbb{1}_{\tau<n} + (P\psi + c)(X_n)\mathbb{1}_{n<\tau} + f(X_\tau)\mathbb{1}_{\tau=n} \\ &\leq \sum_{k=0}^{n-1} c(X_k)\mathbb{1}_{k<\tau} + f(X_\tau)\mathbb{1}_{\tau<n} + \psi(X_n)\mathbb{1}_{n\leq\tau} = M_n.\end{aligned}$$

Probaremos ahora la unicidad de la solución de la ecuación (2.5) bajo el supuesto $\mathbb{P}_i(\tau < \infty) = 1$ para toda i en I. Notemos que si ψ satisface (3.1.8) con igualdad, M_n sería una martingala, por lo que se obtiene lo siguiente:

$$\psi_i = \mathbb{E}_i\big[M_0\big] = \mathbb{E}_i\big[M_n\big] = \phi_i^{(n-1)} + \mathbb{E}_i\big[\psi(X_n)\mathbb{1}_{n\leq\tau}\big],$$

pero por otra parte tenemos

$$|\mathbb{E}_i\big[\psi(X_n)\mathbb{1}_{n\leq\tau}\big]| \leq \mathbb{E}_i\big[C\mathbb{1}_{n\leq\tau}\big] = C\mathbb{E}_i\big[\mathbb{1}_{n\leq\tau}\big] = C\,\mathbb{P}_i\big(n\leq\tau\big)$$

y sabemos que $\mathbb{P}_i(n \leq \tau) = 1 - \mathbb{P}_i(\tau < n)$ y por hipótesis $\mathbb{P}_i(\tau < \infty) = 1$ por lo que obtenemos que

$$\lim_{n\to\infty} |\mathbb{E}_i\big[\psi(X_n)\mathbb{1}_{n\leq\tau}\big]| = \lim_{n\to\infty} C[1 - \mathbb{P}_i(\tau < n)] = 0.$$

Concluimos

$$\begin{aligned} \psi_i &= \lim_{n\to\infty} \phi_i^{(n-1)} + \mathbb{E}_i\big[\psi(X_n)\mathbb{1}_{n\leq\tau}\big] \\ &= \phi_i \\ i.e. \qquad \psi &= \phi. \end{aligned}$$

Por lo tanto, si $\mathbb{P}_i(\tau < \infty) = 1$, ϕ existe, es acotada y es la única solución que satisface (3.1.7).

□

El teorema anterior es válido para el caso discreto. A continuación presentamos la versión análoga para el caso de cadenas a tiempo continuo. Recordatorio: diremos que un proceso $(X_t)_{t\geq 0}$ es minimal si existe ζ un tiempo de paro tal que para toda $t \geq \zeta$, tenemos que $X_t = \infty$ (tal como lo definimos en el capítulo 1, en la Sección 2.2).

Teorema 3.1.4. *Sea $(X_t)_{t\geq 0}$ minimal y $c = (c_i : i \in D)$ y $f = (f_i : i \in \partial D)$ no negativos. Sea ϕ_i definida como en la Definición 3.1.1 para el caso continuo. Entonces $\phi = (\phi_i : i \in I)$ tal que $\phi_i < \infty$ para toda i en I, es la solución mínima no negativa de:*

$$\begin{cases} -Q\phi = c & en \quad D, \\ \quad \phi = f & en \quad \partial D. \end{cases} \tag{3.1.10}$$

Si $\phi_i = \infty$ para alguna $i \in I$ entonces la solución a la ecuación (3.1.10) existe no tiene solución finita no negativa. Más aún, si $\mathbb{P}_i(\tau < \infty) = 1$ para toda i en I entonces (3.1.10) tiene a lo más una solución acotada.

Demostración:
Denotemos por $(Y_n)_{n\geq 0}$ a la cadena de saltos asociada a $(X_t)_{t\geq 0}$, a la sucesión $S_1, S_2, \ldots$ como los tiempos de espera de $(X_t)_{t\geq 0}$ y a Π la matriz de saltos asociada

a la misma, entonces

$$\begin{aligned}
\int_0^\tau c(X_t)dt + f(X_\tau)\mathbb{1}_{\tau<\infty} &= \sum_{i=0}^{N-1}\int_{J_i}^{J_{i+1}} c(X_t)dt + f(X_\tau)\mathbb{1}_{\tau<\infty}\\
&= \sum_{i=0}^{N-1}\int_{J_i}^{J_{i+1}} c(Y_i)dt + f(Y_N)\mathbb{1}_{N<\infty}\\
&= \sum_{i=0}^{N-1} c(Y_i)\int_{J_i}^{J_{i+1}} dt + f(Y_N)\mathbb{1}_{N<\infty}\\
&= \sum_{i=0}^{N-1} c(Y_i)S_{i+1} + f(Y_N)\mathbb{1}_{N<\infty},
\end{aligned}$$

donde N es la primera vez que $(Y_n)_{n\geq 0}$ toca a ∂D. Dada esta consideración, calculemos lo siguiente:

$$\begin{aligned}
&\mathbb{E}_i\left[\int_0^\tau c(X_t)dt + f(X_\tau)\mathbb{1}_{\tau<\infty}\right] = \mathbb{E}_i\left[\sum_{n<N} c(Y_n)S_{n+1} + f(Y_N)\mathbb{1}_{N<\infty}\right]\\
&= \mathbb{E}_i\big[c(Y_0)S_1 + \sum_{1\leq n<N} c(Y_n)S_{n+1} + f(Y_N)\mathbb{1}_{N<\infty}\big]\\
&= \frac{c_i}{q_i} + \mathbb{E}_i\Big[\sum_{1\leq n<N} c(Y_n)S_{n+1} + f(Y_N)\mathbb{1}_{N<\infty}\Big]\\
&= \frac{c_i}{q_i} + \sum_{j\neq i}\mathbb{P}_i\big(Y_1 = j\,|\,Y_0 = i\big)\,\mathbb{E}_i\left[\sum_{1\leq n<N} c(Y_n)S_{n+1} + f(Y_N)\mathbb{1}_{N<\infty}\right]\\
&= \frac{c_i}{q_i} + \sum_{j\neq i}\frac{q_{ij}}{q_i}\,\mathbb{E}\left[\sum_{n<N} c(Y_n)S_{n+1} + f(Y_N)\mathbb{1}_{N<\infty}\,|\,Y_0 = j\right]\\
&= \frac{c_i}{q_i} + \sum_{j\neq i}\frac{q_{ij}}{q_i}\,\phi_j. \qquad (3.1.11)
\end{aligned}$$

Si definimos a $\tilde{c}_i = \frac{c_i}{q_i}$, tenemos que

$$\phi_i = \tilde{c}_i + \sum_{j\in I}\frac{q_{ij}}{q_i}\,\phi_j,$$

$$i.e., \qquad \phi = \tilde{c} + \Pi\phi, \qquad (3.1.12)$$

de manera que por el teorema anterior (la versión discreta), ϕ existe y es la solución mínima no negativa de la ecuación (3.1.8), más aún, si $\mathbb{P}_i(\tau < \infty) = 1$ entonces (3.1.8) tiene a lo más una solución acotada. De esto último, no es díficil ver que, en efecto, ϕ satisface la ecuación (3.1.8) pues de (3.1.12) se concluye

$$q_i\,\phi_i = c_i + \sum_{j\neq i} q_{ij}\,\phi_j.$$

Por lo tanto,

$$\begin{aligned} c_i &= q_i\,\phi_i - \sum_{j\neq i} q_{ij}\,\phi_j \\ &= -q_{ii} - \sum_{j\in I} q_{ij}\,\phi_j \\ &= -\sum_{j\in I} q_{ij}\,\phi_j \end{aligned}$$

$$i.e., \qquad -Q\phi = c.$$

□

El siguiente teorema hace referencia al costo esperado con cierto factor de descuento aplicado a un futuro, como es de esperarse tenemos la versión discreta y continua:

Teorema 3.1.5. *Supongamos que* $(c_i : i \in I)$ *es acotado. Sea* $\alpha \in (0,1)$ *y definamos*

$$\phi_i = \mathbb{E}_i\left[\sum_{n=0}^{\infty} \alpha^n c(X_n)\right]$$

entonces $\phi = (\phi_i : i \in I)$ *es la única solución acotada de*

$$\phi = \alpha P\phi + c. \tag{3.1.13}$$

Demostración:
Supongamos que $|c_i| \leq M < \infty$ para toda i en I, como α está en $(0,1)$ tenemos

$$\begin{aligned} |\phi_i| &= \left|\mathbb{E}_i\left[\sum_{n=0}^{\infty} \alpha^n c(X_n)\right]\right| \\ &\leq \sum_{n=0}^{\infty} \alpha^n M = \frac{M}{1-\alpha} \end{aligned}$$

de manera que $|\phi_i|$ está acotada. Ahora, por el Teorema de Convergencia monótona y la propiedad de Markov se tiene que

$$\begin{aligned}
\phi_i &= \mathbb{E}_i\left[\sum_{n=0}^{\infty}\alpha^n c(X_n)\right] \\
&= \mathbb{E}_i\left[\alpha\left(\sum_{n=0}^{\infty}\alpha^{n-1}c(X_n)\right)\right] \\
&= \alpha\,\mathbb{E}_i\Big[\alpha^{-1}c(X_0)+\sum_{n=1}^{\infty}\alpha^n c(X_n)\Big] \\
&= c_i+\alpha\sum_{j\neq i}p_{ij}\,\mathbb{E}_i\left[\sum_{n=1}^{\infty}\alpha^n c(X_n)\,|X_1=j\right] \\
&= c_i+\alpha\sum_{j\neq i}p_{ij}\,\mathbb{E}_j\left[\sum_{n=0}^{\infty}\alpha^n c(X_n)\right] \\
&= c_i+\alpha\sum_{j\neq i}p_{ij}\,\phi_j
\end{aligned}$$

$$i.e., \qquad \phi = c+\alpha P\phi.$$

Por otra parte, sea ψ otra solución de (3.1.13) acotada y definamos a $C=\sup_i|\psi_i-\phi_i|$, como ϕ es acotada, C existe y es finito. Ahora, observemos que

$$\begin{aligned}
\psi-\phi &= (c+\alpha P\psi)-(c+\alpha P\phi) \\
&= \alpha P(\psi-\phi),
\end{aligned}$$

lo cual implica que

$$\begin{aligned}
|\psi_i-\phi_i| &= \alpha\left|\sum_{j\neq i}p_{ij}(\psi_i-\phi_i)\right| \\
&\leq \alpha\left|\sum_{j\neq i}p_{ij}C\right| \\
&\leq \alpha C,
\end{aligned}$$

mostrando que αC es una cota superior de $|\psi_i-\phi_i|$. Como $C=\sup_i|\psi_i-\phi_i|$ se cumple que

$$|\psi_i-\phi_i|\leq C\leq \alpha C.$$

Por un lado, si α está en $(0,1)$, en particular se cumple que $\alpha C \leq C$ y por otro lado tenemos que $C \leq \alpha C$, por lo tanto se concluye que $C = 0$ y $\psi = \phi$. $\square$

Teorema 3.1.6. *Supongamos que (X_t) no explota y que $(c_i : i \in I)$ está acotada. Sea λ en $(0,\infty)$ y definamos a*

$$\phi_i = \mathbb{E}_i\left[\int_0^\infty e^{-\lambda t} c(X_t)dt\right],$$

entonces $\phi = (\phi_i : i \in I)$ es la única solución acotada de

$$(\lambda - Q)\phi = c. \tag{3.1.14}$$

Demostración:
Definamos a

$$\tilde{X}_t = \begin{cases} X_t & \text{si} \quad t < T, \\ \partial & \text{si} \quad t \geq T \end{cases}$$

y sea la cadena de saltos $(\tilde{Y}_n)_{n\geq 0}$ y los tiempos de espera asociados a $\tilde{X}_t$. Observemos que si condicionamos a $\tilde{Y}_n = i$ tenemos que $\tilde{Y}_n \neq \partial$ si y solo si $S_{n+1} < T$, es decir, $\tilde{S}_{n+1} = T \wedge S_{n+1}$. Como T y S_{n+1} son variables aleatorias independientes entre sí de distribución exponencial de parámetros λ y $q(Y_n)$ respectivamente, obtenemos que $\tilde{S}_{n+1}$ es una variable aleatoria exponencial de parámetro $\tilde{q}(\tilde{Y}_n)$ con $\tilde{q}(\tilde{Y}_n) = q(Y_n) + \lambda$.
Observemos que ∂ es un estado absorbente de $(\tilde{X}_t)_{t\geq 0}$ pues $\mathbb{P}(T < \infty) = 1$. Si definimos a $\tilde{q}_\delta = 0$ y a $\tilde{q}_{i\partial} = \lambda$ el proceso $(\tilde{X}_t)_{t\geq 0}$ es una cadena de Markov con matriz generadora $\tilde{Q} = \lambda + Q$.

