Cadenas de Markov
y Teoría de Potencial

Publicações Matemáticas

Cadenas de Markov y Teoría de Potencial

Johel Beltrán
Pontificia Universidad Católica del Perú

impa

28º Colóquio Brasileiro de Matemática

Impresso no Brasil / Printed in Brazil

Capa: Noni Geiger / Sérgio R. Vaz

28º Colóquio Brasileiro de Matemática

- Cadenas de Markov y Teoría de Potencial - Johel Beltrán
- Cálculo e Estimação de Invariantes Geométricos: Uma Introdução às Geometrias Euclidiana e Afim - M. Andrade e T. Lewiner
- De Newton a Boltzmann: o Teorema de Lanford - Sérgio B. Volchan
- Extremal and Probabilistic Combinatorics - Robert Morris e Roberto Imbuzeiro Oliveira
- Fluxos Estrela - Alexander Arbieto, Bruno Santiago e Tatiana Sodero
- Geometria Aritmética em Retas e Cônicas - Rodrigo Gondim
- Hydrodynamical Methods in Last Passage Percolation Models - E. A. Cator e L. P. R. Pimentel
- Introduction to Optimal Transport: Theory and Applications - Nicola Gigli
- Introdução à Aproximação Numérica de Equações Diferenciais Parciais Via o Método de Elementos Finitos - Juan Galvis e Henrique Versieux
- Matrizes Especiais em Matemática Numérica - Licio Hernanes Bezerra
- Mecânica Quântica para Matemáticos em Formação - Bárbara Amaral, Alexandre Tavares Baraviera e Marcelo O. Terra Cunha
- Multiple Integrals and Modular Differential Equations - Hossein Movasati
- Nonlinear Equations - Gregorio Malajovich
- Partially Hyperbolic Dynamics - Federico Rodriguez Hertz, Jana Rodriguez Hertz e Raúl Ures
- Random Process with Variable Length - A. Toom, A. Ramos, A. Rocha e A. Simas
- Um Primeiro Contato com Bases de Gröbner - Marcelo Escudeiro Hernandes

ISBN: 978-85-244-0316-3

Distribuição: IMPA
Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ
E-mail: ddic@impa.br
http://www.impa.br

Conteúdo

Capítulo 1

Preliminares

En este capítulo introducimos los conceptos fundamentales para estudiar en el próximo capítulo las cadenas de Markov.

1.1 $S^{\mathbb{Z}_+}$, proyecciones y cilindros

Fijaremos un conjunto S enumerable o finito con al menos dos elementos. En toda la monografía, vamos a denotar por $\mathbb{Z}_+$ el conjunto de enteros no negativos: $\mathbb{Z}_+ = \{0, 1, 2 \dots\}$. Para cada entero $k \geq 1$, S^k denotará el producto cartesiano de k copias de S:

$$S^k \ := \ \big\{(a_1, a_2, \dots, a_k) : a_j \in S \ \forall j = 1, \dots, k\big\} \ .$$

Por supuesto, $S^1 = S$. $S^{\mathbb{Z}_+}$ denotará el espacio de todas las sucesiones sobre S:

$$S^{\mathbb{Z}_+} \ := \ \big\{ \, (x_0, x_1, \dots) : x_i \in S \ \forall i \in \mathbb{Z}_+ \, \big\} \ .$$

Usaremos símbolos como $\vec{x}$, $\vec{y}$ para denotar los elementos de $S^{\mathbb{Z}_+}$.

Ejercicio 1.1. *Cada conjunto S^k es enumerable. El conjunto $S^{\mathbb{Z}_+}$ es enumerable? Ni siquiera $\{0, 1\}^{\mathbb{Z}_+}$ es enumerable.*

Definimos las proyecciones

$$X_n : S^{\mathbb{Z}_+} \to S \ , \quad \text{para cada } n \geq 0 \ ,$$

como $X_n(\vec{x}) = x_n$ si $\vec{x} = (x_0, x_1, \dots)$, es decir $X_n(\vec{x})$ indica la n-ésima coordenada de $\vec{x}$. Además, para cualquier subconjunto finito $\{n_0, n_1, \dots, n_k\} \subset \mathbb{Z}_+$, $k \geq 0$, la proyección

$$(X_{n_0}, X_{n_1}, \dots, X_{n_k}) : S^{\mathbb{Z}_+} \to S^{k+1} ,$$

se define como $(X_{n_0}, X_{n_1}, \dots, X_{n_k})(\vec{x}) := \big(X_{n_0}(\vec{x}), X_{n_1}(\vec{x}), \dots, X_{n_k}(\vec{x})\big)$.