Sea $\tilde{N}$ el tiempo de llegada al estado ∂ de $(\tilde{Y}_n)$, entonces

$$
\begin{aligned}
\phi_i &= \mathbb{E}_i\left[\int_0^\infty c(\tilde{X}_t)dt\right] \\
&= \mathbb{E}_i\left[\sum_{n<N} c(\tilde{Y}_n)\tilde{S}_{n+1} + c_\partial \tilde{S}_N\right] \\
&= \mathbb{E}_i\left[\sum_{n<N} c(\tilde{Y}_n)\tilde{S}_{n+1}\right] \\
&= \mathbb{E}_i\big[c(\tilde{Y}_0)S_1 + \sum_{1\le n<N} c(\tilde{Y}_n)\tilde{S}_{n+1}\big] \\
&= \frac{c_i}{\tilde{q}_i} + \mathbb{E}_i\left[\sum_{1\le n<N} c(\tilde{Y}_n)\tilde{S}_{n+1}\right] \\
&= \frac{c_i}{\tilde{q}_i} + \sum_{j\neq i}\frac{q_{ij}}{\tilde{q}_i}\,\mathbb{E}_j\left[\sum_{n<N} c(Y_n)S_{n+1}\,|\,Y_0 = j\right] \\
&= \frac{c_i}{\tilde{q}_i} + \sum_{j\neq i}\frac{q_{ij}}{\tilde{q}_i}\,\phi_j,
\end{aligned}
$$

lo cual implica que

$$\tilde{q}_i\phi_i = c_i + \sum_{j\neq i}\tilde{q}_{ij}\phi_j.$$

Así obtenemos

$$
\begin{aligned}
c_i &= \tilde{q}_i\phi_i - \sum_{j\neq i}\tilde{q}_{ij}\phi_j \\
&= (\lambda - \sum_{i\in I}\tilde{q}_{ij})\phi_j
\end{aligned}
$$

$$i.e., \qquad c = (\lambda - Q)\phi,$$

pues $\tilde{q}_{ij} = q_{ij}$. Por otra parte, notemos que T es el tiempo de llegada a ∂ el cual es finito con probabilidad 1, debido a esto tenemos

$$\begin{aligned}\phi_i &= \mathbb{E}_i\left[\int_0^\infty c(\tilde{X}_t)dt\right]\\ &= \mathbb{E}_i\left[\int_0^T c(\tilde{X}_t)dt + \int_T^\infty c(\tilde{X}_t)dt\right]\\ &= \mathbb{E}_i\left[\int_0^T c(\tilde{X}_t)dt\right]\\ &= \mathbb{E}_i\left[\int_0^\infty c(\tilde{X}_t)\,\mathbb{1}_{t<T}\,dt\right].\end{aligned}$$

En el intervalo $(0,T)$ se cumple $\tilde{X}_t = X_t$ y como c_i es no negativo para toda i en I, por el Teorema de Fubini y la independencia de T se sigue que

$$\begin{aligned}\mathbb{E}_i\left[\int_0^\infty c(\tilde{X}_t)\,\mathbb{1}_{t<T}\,dt\right] &= \int_0^\infty \mathbb{E}_i\big[c(X_t)\big]\mathbb{P}\big(t<T\big)dt\\ &= \int_0^\infty \mathbb{E}_i\big[c(X_t)\big]e^{-\lambda t}dt\\ &= \int_0^\infty \mathbb{E}_i\big[c(X_t)e^{-\lambda t}\big]dt\\ &= \mathbb{E}_i\left[\int_0^\infty c(X_t)e^{-\lambda t}dt\right],\end{aligned}$$

dado que $\phi_i \le C\int_0^\infty e^{-\lambda t}dt \le C/\lambda$ para toda i en I. Por lo tanto, por el Teorema 3.1.4 se concluye que ϕ es la única solución a la ecuación

$$\tilde{Q} = (\lambda - Q)\phi = c.$$

Para el caso en que c es acotada con valores negativos consideremos a

$$\phi_i^{\pm} = \mathbb{E}_i\left[\int_0^\infty e^{-\lambda t}c^{\pm}(X_t)dt\right],$$

donde $c_i^{\pm} = (\pm c_i)\vee 0$ de manera que $\phi = \phi^+ - \phi^-$. Por lo anterior no es difícil ver que

$$(\lambda - Q)\phi^{\pm} = c^{\pm}$$

lo que implica que

$$\begin{aligned}(\lambda-Q)\phi^{+}-(\lambda-Q)\phi^{-} &= (\lambda-Q)(\phi^{+}-\phi^{-})\\ &= c^{+}-c^{-}\\ &= c.\end{aligned}$$

Para finalizar, si ψ es acotada y cumple que $(\lambda-Q)\psi=c$ tenemos que $(\lambda-Q)(\psi-\phi)=0$ lo que implica que $\psi=\phi$. Por lo tanto ϕ es la única solución acotada de (3.1.14). □

De los teoremas anteriores podemos concluir que, dados $c=(c_i : i\in D)$ y $f=(f_i : i\in\partial D)$, para calcular el potencial asociado el problema se puede resumir en resolver un sistema de ecuaciones lineales que a su vez dependerán de las clases de comunicación de una cadena de Markov.

A partir de este momento nos preguntaremos por algunos aspectos estructurales del conjunto de potenciales de una cadena de Markov dada. Sean $(X_n)_{n\geq 0}\sim MarkovTd(\lambda,P)$ una cadena de Markov a tiempo discreto con matriz de probabilidades de transición P y a $(X_t)_{t\geq 0} simMarkovTc(\lambda,Q)$ una cadena de Markov con matriz generadora Q, matriz función de transición $P(t)$.

Definición 3.1.7. *Denotaremos por $G=(g_{ij} : i,j\in I)$ a la matriz de Green dada por*

$$G=\sum_{n=0}^{\infty}P^{n},$$

para el caso discreto. Similarmente, para el caso continuo

$$G=\int_0^{\infty}P(t)dt,$$

donde

$$g_{ij}=\int_0^{\infty}p_{ij}(t)dt.$$

Sean $(X_n)_{n\geq 0}\sim MarkovTd(\lambda,P)$ con valores en el espacio de estados I y $c: I\to\mathbb{R}^{+}$. La función c define un vector columna de N entradas donde N representa la cardinalidad del espacio de estado I. Consideremos a $\phi : I\to\mathbb{R}^{+}$ el potencial asociado a la función c. Por el Teorema 3.1.3 se cumple que $\phi = P\phi + c$ donde

$$\begin{aligned}\phi_i &= c_i + \sum_{j\in I} p_{ij}\phi_j \\ &= c_i + \sum_{j\in I} p_{ij}(c_j + \sum_{k\in I} p_{jk}\phi_k) \\ &= c_i + \sum_{j\in I} p_{ij}c_j + \sum_{j\in I}\sum_{k\in I} p_{ij}p_{jk}\phi_k \\ &= c_i + \sum_{j\in I} p_{ij}c_j + \sum_{j\in I} p_{ij}^{(2)}\phi_j.\end{aligned}$$

Si procedemos de manera progresiva sustituyendo el valor ϕ_j...

$$\begin{aligned}\phi_i &= c_i + \sum_{j\in I} p_{ij}c_j + \ldots + \sum_{j\in I} p_{ij}^{(m)}c_j + \ldots \\ &= \delta_{ii}c_i + \sum_{j\in I} p_{ij}c_j + \ldots + \sum_{j\in I} p_{ij}^{(m)}c_j + \ldots \\ &= \sum_{j\in I} p_{ij}^{(0)}c_i + \sum_{j\in I} p_{ij}c_j + \ldots + \sum_{j\in I} p_{ij}^{(m)}c_j + \ldots \\ &= \sum_{n=0}^{\infty}(P^n c)_i\end{aligned}$$

donde $(P^n c)_i$ denota el producto matricial del $i-$ésimo renglón de la matriz P^n con el vector columna c. Por lo tanto para toda c no negativa se cumple que

$$\begin{aligned}\phi &= \left(\sum_{n=0} P^n\right)c \\ &= Gc. \end{aligned} \tag{3.1.15}$$

En cambio, para el caso continuo por el Teorema 3.1.4 sabemos que el potencial asociado cumple que

$$\begin{aligned}\phi_i &= \frac{c_i}{q_i} + \sum_{j\neq i}\frac{q_{ij}}{q_i}\phi_j \\ &= \frac{c_i}{q_i} + \sum_{i\neq j}\frac{q_{ij}}{q_i}\left(\frac{c_j}{q_j} + \sum_{j\neq k}\frac{q_{jk}}{q_j}\phi_k\right).\end{aligned}$$

Se recuerda que $\pi_{ij} = q_{ij}/q_i$ para toda i distinta de j. Si desarrollamos de manera similar que en el caso discreto considerando a la cadena de saltos asociada al proceso $(X_t)_{t\geq 0}$ se sigue que

$$
\begin{aligned}
&= \frac{c_i}{q_i} + \sum_{j\in I} \pi_{ij}\frac{c_j}{q_j} + \ldots + \sum_{i\neq j} \pi_{ij}^{(m)}\frac{c_j}{q_j} + \ldots \\
&= \sum_{j\in I}\sum_{n=0}^{\infty} \mathbb{E}_i\big[S_{n+1}|Y_n = j\big]\mathbb{P}_i\big(Y_n = j\big)c_j \\
&= \sum_{j\in I} \mathbb{E}_i\left[\sum_{n=0}^{\infty} S_{n+1}\mathbb{1}_{\{Y_n=j\}}\right]c_j \\
&= \sum_{j\in I} \mathbb{E}_i\left[\sum_{n=0}^{\infty} S_{n+1}\mathbb{1}_{\{Y_n=j\}}\right]c_j \\
&= \sum_{j\in I} \mathbb{E}_i\left[\int_0^{\infty} \mathbb{1}_{\{X_t=j\}}dt\right]c_j \\
&= \sum_{j\in I} \int_0^{\infty} \mathbb{E}_i\Big[\mathbb{1}_{\{X_t=j\}}\Big]dt c_j \\
&= \sum_{j\in I} \int_0^{\infty} p_{ij}(t)dt\, c_j \\
&= \left(\left(\int_0^{\infty} P(t)dt\right)c\right)_i .
\end{aligned}
$$

Por lo tanto para toda c no negativa, tenemos que

$$\phi = \left(\int_0^{\infty} P(t)\,dt\right)c.$$

Para potenciales cuyo costo tiene un factor de descuento la situación es similar pues

$$\phi_i = \mathbb{E}_i\left[\sum_{n=0}^{\infty} \alpha^n c(X_n)\right] = \sum_{n=0}^{\infty} \alpha^n \mathbb{E}_i\big[c(X_n)\big] = (R_\alpha c)_i, \tag{3.1.16}$$

donde

$$R_\alpha = \sum_{n=0}^{\infty} \alpha^n P^n.$$

Por el Teorema 3.1.5 ϕ es la única solución acotada al sistema de ecuaciones lineales

$$\phi = \alpha P\phi + c.$$

Denotemos por I la matriz identidad. Por lo anterior, si despejamos al vector c y factorizamos de acuerdo a las leyes del producto de matrices se tiene que

$$\begin{aligned} c &= \phi - \alpha P\phi \\ &= (\mathrm{I} - \alpha P)\phi. \end{aligned} \tag{3.1.17}$$

Tomando en cuenta la Ecuación (3.1.16), en términos matriciales

$$\phi = R_\alpha c,$$

de manera que si sustituimos en la igualdad (3.1.17) se sigue que

$$\begin{aligned} c &= \phi - \alpha P\phi \\ &= (\mathrm{I} - \alpha P)R_\alpha c, \end{aligned} \tag{3.1.18}$$

mostrando que $(\mathrm{I} - \alpha P)R_\alpha = \mathrm{I}$. Otra forma de validar esta afirmación se debe a que

$$\begin{aligned} (\mathrm{I} - \alpha P)(\sum_{k=0}^{n} \alpha^k P^k) &= \mathrm{I}^n - \alpha^n P^n \\ &= \mathrm{I} - \alpha^n P^n. \end{aligned}$$

Si la serie de matrices $(\sum_{k=0}^{n} \alpha^k P^k)$ converge, entonces

$$\begin{aligned} (\mathrm{I} - \alpha P)(R_\alpha) &= \lim_{n\to\infty} (\mathrm{I} - \alpha P)\left(\sum_{k=0}^{n} \alpha^k P^k\right) \\ &= \lim_{n\to\infty} \mathrm{I} - \alpha^n P^n \\ &= \mathrm{I} \end{aligned} \tag{3.1.19}$$

pues α pertenece a $(0,1)$. Por lo tanto, de las Igualdades (3.1.18) y (3.1.19) se concluye que

$$R_\alpha = (\mathrm{I} - \alpha P)^{-1}.$$

En el caso continuo tenemos

$$\phi_i = \mathbb{E}_i\left[\int_0^\infty e^{-\lambda t}c(X_t)dt\right] = \int_0^\infty e^{-\lambda t}\mathbb{E}_i\big[c(X_t)\big]dt$$
$$= \int_0^\infty e^{-\lambda t}p_{ij}(t)dt\, c_i = (R_\lambda c)_i\,,$$

donde

$$R_\lambda = \int_0^\infty e^{-\lambda t}P(t)dt.$$

Si desarrollamos el lado izquierdo de la igualdad

$$\int_0^\infty e^{-\lambda t}p_{ij}(t)dt = -\frac{1}{\lambda}e^{-\lambda t}p_{ij}(t)\Big|_0^\infty + \frac{1}{\lambda}\sum_{j\in I}q_{ij}\int_0^\infty e^{-\lambda t}p(t)dt$$
$$= \frac{1}{\lambda}\delta_{ij} + \frac{1}{\lambda}\sum_{j\in I}q_{ij}\int_0^\infty e^{-\lambda t}p(t)dt\,,$$

lo cual implica que

$$\int_0^\infty e^{-\lambda t}p_{ij}(t)dt\Big(\lambda\,\delta\,jk - \sum_{j\in I}q_{ij}\Big) = \delta_{ij}$$

i.e., $$\int_0^\infty e^{-\lambda t}P(t)dt\,(\lambda \mathrm{I} - Q) = \mathrm{I}.$$

Por lo tanto,

$$R_\lambda = (\lambda \mathrm{I} - Q)^{-1}.$$

A $(R_\alpha : \alpha \in (0,1))$ y a $(R_\lambda : \lambda \in (0,1))$ se les conocen como el resolvente de la cadena de Markov, según sea el caso discreto o continuo.

Regresemos al caso general en el que contamos con una frontera ∂D. Cualquier función acotada $\phi = (\phi_i : i \in I)$ tal que

$$\phi = P\phi \quad \text{en} \quad D,$$

la llamaremos ***harmónica*** en D. Pasaremos a examinar la relación entre las funciones no negativas armónicas en D y con frontera ∂D.

Supongamos que tenemos una cadena de Markov $(X_n)_{n\geq 0}$ con frontera absorbente ∂D. Sea

$$h_i^\partial = \mathbb{P}_i(T < \infty),$$

donde T es el tiempo de llegada a ∂D. Si $\mathbb{P}_i(T < \infty) = 1$ para toda i sabemos que

$$\begin{aligned} \phi &= P\phi + c \quad \text{en} \quad D \\ \phi &= f \qquad\quad\ \ \text{en} \quad \partial D \end{aligned}$$

tiene a lo más una solución acotada. Por otra parte, si ϕ es una función armónica acotada y no negativa en D con valores frontera $\phi_j = f_j$ para toda j en ∂D tenemos por el Teorema de convergencia monótona que para toda i en D

$$\phi_i = \mathbb{E}_i\big[f(X_T)\big] = \sum_{j \in \partial D} f_j \mathbb{P}_i(X_T = j),$$

con lo cual concluimos que toda función armónica es determinada por sus valores frontera y

$$\phi = \sum_{j \in \partial D} f_j h^j,$$

donde $h^j = (h_i^j : i \in I)$. De esta manera las probabilidades de llegada a los estados frontera forman un conjunto de generadores externos para el conjunto de todas las funciones armónicas no negativas y acotadas.