Ejercicio 1.2. *Supongamos que existen $k \geq 1$, $A \subseteq S^k$, $m \geq 1$ y $B \subseteq S^m$ de modo que*

$$\begin{aligned} &\{\vec{x} \in S^{\mathbb{Z}_+} : (X_0, \dots, X_{k-1})(\vec{x}) \in A\} \\ = \; &\{\vec{x} \in S^{\mathbb{Z}_+} : (X_0, \dots, X_{m-1})(\vec{x}) \in B\} . \end{aligned}$$

Pruebe que si $k = m$ entonces $B = A$ y si $k < m$ entonces $B = A \times \underbrace{S \times \cdots \times S}_{(k-m) \text{ veces}}$.

Estaremos interesados en el siguiente tipo de subconjuntos de $S^{\mathbb{Z}_+}$.

Definición 1.1. *Llamaremos cilindros a los subconjuntos de $S^{\mathbb{Z}_+}$ que tienen la siguiente forma:*

$$(X_0, \dots, X_k)^{-1}(A) \; = \; \Big\{\vec{x} \in S^{\mathbb{Z}_+} : (X_0, \dots, X_k)(\vec{x}) \in A\Big\} ,$$

para algún $k \geq 0$ y algún subconjunto $A \subseteq S^{k+1}$. En particular, $S^{\mathbb{Z}_+}$ y $\varnothing$ son cilindros (verifique).

Ejercicio 1.3. *(a) Pruebe que $X_n^{-1}(A) \subseteq S^{\mathbb{Z}_+}$ es cilindro, para todo $n \geq 0$ y $A \subseteq S$.*

(b) Pruebe que para cualquier $k \geq 1$, proyección $(X_{n_0}, X_{n_1}, \dots, X_{n_{k-1}})$ y cualquier $A \subseteq S^k$, el conjunto

$$\Big\{\vec{x} \in S^{\mathbb{Z}_+} : (X_{n_0}, X_{n_1}, \dots, X_{n_{k-1}})(\vec{x}) \in A\Big\}$$

es un cilindro.

Ejercicio 1.4. *(a) Si $C_1, \dots, C_\ell$ son cilindros, pruebe que para todo $m \geq 1$ existe $k \geq m$ y existen $A_1, \dots, A_\ell \in S^k$ de modo que*

$$C_i \; = \; (X_0, \dots, X_{k-1})^{-1}(A_i) \, , \quad \forall i = 1, \dots, \ell \, .$$

(*b*) *Pruebe que la intersección de cualquier número finito de cilindros es un cilindro.*

(*c*) *Pruebe que la unión de cualquier número finito de cilindros es un cilindro.*

Para simplificar la notación, de ahora en adelante escribiremos

$$\begin{aligned}&\big\{(X_{n_0}, X_{n_1}, \ldots, X_{n_k}) \in A\big\}\\ &= \big\{\vec{x} \in S^{\mathbb{Z}_+} : (X_{n_0}, X_{n_1}, \ldots, X_{n_k})(\vec{x}) \in A\big\}\,.\end{aligned}$$

En particular, para $A_i \subseteq S$, $i = 0, \ldots, k$ denotaremos

$$\begin{aligned}&\{X_{n_0} \in A_0\,,\ X_{n_1} \in A_1\,,\ \ldots\,,\ X_{n_k} \in A_k\}\\ &:= \big\{\vec{x} \in S^{\mathbb{Z}_+} : X_{n_0}(\vec{x}) \in A_0, X_{n_1}(\vec{x}) \in A_1, \ldots, X_{n_k}(\vec{x}) \in A_k\big\}\,.\end{aligned}$$