Capítulo 4

Caminatas Aleatorias y Redes Eléctricas

4.1 Introducción a Redes y Funciones Harmónicas

Se define a una red como una gráfica no dirigida G compuesta de un conjunto de vértices V y otro conjunto de aristas E; para referirnos a los vértices de G de manera alternativa les diremos *nodos*. Definamos las conductancias como una función c definida en las aristas de la red G al conjunto de números no negativos dadas por $\{c(e)\}$. Denotaremos por $c(x,y) = c(\{x,y\})$ donde $\{x,y\}$ representa a la arista que conecta a los vértices x,y y $c(x,y)$ representa la conductancia asignada a la arista $\{x,y\}$ donde c cumple que $c(x,y) = c(y,x)$. Al recíproco $r(e) = 1/c(e)$ lo llamaremos como la resistencia de la arista e. Se denota a una red con la dupla $(G,\{c(e)\})$. Para x,y en V escribiremos $x \sim y$ para indicar que $\{x,y\}$ pertenece a E, es decir, x y y están conectados.

Una cadena de Markov es reversible si es que existe una función positiva $x \to \pi(x)$ definida en el espacio de estados tal que las probabilidades de transición satisfacen

$$\pi(x)\,p(x,y) = \pi(y)\,p(y,x) \tag{4.1.1}$$

para cualesquiera dos estados x,y. A la expresión (4.1.1) se le conoce como la ***ecuación de balance***. Si existe π que satisface las ecuaciones de balance, se dice que π es una medida $P-estacionaria$ pues para toda $y \in V$ tenemos que

$$\begin{aligned}\sum_x \pi(x)p(x,y) &= \pi(y)\left(\sum_x p(y,x)\right)\\ &= \pi(y).\end{aligned}$$

Consideremos una cadena de Markov definida sobre los nodos de G con matriz de transición definida por

$$p(x,y) = \frac{c(x,y)}{c(x)},$$

donde $c(x) = \sum_{y:y\sim x} c(x,y)$. A esta cadena también se le conoce como una caminata aleatoria ponderada con pesos $\{c(e)\}$ o también como la cadena de Markov asociada a la red $(G,\{c(e)\})$. Si definimos $\pi(x) = c(x)/c_G$ con $c_G = \sum_{x\in V} c(x)$, la cadena de Markov asociada a $(G,\{c(e)\})$ es una cadena de Markov reversible con respecto a la medida de probabilidad π pues

$$\pi(x)\,p(x,y) = \frac{c(x)}{c_G}\frac{c(x,y)}{c(x)} = \frac{c(x,y)}{c_G} = \frac{c(y)}{c_G}\frac{c(y,x)}{c(y)} = \pi(y)\,p(y,x),$$

mostrando que π es una medida estacionaria.

Se procede a mostrar que toda cadena de Markov reversible es una caminata aleatoria ponderada en una red. Supongamos que P es una matriz de transición definida en un conjunto finito Ω el cual es reversible con respecto a la medida de probabilidad π. Definamos una gráfica con el conjunto de vértices Ω declarando que $\{x,y\}$ es una arista si tenemos que $p(y,x) > 0$. Esto es una definición propia, pues la reversibilidad implica que $p(x,y) > 0$. Se defina a la conductancia en las aristas como $c(x,y) = \pi(x)p(x,y)$ y observemos que esta relación es simétrica debido a la reversibilidad del proceso. Con esta elección de pesos se sigue que

$$\begin{aligned}c(x) &= \sum_{y:y\sim x} \pi(y)\,p(y,x)\\ &= \pi(x)\end{aligned}$$

de manera que la matriz asociada a esta red es P. Con esto concluimos que el estudio de cadenas de Markov reversibles es entonces equivalente al estudio de caminatas aleatorias en redes.

Si una cadena de Markov empieza en un estado x, ¿Cómo podemos determinar si está vinculado a visitar otro estado dado a, es decir, si la posibilidad de visitar eventualmente a a es 1 o menor que 1?

A pesar de nuestro interés por recurrencia y transitoriedad nos enfocaremos primero en las caminatas aleatorias en redes finitas. Sea G una red finita, x un vértice de G y sean A,Z dos subconjuntos disjuntos de vértices de G. Definamos a τ_A la primera vez que la caminata toca algún vértice de A. En ocasiones denotaremos por τ_A^+ como el primer tiempo posterior al tiempo 0 tal que la caminata toca algún vértice de A. Usualmente A y Z serán conjuntos unitarios, es decir, conjuntos con un sólo elemento.

Consideremos la probabilidad de que la caminata visite al conjunto A antes de visitar a Z como una función del punto inicial x, es decir:

$$F(x) := \mathbb{P}_x(\tau_A < \tau_Z). \tag{4.1.2}$$

Claramente $F\big|_A \equiv 0, F\big|_Z \equiv 1$, y para $x \notin A \cup Z$ por la propiedad de Markov tenemos,

$$\begin{aligned} F(x) &= \sum_y P_x(X_1 = y)P_y(\tau_A < \tau_Z | X_1 = y) \\ &= \sum_{y:y\sim x} p(x,y)F(y) = \frac{1}{\pi(x)} \sum_{y:y\sim x} c(x,y)F(y), \end{aligned}$$

donde $x \sim y$ indica que x,y son adyacentes en G. De la última ecuación podemos decir que el valor de $F(x)$ es un *promedio* de los vértices vecinos de x.

Definición 4.1.1. *Sea f una función $f : V \to \mathbb{R}$. Se dice que f es una función harmónica en x si*

$$h(x) = \frac{1}{\pi(x)} \sum_{y:y\sim x} c(x,y)h(y).$$

Si f es harmónica en cada punto del conjunto V entonces se dice que h es harmónica en V.

Las funciones harmónicas satisfacen el principio del máximo. Antes de proceder a la demostración, denotaremos por $\overline{W}$ como al conjunto de vértices que pertenecen a W o son adyacentes a algún vértice en W y por ∂W como al conjunto de vértices adyacentes a algún vértice en W, de esto último observemos que $\overline{W} = W \cup \partial W$.

Principio del Máximo. *Sea G una red finita o infinita. Si H es una subgráfica conexa de G, $f : V(G) \to \mathbb{R}$ una función harmónica en $V(H)$ de manera que f alcanza su máximo en algún vértice de H, entonces f es constante en $\overline{V(H)}$.*

Demostración:
Definamos a $K = \{y \in V(G) : f(y) = M\}$ con $M = \sup f$ y consideremos a $x \in V(H) \cap K$, como f es harmónica en $V(H)$ entonces para toda $x \sim y$

$$\begin{aligned} M = f(x) &= \frac{1}{\pi(x)} \sum_{x \sim y} c(x,y)h(y) \\ &= \sum_{x \sim y} p(x,y)f(y). \end{aligned} \tag{4.1.3}$$

Por otra parte se sabe que $\sum_{x \sim y} p(x,y) = 1$ pues P es una matriz estocástica, de manera que por la igualdad (4.1.3) se sigue que;

$$\begin{aligned} \sum_{x \sim y} p(x,y)f(y) &= M \sum_{x \sim y} p(x,y) \\ &= \sum_{x \sim y} p(x,y)M, \end{aligned} \tag{4.1.4}$$

para obtener que

$$\begin{aligned} 0 &= \sum_{x \sim y} p(x,y)M - \sum_{x \sim y} p(x,y)f(y) \\ &= \sum_{x \sim y} p(x,y)\big[M - f(y)\big]. \end{aligned}$$

Si $x \sim y$, en particular se tiene que $p(x,y) > 0$ y como $x \in V(H) \cap K$ tenemos que $M - f(y) \geq 0$ de manera que si $\sum_{x \sim y} p(x,y)[M - f(y)] = 0$ implica que $M - f(y) = 0$. Por lo tanto, para toda $x \sim y$ se cumple que $f(y) = M$. Ahora como H es conexo y por hipótesis si f alcanza su máximo en algún vértice de H se tiene que $V(H) \cap K \neq \varnothing$ y se sigue que $K \supseteq V(H)$ y por lo tanto $K \supseteq \overline{V(H)}$ mostrando que f es constante en $\overline{V(H)}$. $\square$

De este principio se sigue el siguiente postulado:

Principio de Unicidad. *Sea $G = (V,E)$ una red conexa finita o infinita. Sea W un subconjunto finito propio de V. Si $f, g : V \to \mathbb{R}$ son dos funciones harmónicas en W que coinciden fuera de W (i.e.,$f|_{(V \setminus W)} = g|_{(V \setminus W)}$), entonces $f = g$.*

Demostración:
Definamos a $h = f - g$ y observemos que por definición $h\big|_{(V \setminus W)} = 0$. Por otra parte notemos que h es una función acotada debido al principio del Máximo. Sin pérdida de generalidad, sea $M = \sup h|_W$. Como W es un subconjunto finito, existe

$x \in W$ tal que $h(x) = M$ y consideremos a la componente conexa H que contiene al vértice x. Por el Principio del Máximo, h es constante en $\overline{V(H)}$, en particular en $\partial V(H) \subseteq \partial W$ donde $h = 0$ de manera que $M = 0$. Siguiendo la misma idea para el caso en que $M = \inf h\big|_W$ concluimos que $h = 0$ en W y por lo tanto $f = g$. $\square$

Y a consecuencia de estos dos principios, obtenemos

Principio de Superposición. *Sea $G = (V,E)$ una red conexa finita o infinita. Sea W un subconjunto finito propio de V. Si $f, f_1, f_2 : V \to \mathbb{R}$ son funciones harmónicas en W con $f = \alpha_1 f_1 + \alpha_2 f_2$ fuera de W donde $\alpha_1, \alpha_2 \in \mathbb{R}$ entonces, $f = \alpha_1 f_1 + \alpha_2 f_2$.*

Retomando la ecuación (4.1.2), dada una red eléctrica $G = (V,E)$ la función $F(x)$ es harmónica en $V \setminus (A \cup Z)$ de manera que para determinar los valores de $F(x)$ basta resolver el sistema de ecuaciones:

$$F\big|_A = 1$$
$$F\big|_Z = 0$$
$$F\big|_{(A \cup Z)} = PF,$$

como consecuencia del Principio de Unicidad. De este último sistema de ecuaciones podemos observar que la solución está determinada por los valores establecidos en A y Z.

Pensemos ahora en una función definida en un subconjunto de vértices de una red dada. El ***Problema de Dirichlet*** pregunta cuándo una función dada puede extenderse a todos los vértices de una red de manera que sea harmónica en donde no ha sido definida previamente. El siguiente principio nos dará una idea de la respuesta:

Principio de Existencia. *Sea $G = (V,E)$ una red finita o infinita. Si $W \subsetneq V$ y $f_0 : V \setminus W \to \mathbb{R}$ es una función acotada, entonces existe $f : V \to \mathbb{R}$ tal que $f\big|_{(V \setminus W)} = f_0$ y f es harmónica en W.*

Demostración:
Consideremos a la caminata asociada a la red eléctrica G y al tiempo de llegada $\tau_{V \setminus W}$. Para $x \in V$, definamos a

$$f(x) := \mathbb{E}_x[f_0(X_{\tau_{V \setminus W}})].$$

Observemos que para toda $x \in V \setminus W$ se tiene que $f(x) = f_0(x)$ pues $\tau_{V\setminus W} = 0$. Ahora, para $x \in W$ tenemos

$$\begin{aligned} f(x) &= \mathbb{E}_x[f_0(X_{\tau_{V\setminus W}})] \\ &= \sum_{y:y\sim x} \mathbb{P}_x(X_1 = y)\mathbb{E}_x[f_0(X_{\tau_{V\setminus W}})|X_1 = y] \\ &= \sum_{y:y\sim x} p(x,y)\mathbb{E}_y[f_0(X_{\tau_{V\setminus W}})] = \sum_{y:y\sim x} p(x,y)f(y), \end{aligned}$$

la última desigualdad debido a la propiedad de Markov. Por lo tanto, f es harmónica en W. $\square$

Observemos que la función F definida en (4.1.2) anteriormente es un caso particular del Principio de Existencia donde $W = V \setminus (A \cup Z)$, $f_0\big|_A = 1$, y $f_0\big|_Z = 0$.

Por otro lado, si G es una red finita, el principio de existencia puede deducirse directamente del principio de unicidad: El problema de Dirichlet en un espacio finito de estados consiste en resolver un sistema finito de ecuaciones lineales, una ecuación por cada estado de V. Como el número de incógnitas es igual al número de ecuaciones, el principio de unicidad implica el principio de existencia. Como mencionamos anteriormente, nuestro objetivo consta en determinar cuándo ciertos estados están obligados a ser visitados, por lo que es de vital importancia asumir la reversibilidad de la cadena de Markov asociada a la red eléctrica G, supuesto tomado a partir de ahora, a menos que se establezca lo contrario.