Usando esta notación, es claro que

$$\{X_{n_0} \in A_0\,,\ X_{n_1} \in A_1\,,\ \ldots\,,\ X_{n_k} \in A_k\} \;=\; \cap_{i=0}^{k}\{X_{n_i} \in A_i\}\,.$$

1.2 Trayectorias y tiempos de llegada

Nuestro plan es usar $S^{\mathbb{Z}_+}$ como el conjunto de todas las trayectorias que un individuo que se desplaza sobre S puede realizar. Por ejemplo, dado $(a_0, \ldots, a_k) \in S^{k+1}$, el cilindro $\{X_0 = a_0, X_1 = a_1, \ldots, X_k = a_k\}$ es interpretado como el conjunto de todas las trayectorias que comienzan en a_0 y en cada instante n (desde $n = 0$ hasta $n = k$) se encuentran en a_n. Nada es dicho en el cilindro $\{X_0 = a_0, X_1 = a_1, \ldots, X_k = a_k\}$ sobre la trayectoria después del instante $n = k$.

Ejercicio 1.5. *¿Cómo se escribiría el cilindro que representa el conjunto de trayectorias que comienzan en $a \in S$ y luego de 4 pasos están en $u \in S$?*

Estamos usando el conjunto $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$ para representar el tiempo. Las siguientes funciones serán interesantes.

Definición 1.2. *Dado $A \subseteq S$ no vacío, definimos la función*

$$T_A : S^{\mathbb{Z}_+} \to \mathbb{Z}_+ \cup \{\infty\}$$

del siguiente modo. Fijada una trayectoria $\vec{x} \in S^{\mathbb{Z}_+}$, si el conjunto de instantes $\{n \geq 1 : X_n(\vec{x}) \in A\}$ es no vacío hacemos

$$T_A(\vec{x}) \;:=\; \min\{n \geq 1 : X_n(\vec{x}) \in A\}\,.$$

Si $\{n \geq 1 : X_n(\vec{x}) \in A\} = \varnothing$ hacemos $T_A(\vec{x}) = \infty$.

Notación: En la definición anterior, si $A = \{x\}$ escribiremos $T_x := T_{\{x\}}$.

Ahora es más fácil escribir el siguiente conjunto de trayectorias. Comienzo en $i \in S$ y llego a $a \in S$ antes de llegar a $u \in S$:

$$\{\vec{x} \in S^{\mathbb{Z}_+} : X_0(\vec{x}) = i\,,\; T_a(\vec{x}) < T_u(\vec{x})\} \;\subset\; S^{\mathbb{Z}_+}\,. \tag{1.2.1}$$

Otro ejemplo. Salgo de $e \in S$ y algún día en el futuro ($n \geq 1$) vuelvo a e:

$$\{\vec{x} \in S^{\mathbb{Z}_+} : X_0(\vec{x}) = e\,,\; T_e(\vec{x}) < \infty\} \;\subset\; S^{\mathbb{Z}_+}\,. \tag{1.2.2}$$

Para simplificar la notación, los conjuntos en (1.2.1) y (1.2.2) serán escritos como

$$\{X_0 = i\,,\; T_a < T_u\} \quad \text{y} \quad \{X_0 = e\,,\; T_e < \infty\}\,,$$

respectivamente. Aplicaremos la misma simplificación de notación en conjuntos similares.

Ejercicio 1.6. *Demuestre que los subconjuntos (1.2.1) y (1.2.2) NO son cilindros.*

Muchos subconjuntos de $S^{\mathbb{Z}_+}$ que no son cilindros son bastante interesantes y frecuentes de encontrar cuando uno estudia las trayectorias. Los cilindros son importantes por lo siguiente. Para estudiar una cadena de Markov tendremos que definir una probabilidad encima del espacio de trayectorias. En un principio, los únicos subconjuntos de $S^{\mathbb{Z}_+}$ sobre los cuales podremos asignar una probabilidad serán los cilindros. Fijadas las probabilidades sobre los cilindros, veremos que las probabilidades de una colección mayor de subconjuntos (como (1.2.1) y (1.2.2)) serán determinadas de manera única. Los cilindros funcionarán entonces como subconjuntos básicos para definir probabilidades sobre el espacio de trayectorias.