Matemáticamente, una red eléctrica es una gráfica ponderada. A los pesos asociados a las aristas de una gráfica dada G los llamaremos ***conductancias*** donde el recíproco de estos valores son llamados ***resistencias***. Denotaremos por $r(x,y)$ a $c(x,y)^{-1}$. Consideremos dos subconjuntos A y Z de vértices de una red, el ***voltaje*** es una función definida en los vértices de la red G que es harmónica en el conjunto de vértices $V \setminus (A \cup Z)$. Debido a la harmonicidad del voltaje esta función estará determinada por los valores de v en A y B, usualmente el voltaje será especificado a ser 1 y 0 respectivamente, a estos dos conjuntos los podemos pensar como una ***fuente*** y ***disipador*** respectivamente. Dada una función θ definida para un par ordenado de vértices vecinos, diremos que θ es un ***flujo de corriente*** entre A y Z si $\theta(x,y) = -\theta(y,x)$ para cualesquiera dos vértices vecinos x,y y dado un flujo, definiremos la ***divergencia*** de θ como

$$div\, \theta(x) := \sum_{y:y\sim x} \theta(x,y).$$

Notemos que para todo flujo de corriente θ se tiene que

$$\sum_{x\in V}\sum_{y:y\sim x}\theta(x,y)\sum_{x\in V}\sum_{y:y\sim x}\theta(x,y)=\sum_{<x,y>\in E}[\theta(x,y)+\theta(y,x)]=0. \tag{4.1.5}$$

Dada una función de voltaje v, definiremos a la función asociada ***corriente*** i en las aristas por

$$i(x,y):=c(x,y)[v(x)-v(y)].$$

Observemos que por definición, i es antisimétrica (es decir, $i(x,y)=-i(x,y)$) y, por otra parte, que la corriente fluye en la dirección de los vértices de menor voltaje o dicho de otra forma $i(x,y)>0$ si y solo si $v(x)>v(y)$. Consideremos ahora un vértice x de G y observemos que por la harmonía de $v(\cdot)$ tenemos

$$\begin{aligned}\sum_{y:y\sim x} i(x,y) &= \sum_{y:y\sim x} c(x,y)[v(x)-v(y)]\\ &= \sum_{y:y\sim x} c(x,y)v(x)-\sum_{x\sim y} c(x,y)v(y)\\ &= c(x)v(x)-c(x)\left(\sum_{y:y\sim x} p(x,y)v(y)\right)\\ &= 0.\end{aligned}$$

Entonces, un flujo de A a Z es un flujo θ que satisface la: **Ley de nodos de Kirchhoff:**

$$div\,\theta(x)=0 \quad \text{para toda } x\notin A\cup B.$$

Observemos que por definición $i(x,y)=c(x,y)[v(x)-v(y)]$ para deducir que la corriente cumple la

Ley de Ohm: Si $x\sim y$, la corriente fluyendo de x a y satisface

$$v(x)-v(y)=i(x,y)r(x,y).$$

Finalmente, consideremos la trayectoria $x_1\sim x_2\sim\ldots\sim x_n\sim x_{n+1}=x_1$. Por la Ley de Ohm la suma $\sum_{k=1}^{n+1} i(x_k,x_{k+1})r(x_k,x_{k+1})=v(x_1)-v(x_{n+1})=0$ para deducir la siguiente ley

Ley de ciclos de Kirchhoff: Si $x_1\sim x_2\sim\ldots\sim x_n\sim x_{n+1}=x_1$, es un ciclo entonces

$$\sum_{k=1}^{n+1} i(x_k,x_{k+1})r(x_k,x_{k+1})=0.$$

Proposición 4.1.2. *Supongamos que una función antisimétrica j definida en las aristas de una red finita conectada satistface la Ley de ciclos de Kirchhoff y satisface la Ley de nodos de Kirchhoff de la forma $\sum_{x\sim y} j(x,y) = 0$ para toda x en W. Entonces, j es la corriente asociada a algún voltaje cuyos valores están especificados fuera de W, además esta función voltaje es única salvo una constante aditiva.*

Demostración:
Sea $G = (V,E)$ una red finita conectada y una función antisimétrica $j : E \to \mathbb{R}$ tal que satisface la Ley de ciclos de Kirchhoff y satisface la Ley de nodos de Kirchhoff de la forma $\sum_{x\sim y} j(x,y) = 0$ para toda x en W. Observemos que dada esta suposición, la cantidad $r(x,y)j(x,y)$ está bien definida pues para toda x en W, si $x \sim y$ por la Ley de ciclos en G, tenemos que para toda trayectoria de y a x $x_1 = y \sim x_2 \sim ... \sim x_{n+1} = x$ se cumple

$$j(x,y)r(x,y) = \sum_{k=1}^{n} r(x_k,x_{k+1})j(x_k,x_{k+1}).$$

Impongamos una función $f_0 : V \setminus W \to \mathbb{R}$ y definamos otra función $w : V \to \mathbb{R}$ tal que cumpla que para toda x en W, si $x \sim y$

$$w(x) - w(y) = r(x,y)j(x,y),$$

con $W\big|_{V\setminus W} = f_0$. Por construcción, tenemos que $j(x,y) = c(x,y)[w(x) - w(y)]$. Ahora, por la Ley de nodos de Kirchhoff para toda x en W se cumple la siguiente igualdad

$$w(x)\sum_{x\sim y} c(x,y) = \sum_{x\sim y} c(x,y)w(y),$$

para así obtener que

$$w(x) = \frac{1}{\pi(x)}\sum_{x\sim y} c(x,y)w(y).$$

Por lo tanto dada una función j podemos construir una función harmónica tal que en W, $j(x,y)$ es la corriente asociada al voltaje w, es decir, si $x \sim y$ se cumple que $j(x,y) = c(x,y)[w(x) - w(y)]$ para toda x en W.

Ahora, supongamos que existe v tal que

$$v(x) - v(y) = r(x,y)j(x,y) = w(x) - w(y). \tag{4.1.6}$$

Por el principio de superposición la función $h = v - w$ también es harmónica donde por la igualdad (4.1.6) podemos deducir que $v(x) - w(x) = v(y) - w(y)$, es decir, h es una función constante. Dicho esto, concluimos que w es única salvo una constante.

□

De esta última construcción, obtenemos el siguiente corolario

Corolario 4.1.3. *Dadas la Ley de nodos y la Ley de ciclos de Kirchhoff siempre podemos deducir la Ley de Ohm.*

Definamos la **intensidad** de un flujo θ de a a z como $||\theta|| := div\theta(a)$. Diremos que un flujo es unitario si la intensidad de dicho flujo es uno.

En la siguiente sección analizaremos parte de estas ideas y daremos una interpretación bajo un punto de vista probabilístico.

4.2 Interpretaciones Probabilísticas

Dada una red, la razón $[w(a) - w(z)]/||I||$, donde I es el flujo de la corriente correspondiente a un voltaje w, es independiente del voltaje aplicado a la red. Definamos la **resistencia efectiva** entre 2 vértices a y z como

$$\mathscr{R}(a \leftrightarrow z) := \frac{w(a) - w(z)}{||I||}. \tag{4.2.1}$$

De manera paralela, definamos la **conductancia efectiva** como $\mathscr{C}(a \leftrightarrow z) = 1/\mathscr{R}(a \leftrightarrow z)$. Imaginemos una red nueva compuesta únicamente por una arista que conecta los vértices a y z con resistencia $\mathscr{R}(a \leftrightarrow z)$, la idea detrás de esta cantidad es que al aplicar el mismo voltaje a a y z en ambas redes, la cantidad de corriente fluyendo de a a z debe ser la misma que en la red inicial. Por otro lado, como $||I||$ es la cantidad total de corriente que fluye en el circuito en a, podemos considerar el circuito entero entre a y z como un solo conductor con la conductancia efectiva $\mathscr{C}(a \leftrightarrow z)$.

A continuación discutiremos la conexión entre resistencia efectiva y la probabilidad de escape $\mathbb{P}_a(\tau_z < \tau_a)$, es decir, la probabilidad de que una caminata aleatoria empezada en a toque a z antes de regresar a a, en ocasiones denotaremos por $\mathbb{P}(a \to z)$ a esta probabilidad.

Proposición 4.2.1. *Para cualesquiera x, a en Ω y $z \in Z \subseteq V$ se tiene*

$$\mathbb{P}_a(\tau_z < \tau_a^+) = \frac{1}{\pi(a)\mathscr{R}(a \leftrightarrow z)} = \frac{\mathscr{C}(a \leftrightarrow z)}{\pi(a)}$$

Demostración:
Consideremos la función

$$x \to \mathbb{E}_x \mathbb{1}_{X_{\tau_{a,z}=z}} = \mathbb{P}_x(\tau_z < \tau_a),$$

de manera que por el principio de existencia, es la única función harmónica en $\Omega \setminus a, z$ con valores 0 en a y 1 en z. Por otra parte, dado un voltaje cualquiera v, la función

$$x \to \frac{w(a) - w(x)}{w(a) - w(z)},$$

también es harmónica en $\Omega \setminus a, z$ con los mismos valores frontera que $\mathbb{P}_x(\tau_z < \tau_a)$, a consecuencia del principio de unicidad tenemos que

$$\mathbb{P}_x(\tau_z < \tau_a) = \frac{w(a) - w(x)}{w(a) - w(z)}.$$

Por lo tanto,

$$\begin{aligned}\mathbb{P}_a(\tau_z < \tau_a^+) &= \sum_{x:x\sim a} p(a,x)\mathbb{P}_x(\tau_z < \tau_a) \\ &= \sum_{x:x\sim a} \frac{c(a,x)}{\pi(a)} \frac{w(a) - w(x)}{w(a) - w(z)},\end{aligned}$$

donde por definición de intensidad de corriente, concluimos que

$$\mathbb{P}_a(\tau_z < \tau_a^+) = \frac{1}{\pi(a)\mathscr{R}(a \leftrightarrow z)} = \frac{\mathscr{C}(a \leftrightarrow z)}{\pi(a)}.$$

$\square$

Definición 4.2.2. *La función de Green de una variable aleatoria parada en un tiempo de paro τ se define como*

$$\mathscr{G}_\tau(x, y) = \mathbb{E}_x\left[\sum_{k=0}^{\infty} \mathbb{1}_{\{X_k = y, \tau > k\}}\right]. \tag{4.2.2}$$

Lema 4.2.3. *Sea $G=(V,E)$ una red finita y conexa y consideremos la caminata aleatoria asociada a la red. Entonces la función de Green satisface*

$$\mathscr{G}_{\tau_x}(a,a)=\pi(a)\mathscr{R}(a\leftrightarrow Z).$$

Demostración:
Se recuerda al lector que τ_Z es el tiempo de llegada a Z. Definamos a la variable aleatoria $N_a^{(Z)}=\sum_{k=0}^{\infty}\mathbb{1}_{\{X_k=y,\tau_Z>k\}}$ y calculemos la siguiente probabilidad

$$\begin{aligned}
&\mathbb{P}_a(N_a^{(Z)}\leq k)\\
&\quad=\sum_{m=1}^{\infty}\mathbb{P}_a(N_a^{(Z)}\leq k-1,\, x_m=a,x_{m-1}\neq a,\ldots,x_1\neq a,\tau_Z>k)\\
&\quad=\sum_{m=1}^{\infty}\mathbb{P}_a(N_a^{(Z)}\leq k-1,|\, x_m=a,x_{m-1}\neq a,\ldots,x_1\neq a,\tau_Z>k)\cdot\\
&\qquad\qquad\qquad\mathbb{P}_a(x_m=a,x_{m-1}\neq a,\ldots,x_1\neq a,\tau_Z>k)\\
&\quad=\mathbb{P}_a(N_a^{(Z)}\leq k-1)\sum_{m=1}^{\infty}P_a(x_m=a,x_{m-1}\neq a,\ldots,x_1\neq a,\tau_Z>k)\\
&\quad=\mathbb{P}_a(N_a^{(Z)}\leq k-1)\,\mathbb{P}_a(\tau_a<\tau_z^+)\\
&\qquad\vdots\\
&\quad=\mathbb{P}_a(N_a^{(Z)}\leq 0)\left(\mathbb{P}_a\big(\tau_a<\tau_z^+\big)\right)^k,
\end{aligned}$$

pero

$$\mathbb{P}_a(N_a^{(Z)}\leq 0)=\mathbb{P}_a\big(\tau_z<\tau_a^+\big)+P_a(\tau_a=\infty)=\mathbb{P}_a\big(\tau_z<\tau_a^+\big),$$

pues la cadena de Markov asociada a la red finita G es reversible. Por lo tanto, el número de visitas al estado a estrictamente antes de llegar a Z es una variable aleatoria geométrica con esperanza $\mathbb{P}_a\big(\tau_z<\tau_a^+\big)^{-1}$. El lema se sigue de la Proposición 4.2.1. □

Por definición de corriente, al escalar un voltaje multiplicándolo por una constante, como resultado obtenemos que la corriente será escalada por el mismo factor. En ocasiones convendrá escalarlo de manera que la corriente sea una unidad de flujo.

Proposición 4.2.4. *Sea G una red finita y conectada. Impongamos un voltaje en $\{a\} \cup Z$ de manera que fluya la corriente unitaria de a a Z con 0 en Z, entonces, el voltaje satisface para toda x en V*

$$v(x) = \frac{\mathscr{G}_Z(a,x)}{\pi(x)}$$

Demostración:
Ya se demostró que la proposición es verdadera para x en $\{a\} \cup Z$, basta demostrar que la función $\mathscr{G}_Z(a,x)/\pi(x)$ es harmónica fuera. Por el teorema de Fubini, la función de Green se puede representar de la siguiente manera:

$$\mathbb{E}_x \left[\sum_{k=0}^{\infty} \mathbb{1}_{\{X_k = y, \tau > k\}} \right] = \sum_{k=1}^{\infty} \mathbb{P}_x(X_k = y, \tau_Z > k).$$

En gráficas finitas conectadas, $\mathbb{E}_x \, \tau_Z < \infty$ para toda x (*ver* [Nor97, p. 37]) donde $\mathscr{G}_Z(x,y) \leq \mathbb{E}_x \, \tau_Z$, por otra parte, si una cadena de Markov es reversible, entonces para toda $x_1, x_2, \ldots, x_n$ tenemos la siguiente igualdad

$$\pi(x_1) \prod_{i=1}^{n-1} p(x_i, x_{i+1}) = \pi(x_n) \prod_{i=1}^{n-1} p(x_{n+1-i}, x_{n-i}), \tag{4.2.3}$$

y por consiguiente $\prod_{i=1}^{n-1} p(x_i, x_{i+1}) = \prod_{i=1}^{n-1} p(x_{n+1-i}, x_{n-i})$ siempre que $x_1 = x_n$ pues con $n = 1$, por (4.1.1)

$$\pi(x_1) p(x_1, x_2) = \pi(x_2) p(x_2, x_1).$$

De forma inductiva, tenemos que para toda k la Ecuación (4.2.3) se cumple. Consideremos ahora el caso $k+1$, por hipótesis de inducción y la Ecuación (4.1.1) tenemos

$$\begin{aligned}
\pi(x_1) \prod_{i=1}^{k} p(x_i, x_{i+1}) &= \big(\pi(x_1) \prod_{i=1}^{k-1} p(x_i, x_{i+1})\big) \big(p(x_k, p_{k+1})\big) \\
&= \Big(\pi(x_k) \prod_{i=1}^{k-1} p(x_{n+1-i}, x_{n-i})\Big) \Big(p(x_k, p_{k+1})\Big) \\
&= \Big(\pi(x_{k+1}) p(x_{k+1}, p_k)\Big) \left(\prod_{i=1}^{k-1} p(x_{n+1-i}, x_{n-i}) \right) \\
&= \pi(x_{k+1}) \prod_{i=1}^{k} p(x_{n+1-i}, x_{n-i}),
\end{aligned}$$

demostrando la Igualdad (4.2.3). En consecuencia, para toda $k \geq 0$ se cumple

$$\begin{aligned}
&\pi(x)\mathbb{P}_x(X_k = y, \tau_Z > k) \\
&\quad = \sum_{x_1,\ldots,x_{k-1}\in(A\setminus Z)} \pi(x)\mathbb{P}_x(X_1 = x_1, \ldots, X_{k-1} = x_{k-1}, X_k = y, \tau_Z > k) \\
&\quad = \sum_{x_1,\ldots,x_{k-1}\in(A\setminus Z)} \pi(y)\mathbb{P}_y(X_1 = x_{k-1}, \ldots, X_{k-1} = x_1, X_k = x, \tau_Z > k) \\
&\quad = \pi(y)\mathbb{P}_y(X_k = x, \tau_Z > k)
\end{aligned}$$

Por consiguiente,

$$\frac{\mathscr{G}_Z(a,x)}{\pi(x)} = \frac{\mathscr{G}_Z(x,a)}{\pi(a)}, \tag{4.2.4}$$

donde la harmonicidad de la función $\mathscr{G}_Z(a,x)$ se deduce de la misma forma que se hizo en el Principio de Existencia. Por definición de resistencia efectiva se tiene que $\mathscr{R}(a \leftrightarrow Z) = w(a)$, así en conjunto con el lema anterior y la Igualdad (4.2.4) concluimos

$$v(x) = \frac{\mathscr{G}_Z(a,x)}{\pi(x)}. \tag{4.2.5}$$

□

Hasta este momento, hemos mostrado dos interpretaciones del voltaje, sin embargo es natural preguntarse si existe alguna interpretación probabilística de la corriente, la respuesta es sí, para esto consideraremos a la variable aleatoria dada por dos vertices $x \sim y$ al contar el número de transiciones de un vértice a otro estrictamente antes de visitar a Z, denotada por $S^{(Z)}_{x,y}$.

Proposición 4.2.5. *Sea G una red finita conexa y la cadena de Markov asociada a la red parada en Z, entonces para toda $x \sim y$ ocurre que*

$$\begin{aligned}
\mathbb{E}_a[S^{(Z)}_{x,y}] &= \mathscr{G}_Z(a,x)p(x,y) \qquad y \\
\mathbb{E}_a[S^{(Z)}_{x,y} - S^{(Z)}_{y,x}] &= i(x,y),
\end{aligned}$$

donde i es la corriente en G cuando un potencial es aplicado entre a y Z de manera que la corriente unitaria fluya de a a Z.

Demostración:
Notemos que para $k = 1,2,\ldots$ el evento $k < \tau_Z$ depende únicamente de $X_0,\ldots,X_k$ de manera que por la propiedad de Markov en n se tiene

$$\mathbb{P}(X_k = x, X_{k+1} = y, \tau_Z > k) = \mathbb{P}(X_k = x, \tau_Z > k)\, p(x,y).$$

Dicho esto, calculemos el número esperado de transiciones de x a y antes de llegar a Z:

$$\begin{aligned}
\mathbb{E}_a[S_{x,y}^{(Z)}] &= \mathbb{E}_a\left[\sum_{k=0}^{\infty} \mathbb{1}_{\{X_k=x, X_{k+1}=y, \tau_Z>k\}}\right] \\
&= \sum_{k=0}^{\infty} \mathbb{P}_a(X_k = x, X_{k+1} = y, \tau_Z > k) \\
&= \sum_{k=0}^{\infty} \mathbb{P}_a(X_k = x, \tau_Z > k) p(x,y) \\
&= p(x,y)\mathbb{E}_a\left[\sum_{k=0}^{\infty} \mathbb{1}_{\{X_k=x, \tau_Z\}}\right] = \mathscr{G}_Z(a,x)p(x,y).
\end{aligned}$$

Por lo tanto, por la Igualdad (4.2.5) para toda x, y tenemos:

$$\begin{aligned}
\mathbb{E}_a[S_{x,y}^{(Z)} - S_{y,x}^{(Z)}] &= \mathscr{G}_Z(a,x)p(x,y) - \mathscr{G}_Z(a,y)p(y,x) \\
&= v(x)\pi(x)p(x,y) - v(y)\pi(y)p(y,x) \\
&= c(x,y)[v(x) - v(y)] \\
&= i(x,y).
\end{aligned}$$

□

De esta sección, podemos concluir que la conductancia efectiva es una cantidad clave debido a las relaciones de voltaje y probabilidades de escape cuyos valores dependen de la recurrencia o transitoriedad de la cadena. Las proposiciones anteriores nos serán útiles para analizar el caso en que G es una red infinita. Para una red infinita G, supongamos que para toda x en V se cumple que

$$\sum_{y:y\sim x} c(x,y) < \infty,$$

de manera que la caminata aleatoria asociada a la red G esté bien definida. Esto siempre es cierto cuando G es **localmente finita**, es decir, cuando el número de aristas incidentes a todo vértice dado en G es finito.

Dada una red infinita, diremos que $< G_n >$ es una sucesión de subgráficas finitas de G que **exhausta** a G si $G_{n+1} \supseteq G_n$ y $G = \bigcup G_n$. Cada arista en G_n es una arista en G, así que de manera simple les damos las mismas conductancias que en G. Asumamos que G_n es la gráfica inducida en G por $V(G_n)$. Sea Z_n el conjunto de vértices en $G \setminus G_n$. Definamos $W_n^{(G)}$ la gráfica obtenida de identificar Z_n con un sólo vértice z_n y de remover las aristas autodirigidas que se pueden llegar a generar. Observemos que un vértice en G_n puede comunicarse con más de un vértice de Z_n, para este caso definiremos a la conductancia $c(x,z_n) = \sum_{z\in Z_n} c(x,z)$ de manera que las conductancias se pueden ver afectadas en la red $W_n^{(G)}$. Ahora, dada una caminata aleatoria sobre la red G, si la paramos la primera vez que llega a Z_n, esta induce una caminata aleatoria en la red $W_n^{(G)}$ hasta que llegue a z_n. Sea $a \in W_n^{(G)}$ e impongamos un voltaje v en $\{a\} \cup \{z_n\}$ de manera que fluya la corriente unitaria de a a z_n con 0 en z_n, como $W_n^{(G)}$ es finita, dicho voltaje existe y está bien definido en $W_n^{(G)}$. Consideremos lo siguiente, por la Proposición 4.2.4 sabemos que

$$P[a \to z_n] = \frac{\mathscr{C}(a \leftrightarrow z_n)}{\pi(a)}. \tag{4.2.6}$$

Por otra parte, del Lema 4.2.3 despejando obtenemos lo siguiente

$$\mathscr{C}(a \leftrightarrow z_n) = \frac{\pi(a)}{\mathscr{G}_{\tau_{z_n}}(a,a)},$$

de manera que si sustituimos a $\mathscr{C}(a \leftrightarrow z_n)$ en (4.2.6), tenemos la siguiente igualdad

$$\mathbb{P}[a \to z_n] = \frac{1}{\mathscr{G}_{\tau_{z_n}}(a,a)}.$$

Como G es una gráfica conectada, para toda n, se tiene que $\mathscr{G}_{\tau_{z_n}} > 0$. Por otra parte no es difícil ver que $\lim_{n\to\infty} \tau_{z_n} = \infty$ pues la sucesión de gráficas definida anteriormente exhausta a G. Dicho esto, entonces para $\mathscr{G}_{\tau_{z_n}}(a,a)$ tenemos el siguiente límite:

$$\begin{aligned}
\lim_{n\to\infty} \mathscr{G}_{\tau_{z_n}}(a,a) &= \lim_{n\to\infty} \mathbb{E}_a\left[\sum_{k=0}^{\infty} \mathbb{1}_{\{X_k=a,\tau_{z_n}>k\}}\right] \\
&= \mathbb{E}_a\left[\sum_{k=0}^{\infty} \mathbb{1}_{\{X_k=a\}}\right],
\end{aligned}$$

de lo cual se sigue que:

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}[a \to z_n] &= \lim_{n\to\infty} \frac{1}{\mathscr{G}_{\tau_{z_n}}(a,a)} \\
&= \frac{1}{\mathbb{E}_a\left[\sum_{k=0}^{\infty} \mathbb{1}_{\{X_k=a\}}\right]}.
\end{aligned}$$

En este contexto, denotaremos por $\mathbb{P}[a \to \infty] = \lim_{n\to\infty} \mathbb{P}[a \to z_n]$, a este límite lo llamaremos como la probabilidad de escape de a en G, o dicho de otro modo, la probabilidad de nunca llegar a a. Por lo tanto, $\mathbb{P}[a \to \infty] > 0$ para toda a en G si y solo si $\mathbb{E}_a\left[\sum_{k=0}^{\infty} \mathbb{1}_{\{X_k=a\}}\right] < \infty$, es decir, si y solo si la cadena de Markov asociada a la red conectada infinita G es transitoria. Por lo tanto, se demuestra el siguiente Teorema:

Teorema 4.2.6. *(Transitoriedad y Conductancia efectiva) Una caminata aleatoria en una red infinita conectada es transitoria si y solo si la conductancia efectiva de uno de sus vértices a infinito es positiva.*

Capítulo 5

Simulaciones

Este capítulo muestra un ejemplo del Problema de Dirichlet resuelto con un método Monte Carlo (i.e. simulación estocástica) que proporciona una aproximación a la solución del problema (la cual sabemos es una función harmónica) dados una red finita $(G, c\{e\})$ y un vector función $f_0 : \partial W \to \mathbb{R}$ mediante un algoritmo numérico implementado en el lenguaje de programación R con la intención de corroborar la teoría mostrada e evidenciar el comportamiento harmónico de las cadenas de Markov reversibles como parte de la interpretación probabilística de la Teoría del Potencial. Para esto, es necesario establecer un criterio que nos permita evaluar la calidad de nuestra simulación con respecto a la extensión harmónica deseada; la manera natural de evaluar la calidad de nuestra aproximación será calculando la divergencia del voltaje asociado a la función harmónica el cual esperamos sea 0.

Esta aproximación se basa en un método de simulación estocástica que de manera general implementa un algoritmo para simular una caminata aleatoria para realizar una primera estimación de la función harmónica a calcular en la red G para así después mejorar este resultado.

A su vez, este capítulo se divide en 2 partes; la primera describe y justifica los algoritmos o cálculos ocupados y en la segunda implementaremos los algoritmos generados para resolver un ejemplo.

5.1 Introducción

Por un lado, del capítulo anterior sabemos que toda cadena de Markov Reversible es una caminata aleatoria pondera y que la extensión harmónica que resuleve al Problema de Dirichlet en el caso finito existe siempre y cuando estén definidos los siguientes elementos:

- La red $(G,\{e\})$ finita con conductancias $c(x,y)>0$.
- La función $f_0:\partial W\to\mathbb{R}$ acotada y el subconjunto $\partial W\subsetneq V$.

Dados estos elementos, podemos definir el vector de cargas c dado por la relación $c(x)=\sum_{y:x\sim y}c(x,y)$ y la distribución estacionaria $\pi(x)=\frac{c(x)}{c_G}$ con $c_G=\sum_{x\in V}c(x)$. Con estos datos podemos construir la matriz de transición dada por:

$$p(x,y)=\frac{c(x,y)}{c(x)}.$$

Por otro lado, si definimos $c(x,y)=1$ para toda $x\sim y$, la caminata aleatoria simple en G con probabilidades de transición dada por

$$p(x,y)=\begin{cases}\frac{1}{c(x)} & si \quad x\sim y,\\ 0 & e.o.c,\end{cases} \tag{5.1.1}$$

es un caso especial de una camianta aleatoria ponderada en la que en particular, con estas consideraciones, el grado del vértice x o número de adyacencias del vértice x resulta ser igual a $c(x)$.

Definido lo anterior para resolver el problema, por el Principio de Existencia se tiene que la extensión harmónica que resuelve el problema está dada por

$$w(x)=\mathbb{E}_x[f_0(X_{(\tau_{\delta W})})],$$

donde $X_{(\tau_{\delta W})}$ denota el valor de la caminata en el tiempo $\tau_{\delta W}$. Con esto en mente, para calcular (o estimar) $w(x)$ para toda x en W debemos ser capaces de simular una caminata aleatoria iniciada en el vértice x, pararla al momento de llegar al conjunto δW, evaluar $f_0(X_{(\tau_{\delta W})})$ y repetir hasta obtener M valores de la variable aleatoria $f_0(X_{(\tau_{\delta W})})$ con la intención de promediar los distintos valores obtenidos. Mencionado lo anterior, para iniciar con los algoritmos se asume que podemos generar una variable aleatoria $U\sim unif(0,1)$.

En general para simular cada salto dada una matriz de transición $P=(p_{i,j})=\mathbb{P}_x(X_1=y)$ para alguna x,y en V e i,j en $I=1,2,\ldots,S$ y S el número de elementos en V; en nuestro algoritmo consideramos el método de la transofrmación inversa.

Definamos

$$Y_i = \begin{cases} x_0 & \text{si} \quad U < p_{i,1}, \\ x_1 & \text{si} \quad p_{i,1} \leq U < p_{i,1} + p_{i,2}, \\ \vdots & \\ x_j & \text{si} \quad \sum_{k=0}^{j-1} p_{i,k} \leq U < \sum_{k=0}^{j} p_{i,k}, \\ \quad \vdots & \end{cases}$$

una *v.a.* genérica distribuida conforme al *i-ésimo* renglón de P. Dada $U \sim unif(0,1)$ se cumple que $\mathbb{P}(a \leq U < b) = b - a$ y por lo tanto se obtiene que

$$\begin{aligned} \mathbb{P}(Y_i = j) &= \mathbb{P}\left(\sum_{k=0}^{j-1} p_{i,k} < U < \sum_{k=0}^{j}\right) \\ &= p_{i,j} \\ &= \mathbb{P}_x(X_1 = y), \end{aligned}$$

para alguna x, y en V. De manera que, se deduce $X_n \sim MarkovTd(\delta_{X_{n-1}}, P)$.

Algoritmo 5.1.1. Simulación de una caminata aleatoria $MarkovTd(\delta_x, P)$ parada en $\tau_{\delta W}$ con $W = \{z_1, ..., z_n\}$.

1. Esccoger un estado inicial x de V, $X_0 = x$. Definir $n = 1$.
2. Sea N suficientemente grande, definir $\tau = \tau_{\delta W} \wedge N$.
3. Generar $Y_{i_1} \sim Markovdis(\delta_{X_0}, P)$, definir $X_1 = Y_{i_1}$.
4. Si $n < N$ y $X_n \notin \partial W$, generar $Y_{i_{n+1}}$ y definir $n = n+1$ y definir $X_{n+1} = Y_{i_{n+1}}$, cualquier otro caso parar.
5. Regresar al paso 3.

Con respecto al último algoritmo, para simular la caminata asociada a la red G basta establecer N suficientemente grande para asegurarnos que $\tau_{\delta W}$ se alcance. Si asumimos G es conexa y finita, se tiene que P es irreducible y dado que π es una distribución invariante de ésta, implica que todo vértice es recurrente positivo para así obtener que $\mathbb{E}_x \tau_{\partial W} < \mathbb{E}_x \tau_z < \infty$ para alguna $z \in \partial W$. En particular se cumple que $\mathbb{P}(\tau_{\partial W} < \infty) = \mathbb{P}(\tau_z < \infty) = 1$ y a consecuencia de esto, sin importar la elección de N con este algoritmo podemos estimar la media de la variable aleatoria $\tau_{\partial W}$ la cuál nos puede dar una idea del parámetro N a elegir. A priori, la elección de N dependerá del tamaño de la red y se puede deducir ocupando el número máximo obtenido de una muestra de trayectorias simuladas con el algoritmo anterior.

Algoritmo 5.1.2. Estimación puntal de una función harmónica por Método de Monte Carlo.

1. Definir x en V y M en $\mathbb{N}$.

2. Simular M caminatas aleatorias paradas en $\tau_{\delta W}$.

3. Definir
$$\hat{w}(x) = \frac{1}{M}\sum_{k=1}^{M} f_0(X_{(\tau_{\delta W})_k}),$$
donde $X_{(\tau_{\delta W})_k}$ denota el valor alcanzado en la *k-ésima* caminata aleatoria generada.

En resumen, por la Ley Débil de los Grandes Números, la estimación $\hat{w}(x)$ converge en probabilidad a la función harmónica $w(x)$ ya que cada trayectoria generada es independiente de las demás y que la variable aleatoria $\tau_{\delta W}$ es finita con esperanza y varianza acotada. Para evaluar la eficacia de este estimador es necesario calcular $div_{\hat{j}}(x)$, para toda x en el conjunto $W = V \setminus \delta W$ con

$$\hat{j}_k(x,y) = \sum_{y:y\sim x} c(x,y)[\hat{w}_k(x) - \hat{w}_k(y)].$$

En la Sección 4.2 podemos observar que la divergencia en W se concentra alrededor de 0, sin embargo, buscamos w tal que $div_{j_w}(x) = 0$ en W. No obstante, cabe preguntarse si es posible mejorar este resultado.

Definamos

$$\begin{aligned}\psi^{(n)}(x) &= \mathbb{E}_x[\hat{w}(X_n)\mathbb{1}_{\tau_{\partial W}\leq n}]\\ &= \mathbb{E}_x[\hat{w}(X_n)\mathbb{1}_{\tau_{\partial W}<n}] + \mathbb{E}_x[\hat{w}(X_n)\mathbb{1}_{n=\tau_{\partial W}}].\end{aligned}$$

Por un lado, por construcción $w(x) = \mathbb{E}_x[w(X_{\tau_{\partial W}})] = \mathbb{E}_x[\hat{w}(X_{\tau_{\partial W}})]$ y $|\hat{w}(x)| < C < \infty$, para así obtener que

$$\begin{aligned}|\psi^{(n)}(x)| &= \left|\mathbb{E}_x[\hat{w}(X_n)\mathbb{1}_{\tau_{\partial W}<n}] + \mathbb{E}_x[\hat{w}(X_{\tau_{\partial W}})]\right|\\ &\leq C[1 - \mathbb{P}_x(\tau_{\partial W} < n)] + |w(x)|.\end{aligned}$$

Como P es reversible y G es finita, $1 = \mathbb{P}_x(\tau_{\partial W} < \infty) = \lim_{n\to\infty} \mathbb{P}_x(\tau_{\partial W} < n)$. Por lo tanto, por los Teoremas Existencia y Unicidad, se concluye que

$$\lim_{n\to\infty} \psi^{(n)}(x) = w(x). \tag{5.1.2}$$

Por otro lado, a consecuencia de la propiedad de Markov

$$\begin{aligned}\psi^{(n)}(x) &= \mathbb{E}_x[\hat{w}(X_n)\mathbb{1}_{\tau_{\partial W}\leq n}] \\ &= \sum_{y:y\sim x} p(x,y)\mathbb{E}_x[\hat{w}(X_{n-1})\mathbb{1}_{\tau_{\partial W}\leq n-1}|X_1 = y] \\ &= \sum_{y:y\sim x} p(x,y)\mathbb{E}_y[\hat{w}(X_{n-1})\mathbb{1}_{\tau_{\partial W}\leq n-1}],\end{aligned}$$

para así obtener que

$$\psi^{(n)} = P\psi^{(n-1)} = P^{(n-1)}\hat{w}. \tag{5.1.3}$$

Mostrado lo anterior, para mejorar nuestra estimación, por las igualdades (5.1.2) y (5.1.3) basta establecer N suficientemente grande de manera que $\psi^{(N)}(x)$ será la aproximación deseada a la función harmónica $w(x)$.

5.2 Ejemplo

Primero, definamos la la red finita $G = (V, c\{e\})$ inducida por $\mathbb{Z}_{40}^2$; sea $V = \mathbb{Z}_{40}^2$, cada vértice en este conjunto es representado por $v = (m,n)$ con m,n en $\mathbb{Z}_{40}$. Diremos que dos vértices x,y son adyacentes si la distancia euclideana $d(x,y) = 1$. Si $x \sim y$, definamos la conductancia dada por $c(x,y) = 1 = d(x,y)$. Dada esta información, podemos calcular el vector de cargas $c : V \to \mathbb{R}$, la distribución estacionaria π y la matriz de transición como se mostro en la sección anterior.

Ahora, definamos $a = \{(13,13)\}$ y $z = \{(20,30)\}$, ambos puntos en nuestro conjunto de vértices. En la sección 3.2, mostramos que la función dada por la probabilidad de escape:

$$w(x) = \mathbb{P}_x(\,\tau_z < \tau_a), \tag{5.2.1}$$

es harmónica en $W = V \setminus \partial W$ con $\partial W = (a \cup z)$. En particular, tenemos que $w(a) = 0$ y $w(z) = 1$ y por otro lado por el Principio de Existencia se cumple que

$$\begin{aligned}w(x) = \mathbb{E}_x[\,f_0(X_{(\tau_{\delta W})})] &= \mathbb{E}_x[\mathbb{1}_{X_{(\tau_{\delta W})}=z)}] \\ &= \mathbb{P}_x(\,\tau_z < \tau_a),\end{aligned}$$

con $f_0 : \partial W \to \mathbb{R}$ definida como $f_0 = \mathbb{1}_{X_{(\tau_{\delta W})}=z}$.

En el Apéndice B, el Código(R) 1 muestra las líneas de comando ocupadas en el software de programación R para realizar una aproximación a la función $w(x)$ definida en la Ecuación (5.2.1). Los resultados de esta aproximación se muestran de manera gráfica en las siguientes figuras:

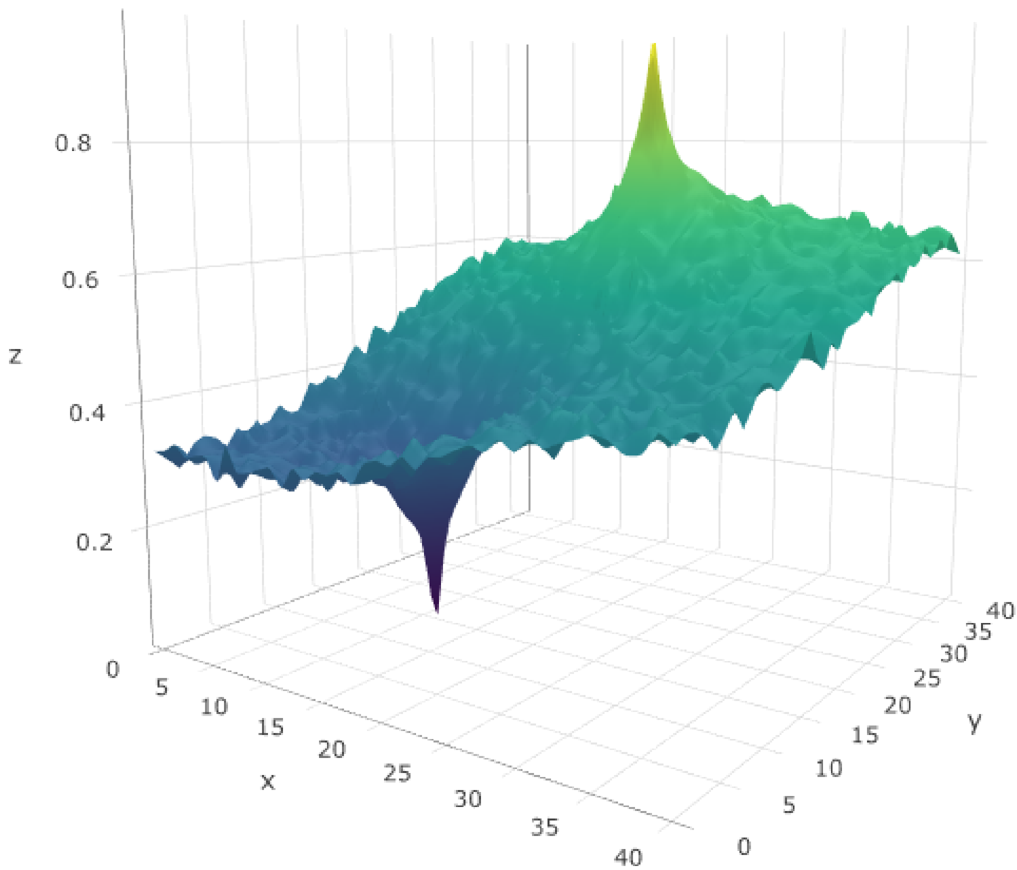

Figura 5.1: Simulación del problema de Dirchlet $\hat{w}(x)$ con dos valores frontera.

Esta figura muestra la gráfica $(x, \hat{w}(x))$ de la aproximación denotada por $\hat{w}$. Esta función se espera que que resuelva el problema, es decir, que sea harmónica en el conjunto de vértices W. En este caso, los valores frontera que determinan a la función son 0 y 1.

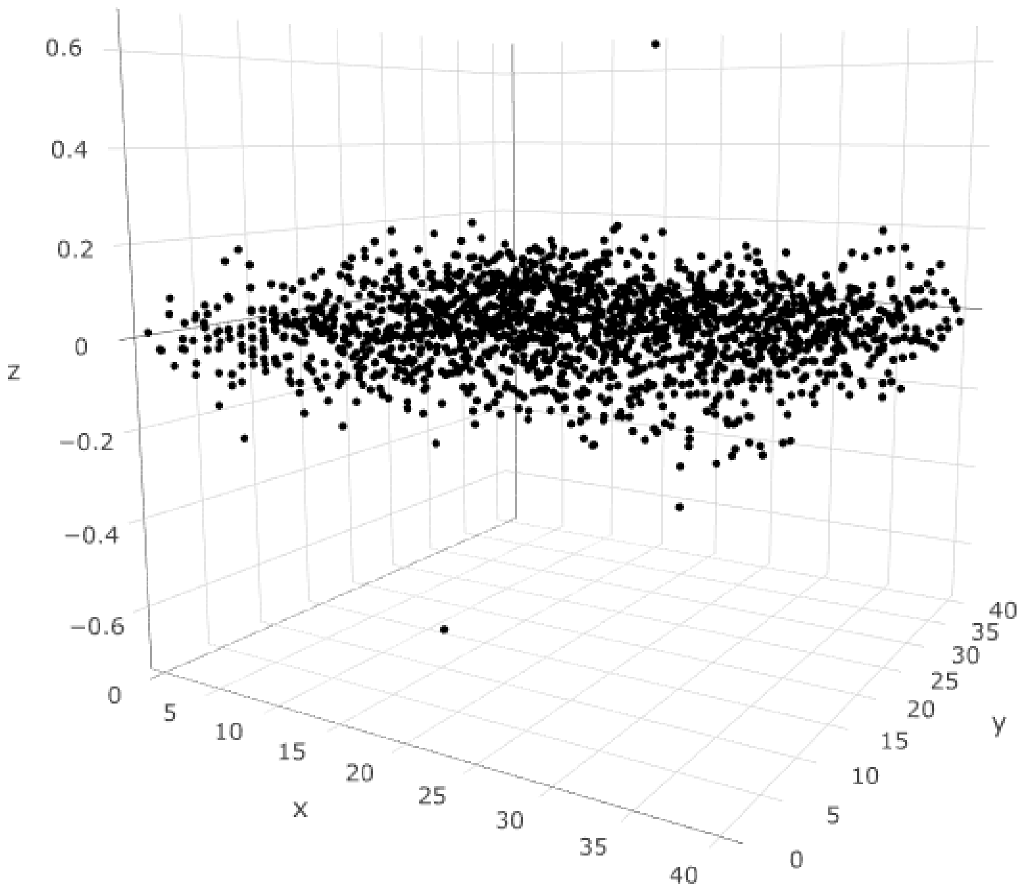

Figura 5.2: Divergencia asociada a $\hat{w}(x)$.

Definamos

$$div\,\hat{\theta}\,(x) = \sum_{y:y\sim x} c(x,y)[\hat{w}(x) - \hat{w}(y)].$$

Si $\hat{w}$ es harmónica en W, es equivalente a que $div\,\hat{\theta}\,(x) = 0$ en W. Como podemos observar, la aproximación $\hat{w}(x)$, no es harmónica en W, sin embargo, esta cantidad se concentra alrededor del 0.

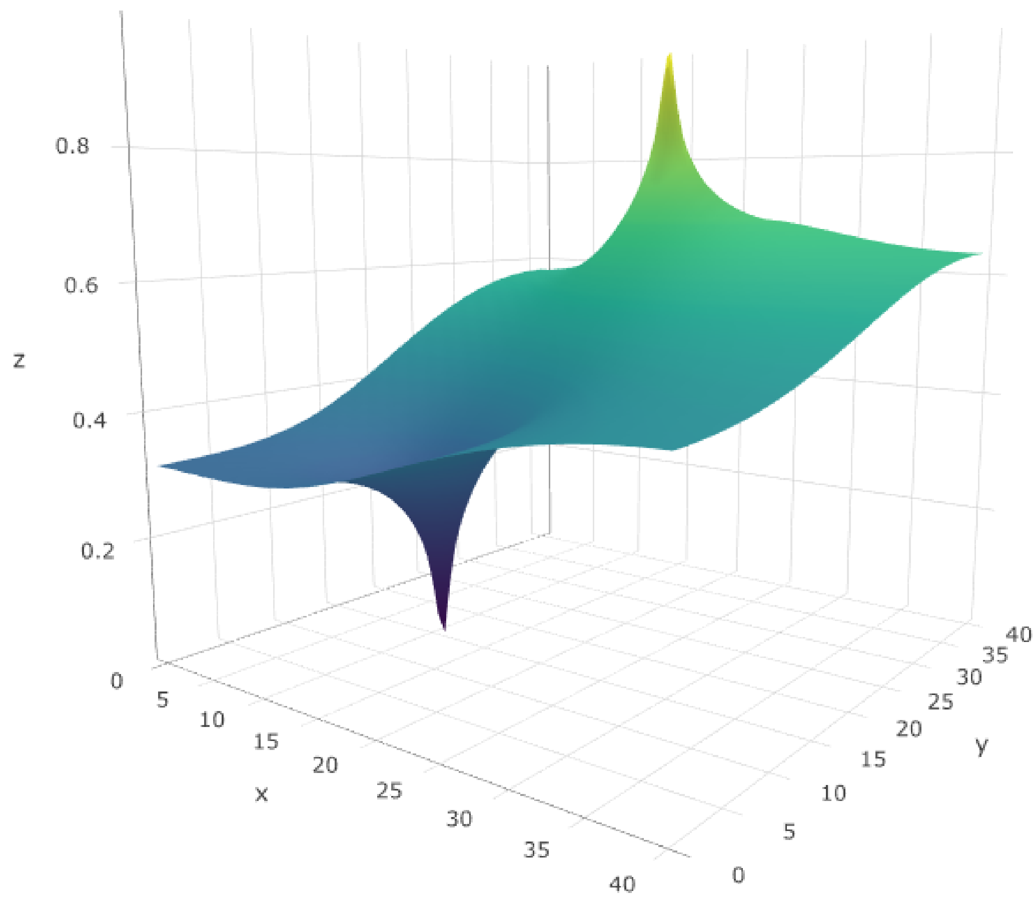

Figura 5.3: Ajuste harmónico de la función $\hat{w}(x)$.

Esta figura muestra la gráfica $(x,\phi^{(N)}(x))$, donde $\phi^{(N)}$ está definida como en la Ecuación (5.1.3). Esta figura muestra una función más suave y regular a comparación con la Figura 5.1, y en particular, se espera que esta función resuelva el problema.

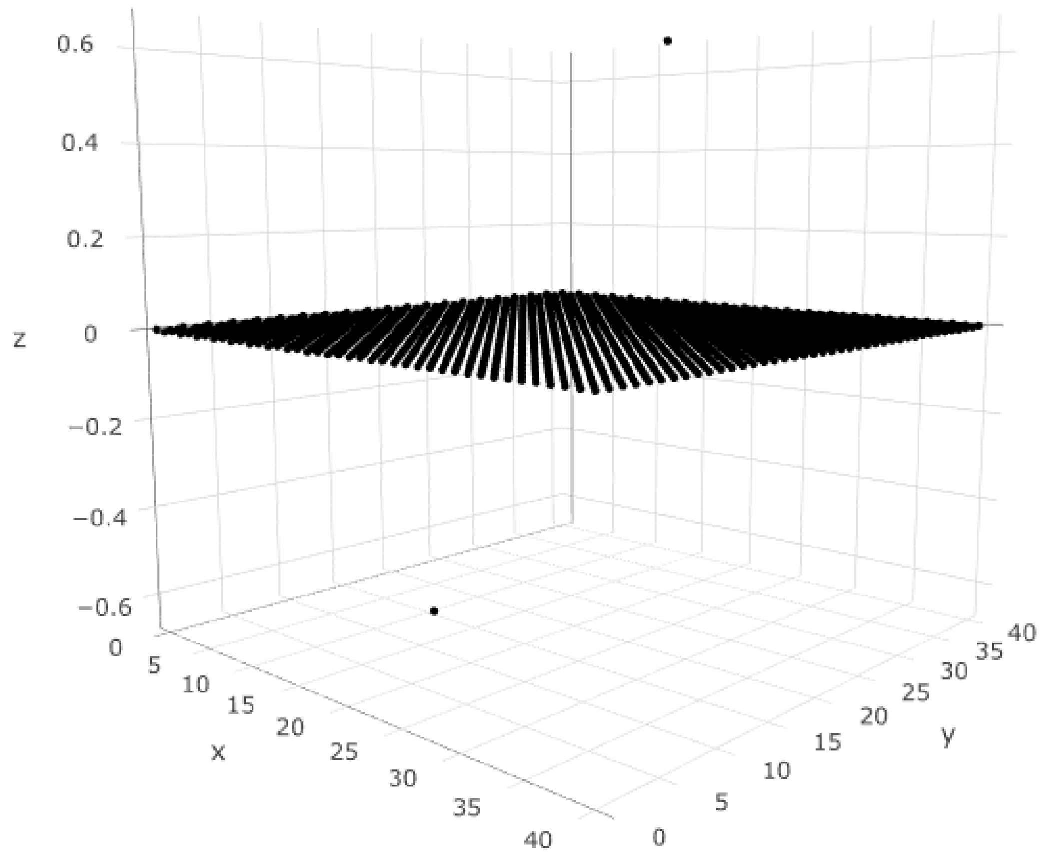

Figura 5.4: Divergencia asociada al ajuste harmónico de $\hat{w}(x)$.

Definamos $div\,\hat{\theta}\,(x) = \sum_{y:y\sim x} c(x,y)\,[\,\phi^{(N)}(x) - \phi^{(N)}(y)\,]$. Como podemos observar, esta función muestra una divergencia más uniforme, casi constante en 0 en el conjunto W. De este resultado podemos concluir que la aproximación se acerca a la función deseada, mostrando que $\phi^{(N)}$ es la función que resuelve el problema, es decir, $\phi^{(N)}$ es la extensión harmónica que cumple ser harmónica en W tal que $\phi^{(N)}(z) = 1$ y $\phi^{(N)}(a) = 0$.

Conclusiones

La Teoría de las cadenas de Markov, a pesar de ser un caso particular de los procesos de Markov estas toman un rol especia en la Teoría general de los procesos estocásticos. En general, la hipótesis sobre un espacio numerable finito o infinito de estados es la hipótesis que define a lo que llamamos como "cadena" en este trabajo de manera que genera en si una serie de preguntas acompañada de la demanda de definiciones claras y precisas. En particular, el estudio de esta clase de procesos puede presentar dificultades técnicas de manera que este texto aborda estas dificultades hasta cierto punto. La existencia de estos procesos queda establecida a través de su contrucción y sus caracterizaciones.

De este trabajo se espera que sirva de ilustración al enfoque moderno y riguroso de los procesos estocásticos presentando una introducción al estudio de las cadenas de Markov a tiempo continuo como también mostrando la conexión de la Teoría del Potencial con los procesos estocásticos a través de las cadenas de Markov. En conclusión, este trabajo presenta las ideas introductorias relacionadas con la Teoría del Potencial para Cadenas de Markov con la intención de aportar un mejor entendimiento de ésta teoría para cualquier estudiante a nivel universitario.

Apéndice

A Matriz Exponencial

Sea $M_{N\times N}(\mathbb{R})$ el espacio de matrices cuadradas con entradas en los reales de $N \times N$. Sea I un conjunto numerable finito de cardinalidad N. Dada cualquier matriz real $Q = (q_{ij})_{i,j\in I}$, la matriz Q define un operador $T : \mathbb{R}^N \to \mathbb{R}^N$ dado por

$$Tx = y = Qx \tag{A.1}$$

donde $x = (x_j)$ y $y = (y_j)$ son vectores columna de N componentes respectivamente. De esta forma, si consideramos el producto de matrices, entonces en la ecuación (A.1) se obtiene que y se puede expresar en términos de las componentes de la matriz Q y el vector x como sigue

$$y_j = \sum_{k=1}^{N} q_{jk}x_k. \tag{A.2}$$

Definamos dos normas en el espacio de matrices dadas por

$$\|Q\| = \sup_{x\neq 0} \frac{\|Qx\|}{\|x\|} \qquad , \qquad \|Q\|_\infty = \sup_{i,j} |q_{ij}|.$$

Claramente la norma $\|Q\|_\infty$ es finita ya que la matriz Q está definida a partir de un conjunto numerable finito y cuyo valor se alcanza en alguna componente de la matriz en cuestión. A la norma $\|Q\|$ la definimos a partir de la norma usual de $\mathbb{R}^N$ dada por

$$\|x\| = \Big(\sum_{j=1} x_j^2\Big)^{1/2}. \tag{A.3}$$

A continuación mostraremos que ambas normas son equivalentes y que $\|Q\|$ está bien adaptada a la suma y producto de matrices.

Lema A.1. *Para cualquier matriz Q, Q_1 y Q_2 tenemos lo siguiente:*

(a) $\|Q\|_\infty \leq \|Q\| \leq N\|Q\|_\infty$*;*

(b) $\|Q_1 + Q_2\| \leq \|Q_1\| + \|Q_2\|$ *y* $\|Q_1Q_2\| \leq \|Q_1\|\,\|Q_2\|$*.*

Demostración:

(a) Definamos $y = Qx$. Sea y_j como se muestra en (A.2), si calculamos la norma de $||y||$ como se muestra en (A.3), para cualquier vector x y por la desigualdad de Cauchy-Schwarz se tiene que

$$\begin{aligned}\|Qx\|^2 &= \sum_{j=1}^{N}\left(\sum_{k=1}^{N} q_{jk}x_k\right)^2 \\ &\leq \sum_{j=1}^{N}\left(\left(\sum_{k=1}^{N} q_{jk}^2\right)^{1/2}\left(\sum_{k=1}^{N} x_m^2\right)^{1/2}\right)^2, \end{aligned} \tag{A.4}$$

como $\|x\| = \left(\sum_{j=1} x_j^2\right)^{1/2}$ es una constante real no negativa que no depende de los índices de cada una de las dos sumas de la desigualdad anterior. De esta forma podemos factorizar de manera que de la desigualdad (A.4) se sigue que

$$\|Qx\|^2 \leq \|x\|^2\left(\sum_{j=1}^{N}\sum_{k=1}^{N} q_{jk}^2\right). \tag{A.5}$$

Por lo anterior

$$\|Qx\|^2 \leq c^2\,\|x\|^2 \qquad \text{donde} \qquad c = \left(\sum_{j=1}^{N}\sum_{k=1}^{N} q_{jk}^2\right)^{1/2}.$$

Sea x con $||x|| \neq 0$. Por la desigualdad (A.5) tenemos que c es una cota superior de la cantidad $\|Qx\|/\|x\|$, de manera que por propiedades del supremo deducimos que

$$\frac{||Qx||}{||x||} \leq \sup_{x\neq 0}\frac{||Qx||}{||x||} \leq c. \tag{A.6}$$

Como $\|Q\| = \sup_{x\neq 0} \|Qx\| / \|x\|$ y por la desigualdad (A.6) concluimos que $\|Q\|$ es una norma finita. Finalmente, si $||x|| \neq 0$ por la desigualdad (A.6) y la definición de $\|Q\|$ tenemos que $\|Qx\| / \|x\| \leq \|Q\|$, implica que $\|Qx\| \leq \|Q\| \|x\|$. En particular, para los vectores $\varepsilon_j = (0,...,1,...,0)$, con 1 en la $j-$*ésima* entrada tenemos que $\left\|Q\varepsilon_j\right\| \leq \|Q\|$. Como I es un conjunto numerable finito, el supremo que define a la norma $\|Q\|_\infty$, se alcanza en algún renglón j de la matriz Q de manera que

$$\|Q\|_\infty^2 \leq \sum_{i=1}^{N} (q_{ij})^2 = \left\|Q\varepsilon_j\right\|^2 \leq \|Q\|^2,$$

mostr que

$$\|Q\|_\infty \leq \|Q\| \tag{A.7}$$

Retomando la ecuación (A.5), en particular para toda $i, j \in I$ se tiene que $q_{ij} \leq \|Q\|_\infty$ de manera que

$$\begin{aligned} \|x\|^2 \left(\sum_{j=1}^{N} \sum_{k=1}^{N} q_{jk}^2 \right) &\leq \|x\|^2 \left(\sum_{j=1}^{N} \sum_{k=1}^{N} \|Q\|_\infty^2 \right) \\ &= \|x\|^2 \|Q\|_\infty^2 N^2, \end{aligned}$$

para así obtener que

$$\|Q\| \leq N \|Q\|_\infty. \tag{A.8}$$

Por lo tanto, por la desigualdad (A.7) en conjunto con (A.8) se demuestra el inciso (a) del teorema.

(b) Para cualquier vector x en $\mathbb{R}^N$ y dos matrices Q_1 y Q_2, por la desigualdad del triángulo y el inciso anterior tenemos que

$$\begin{aligned} \|(Q_1 + Q_2)x\| &= \|Q_1 x + Q_2 x\| \\ &\leq \|Q_1 x\| + \|Q_2 x\| \\ &\leq \|Q_1\| \|x\| + \|Q_2\| \|x\| \\ &= (\|Q_1\| + \|Q_2\|) \|x\|; \end{aligned} \tag{A.9}$$

por otra parte, de manera similar

$$\begin{aligned} \|Q_1 Q_2 x\| &= \|Q_1(Q_2 x)\| \\ &\leq \|Q_1\| \|Q_2 x\| \leq \|Q_1\| \|Q_2\| \|x\|,. \end{aligned} \tag{A.10}$$

Por (A.9) y (A.10) se demuestra el inciso (b) del teorema.

$\square$

Ahora, procederemos a demostrar que la matriz exponencial definida como

$$e^{tQ} = \sum_{k=0}^{\infty} \frac{(tQ)^k}{k!},$$

converge uniformemente en $[0,\infty)$ en el espacio de matrices reales. Definamos una función de t al espacio de matrices reales dada por $t \to e^{tQ}$. Para $n = 0, 1, 2, ...,$ consideremos la suma finita

$$E^{(n)}(t) = \sum_{k=0}^{n} \frac{(tQ)^k}{k!}.$$

Antes de proceder, se considera el siguiente teorema cuya demostración se omite y se refiere a [Apostol, pps 234]

Teorema A.2 (Prueba de la razón para convergencia de series de funciones)**.** *Dada un serie de potencias $\sum_{n=0} a_n(x-x_0)^n$, definamos*

$$\lambda = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|, \qquad\qquad r = \frac{1}{\lambda},$$

(donde $r = 0$ si $\lambda = \infty$ y $r = \infty$ si $\lambda = 0$). Entonces la serie converge absolutamente si $|x-x_0| < r$ y diverge si $|x-x_0| > r$. Además, la serie converge uniformemente en todo subconjunto compacto dentro del intervalo de convergencia dado por $|x-x_0| < r$.

Para toda i, j y para $m \leq n$, por el Lema A.1 inciso (a) se cumple que

$$\begin{aligned} |(E^{(n)}(t) - E^{(m)}(t))_{i,j}| &\leq \left\| E^{(n)}(t) - E^{(m)}(t) \right\|_\infty \\ &\leq \left\| E^{(n)}(t) - E^{(m)}(t) \right\|, \end{aligned}$$

de manera que si sustituimos en la última desigualdad

$$\begin{aligned} \left\| E^{(n)}(t) - E^{(m)}(t) \right\| &= \left\| \sum_{k=m+1}^{n} \frac{(tQ)^k}{k!} \right\| \\ &\leq \sum_{k=m+1}^{n} \frac{\|tQ\|^k}{k!}. \end{aligned} \tag{A.11}$$

Debido al inciso (a) del Lema A.1, $\|Q\| < N\|Q\|_\infty < \infty$ mostrando que la serie (A.11) es una suma finita. Por el Lema A.1 inciso (a), tenemos que

$$\begin{aligned}\sum_{k=0}^{n} \frac{\|t\,Q\|^k}{k!} &\leq \sum_{k=0}^{n} \frac{(N\,\|t\,Q\|_\infty)^k}{k!} \\ &= \sum_{k=0}^{n} \frac{(N\,\|Q\|_\infty)^k}{k!} t^k. \end{aligned} \tag{A.12}$$

Definamos $a_n = (N\,\|Q\|_\infty)^k/k!$, por la prueba de la razón para la convergencia de series de funciones,

$$\begin{aligned}\lambda &= \lim_{n\to\infty} \frac{\frac{(N\|Q\|_\infty)^{n+1}}{n+1!}}{\frac{(N\|Q\|_\infty)^{n}}{n!}} \\ &= \lim_{n\to\infty} \frac{N\,\|Q\|_\infty}{n} \\ &= N\,\|Q\|_\infty \left(\lim_{n\to\infty} \frac{1}{n}\right) \\ &= 0.\end{aligned}$$

A consecuencia del Teorema A.2, por lo anterior tenemos que el radio de convergencia de la serie (A.12) es infinito mostrando que la serie converge absolutamente y en particular, converge uniformemente en $[0,\infty)$. Debido a este hecho, cuando $m,n \to \infty$

$$\sum_{k=m+1}^{n} \frac{\|t\,Q\|^k}{k!} \to 0,$$

pues

$$\sum_{k=m+1}^{n} \frac{\|t\,Q\|^k}{k!} \leq \sum_{k=m+1}^{n} \frac{(N\,\|Q\|_\infty)^k t^k}{k!} \to 0,$$

cuando $m,n \to \infty$. Por lo tanto para toda $i,j \in I$, para toda $\varepsilon > 0$, existe M tal que para toda $m > M$ y $n > M$ se tiene que

$$|(E^{(n)}(t) - E^{(m)}(t))_{i,j}| < \varepsilon,$$

mostrando que la sucesión $E^{(n)}$ define una sucesión de funciones que satisface la condición de Cauchy. De este hecho en particular se desprende que

$$e^Q = \sum_{k=0}^{\infty} \frac{Q^k}{k!},$$

está bien definida y que la serie de potencias

$$\sum_{k=0}^{\infty} \frac{(tQ)_{ij}^k}{k!}$$

converge uniformemente a $(e^{tQ})_{ij}$. La continuidad de la función $(e^{tQ})_{ij}$ en $[0,\infty)$ se sigue del siguiente teorema cuya demostración se omite y se refiere a [Apostol, pps 234] para su demostración.

Teorema A.3. *Asumamos que la serie de potencias* $\sum_{n=0}^{\infty} a_n(x-x_0)$ *converge para cada* x *en* (x_0-r, x_0+r). *Entonces la función* f *definida por la ecuación*

$$f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n,$$

con x *en* (x_0-r, x_0+r), *es continua en* (x_0-r, x_0+r).

Finalmente, para 2 matrices Q_1 y Q_2 que conmutan se tiene que

$$\begin{aligned} e^{Q_1+Q_2} &= \sum_{k=0}^{\infty} \frac{(Q_1+Q_2)^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{n=0}^{k} \frac{k!}{n!(k-n)!} Q_1^n Q_2^{k-n} \\ &= \sum_{n=0} \frac{Q_1^n}{n!} \sum_{k=n}^{\infty} \frac{Q_2^{k-n}}{(k-n)!} \\ &= e^{Q_1} e^{Q_2}. \end{aligned}$$

B Código implementado

```
### Red G = (V,E)
##  Conjunto de V\'ertices V
Zn <- function(n){
  a = seq(0,n)
  Zn = as.data.frame.list(expand.grid(a,a))
  Zn = data.frame(matrix(unlist(Zn), ncol=2, byrow=F),
    stringsAsFactors=FALSE)
  v = rep("v",length(Zn[,1]))
  row.names(Zn) <- paste(v,rownames(Zn),sep = "")
```

```
  return(Zn)
}
##  Auxiliar: Para control del vector f_{0} = (f(v1), ..., f(vn)).
Generadorh0 <- function(x,y){
  V = get("V",1)
  dW <- x
  hVminusW <- t(as.matrix(y))
  colnames(hVminusW) <- dW
  return(hVminusW)
}
##  Auxiliar: Para generar W = V \ dW
GeneradorW <- function(x){
  V = get("V",1)
  dW <- x
  W = setdiff(rownames(V),dW)
  return(W)
}
## Matriz de conductancias.
matriz.adyacenciaZn <- function(A){
  n = length(row.names(A))
  v = rep(0,n)
  AZn = matrix(0,n,n)
  colnames(AZn) <- row.names(A)
  rownames(AZn) <- row.names(A)
  for(i in 1:n){
    for(k in 1:n){
      v[k] = (A[i,1] - A[k,1])^2 + (A[i,2] - A[k,2])^2
      if(!v[k] == 1){
        v[k] = 0
      }
    }
    AZn[i,] = v
  }
  return(AZn)
}
## Matriz transici\'on P asociada a la red G.
matriz.transcXn <- function(A){
  n = length(rownames(A))
  Pi = rep(0,n)
  P = A
  for(i in 1:n){
    c = sum(A[i,])
    Pi[i] = c
    P[i,] = A[i,]/c
```

```
  }
  return(P)
}
## Vector distribuci\'on estacionaria pi
EstacionariaP <- function(A){
  n = length(rownames(A))
  C = rep(0,n)
  Pi = rep(0,n)
  for(i in 1:n){
    cx = sum(A[i,])
    C[i] = cx
    Pi[i] = cx
  }
  cG = sum(C)
  Pi = t(as.matrix((1/cG)*Pi))
  colnames(Pi) = rownames(A)
  return(Pi)
}
### Simulaci\'on
## Caminata Aleatoria iniciada en "x" parada en un conjunto dado
    W subconjunto de V.
Hit.Time.VsetminusW <- function(P,n,x,W){
x0 = x
xn = x
Xn = x
m = which(rownames(P) == x)
cg = which(P[m,] > 0)
G = as.matrix(t(P[m,cg]))
for(i in 1:n){
  if(xn %in% W){
    break()
  }
  else{
    k = 1
    B = 0
    U = runif(1,0,1)
    while(B < U){
      B = sum(G[1:k])
      k = k+1
    }
    m = which(rownames(P) == colnames(G)[k-1])
    xn = rownames(P)[m]
    Xn = c(Xn,xn)
    cg = which(P[m,] > 0)
    G = as.matrix(t(P[m,cg]))
```

```
    }
}
return(Xn)
}
## Hace lo mismo que la funci\'on anterior pero S\'oLO simula
    Xtau_dW.
HHit.Time.VsetminusW <- function(P,n,x,W){
  xn = x
  m = which(rownames(P) == x)
  cg = which(P[m,] > 0)
  G = as.matrix(t(P[m,cg]))
  for(i in 1:n){
    if(xn %in% W){
      break()
    }
    else{
      k = 1
      B = 0
      U = runif(1,0,1)
      while(B < U){
        B = sum(G[1:k])
        k = k+1
      }
      m = which(rownames(P) == colnames(G)[k-1])
      x = rownames(P)[m]
      xn = x
      cg = which(P[m,] > 0)
      G = as.matrix(t(P[m,cg]))
    }
  }
  return(xn)
}
### Funciones Harm\'onicas
## Estimador puntual de w(x) = E_{x} [f_{0} (X_tau_VminusdW)]
    (w(x) gorro).
Harmonica <- function(n,m,x,W,hW,P){
  h = rep(x,n+10)
  H = sapply(h, function(y) HHit.Time.VsetminusW(P,m,y,W))
  m = sapply(H, function(y) which(colnames(hW) == y))
  HA = sapply(m, function(y) hW[y])
  hx = mean(HA[10:(n+10)])
  return(hx)
}
## Extensi\'on harm\'onica (w gorro, definida en toda la red G
    ).
```

```
Ext.Harmonica <- function(V,hW){
  W = colnames(hW)
  h = sapply(rownames(V), function (y) Harmonica(100,50000,x,W,
    hW,P))
  return(h)
}
##  Para calcular la divergencia en cada v\'ertice de V.
#   E denota la matriz de aristas.
divergencia <- function(P,E,w){
  h = names(w)
  n = length(h)
  m = seq(1,n)
  r = rownames(P)
  s = sapply(h, function(x) which(r == x))
  vecinos = sapply(s, function(x) which(E[x,] > 0))
  carga = sapply(vecinos, function(x) length(x))
  sumpxyvy = sapply(h,function(x) P[x,] %*% w )
  cxvx = sapply(m, function(x) carga[x]*w[x])
  divx = sapply(m, function(x) cxvx[x] - carga[x]*sumpxyvy[x])
  return(divx)
}
##  Pre-ajuste harm\'onico.
#   Producto Pw. Pre-ajuste harm\'onico para la funci\'on w
    gorro en el conjunto W.
ajusteharmonicoW <- function(P,E,w,W){
  h = W
  vol = w
  n = length(h)
  m = seq(1,n)
  r = rownames(P)
  S = which(r %in% h)
  vecinos = sapply(S, function(x) which(E[x,] > 0))
  sumpxyvy = sapply(S,function(x) P[x,] %*% w )
  vol[S] = sumpxyvy
  return(vol)
}
##  Ajuste harm\'onico.
#   Producto P^{N} w. Ajuste harmonico para la funci\'on w gorro
    en el conjunto W.
harm <- function(n,P,E,w,W){
  aj = ajusteharmonicoW(P,E,w,W)
  for(i in 1:n){
    aj = ajusteharmonicoW(P,E,aj,W)
  }
  return(aj)
```

```
}
```

Código (R) 1: Código implementado.

```
# V\'ertices de la red G = (v,{e}).
V = Zn(40)
head(V)
# Aristas, en este caso la conductancia es igual a 1.
E = matriz.adyacenciaZn(V)
E[1:3,1:3]
# Matriz de transici\'on de la caminata aleatoria asociada a G.
P = matriz.transcXn(E)
P[1:3,1:3]
# Distribuci\'on estacionaria de P.
Pi = EstacionariaP(P)
Pi[,1:6]
sum(Pi)
# Conjunto de ejemplo.
dW1 = c("v547","v1251")
h01 = Generadorh0(dW1,c(0,1))
W1 = GeneradorW(dW1)
Xtau_dW1 = Hit.Time.VsetminusW(P,50000,"v250",dW1)
which(Xtau_dW1 %in% dW1)
length(Xtau_dW1)
# ADVERTENCIA: La siguiente funci\'on puede tomar mucho tiempo.
    Se recomienda usar par\'ametro "n" menor o igual a  100 en
    una red MUY grande como la ocupada.
voltaje1 = Harmonica(10,50000,W1,dW1,h01,P) # Estimaci\'on de w(
    x).
har1 = matrix(as.vector(voltaje1),41,41,byrow = TRUE) # Matrix
    auxiliar para graficar.
Ajharm1 = harm(1000,P,E,voltaje1,W1) # w(x). Soluci\'on al
    Problema de Dirichlet.
H1 = matrix(haraj1,41,41,byrow = TRUE) # Matrix auxiliar para
    graficar.
# Librer\'ia para graficar
library(plotly)
x = seq(0,40)
## Gr\'aficas
# Estimaci\'on w(x).
p1 = plot_ly(x = x, y = x, z = har1)  %>% add_surface()
# Soluci\'on.
p2 = plot_ly(x = x, y = x, z = H1)  %>% add_surface()
# Divergencia de la estimaci\'on.
j1 = plot_ly(x = V[,1], y = V[,2], z = divergencia(P,E,voltaje1)
    , color = I("black"), marker = list(size = 2.5)) %>%
```

```
  add_markers()
j2 = plot_ly(x = V[,1], y = V[,2], z = divergencia(P,E,har1),
    color = I("black"), marker = list(size = 2.5)) %>%
  add_markers()
```

Código (R) 2: Ejemplos.

Bibliografía

[Chu60] Kai Lai Chung. *Markov Chains With Stationary Transition Probabilities*. Primera ed. Berlín, Alemania: Springer-Verlag, 1960.

[Apo74] Tom M. Apostol. *Mathematical Analysis*. Segunda ed. Estados Unidos: Addison-Wesley Publishing Company,Inc., 1974.

[Sam75] Howar M. Taylor Samuel Karlin. *A First Course In Stochastic Processes*. Quinta ed. Nueva York, Estados Unidos: Academic Press, 1975.

[Spi75] Michael Spivak. *Calculus Cálculo Infinitesimal*. Segunda ed. Nueva York, Estados Unidos: Academic Press, 1975.

[Kre78] Erwin Kreyszig. *Introductory Functional Analysis with applications*. Primera ed. Estados Unidos: John Wiley and Sons, 1978.

[Doo84] J.L. Doob. *Classical Potential Theory and Its Probabilistic Counterpart*. Primera Ed. Nueva York, Estados Unidos: Springer-Verlag, 1984.

[Fre89] David Freedman. *Markov Chains*. Primera ed. Nueva York, Estados Unidos: Springer-Verlag, 1989.

[Doo90] J.L. Doob. *Stochastic Processes*. Edición revisada. Estados Unidos: Wiley Classics Library, 1990.

[Nor97] J.R. Norris. *Markov Chains*. revisar. Cambridge, Inglaterra: Cambridge University Press, 1997.

[Rus97] Yuval Peres Russel Lyon. *Probability on Trees and Networks*. Primera ed. Cambridge, Inglaterra: Cambridge University Press, 1997.

[DS00] P.G. Doyle and J.L. Snell. "Random Walks and Electric Networks". Version 3.02. In: (Jan. 2000).

[Car04] N.L. Carothers. *Real Analysis*. Primera ed. Nueva York, Estados Unidos: Cambridge University Press, 2004.

[Ros13] Sheldon Ross. *Simulation*. Quinta ed. Oxford, Inglaterra: Elsevier, 2013.

[Sam15] Robert J. Elliott Samuel N. Cohen. *Stochastic Calculus and Applications*. Segunda ed. Estados Unidos: Birkhäuser, 2015.

[Dan17] Yuval Peres Daniel A. Levin Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. Segunda ed. Estados Unidos: American Mathematical Society, 2017.