1.3 Probabilidad

Para definir probabilidad fijemos un contexto más abstracto. Sea Ω un conjunto cualquiera. En un caso simple, en que Ω es finito o enumerable, fijar una probabilidad sobre Ω consiste en indicar una probabilidad por cada elemento de Ω (sumando un total de uno) y luego declarar que la probabilidad de cada subconjunto de Ω es la suma de las probabilidades de sus elementos.

Por ejemplo, para cada $k \geq 1$, S^k es un conjunto enumerable. Denotemos por $\mathcal{P}(S^k)$ el conjunto potencia de S^k. Cualquier función $p : S^k \to [0,1]$, tal que $\sum_{u \in S^k} p(u) = 1$, determina una probabilidad $\mathbf{P} : \mathcal{P}(S^k) \to [0,1]$ como siendo

$$\mathbf{P}(E) = \sum_{u \in E} p(u) , \quad \forall E \in \mathcal{P}(S^k) . \tag{1.3.1}$$

Es decir, la probabilidad de cada subconjunto $E \subseteq S^k$ es declarada como $\sum_{u \in E} p(u)$, donde la suma es cero si E es vacío. Eso nos funciona para cada S^k, pero no para $S^{\mathbb{Z}_+}$ que no es enumerable.

Una probabilidad será una función que asocia a ciertos subconjuntos de Ω un número en $[0,1]$ y cumplirá ciertas propiedades. Denotemos por $\mathcal{C}$ una colección de subconjuntos de Ω. Queremos definir una aplicación

$$\mathbf{P} : \mathcal{C} \to [0,1] .$$

¿Has intentado definir una transformación lineal sobre un conjunto cualquiera? Para que transformación lineal sea un concepto interesante, necesitas que en ese conjunto se pueda (al menos) sumar. Para las propiedades que pediré a $\mathbf{P}$ necesito que $\mathcal{C}$ cumpla (al menos) con lo siguiente:

Definición 1.3. *Una colección $\mathcal{C}$ de subconjuntos de Ω es llamado λ-sistema si cumple con las siguientes propiedades.*

(λ1) $\Omega \in \mathcal{C}$

(λ2) *Si $E, F \in \mathcal{C}$ y $E \subseteq F$ entonces $F \setminus E \in \mathcal{C}$.*

(λ3) *Si $(E_n : n = 1, 2, \dots)$ es una sucesión de elementos de $\mathcal{C}$ tales que*

$$E_n \subseteq E_{n+1} , \quad \forall n \geq 1 , \quad \textit{(sucesión creciente)} \tag{1.3.2}$$

entonces $\cup_{n=1}^{\infty} E_n$ *también debe pertenecer a* $\mathcal{C}$.

Ejercicio 1.7. *Prueba que si* $\mathcal{C}$ *es un* λ*-sistema entonces valen las siguientes propiedades:*

(a) $\varnothing \in \mathcal{C}$.

(b) *Si* $E \in \mathcal{C}$ *entonces* $E^c \in \mathcal{C}$.

(c) *Si* E_1, E_2, $\ldots$, E_m *son elementos de* $\mathcal{C}$ *disjuntos dos a dos entonces* $\cup_{i=1}^{m} E_i \in \mathcal{C}$.

(d) *Si* $\{E_i : i \in I\}$ *es una familia enumerable de elementos de* $\mathcal{C}$*, disjuntos dos a dos, entonces* $\cup_{i\in I} E_i \in \mathcal{C}$.

Definamos probabilidad

Definición 1.4. *Sea* $\mathcal{C}$ *un* λ*-sistema de subconjuntos de* Ω*. Diremos que la aplicación* $\mathbf{P} : \mathcal{C} \to [0,1]$ *es una probabilidad si vale lo siguiente.*

(P1) $\mathbf{P}(\Omega) = 1$.

(P2) Si $E, F \in \mathcal{C}$ *son tales que* $E \subseteq F$ *entonces*

$$\mathbf{P}(F \setminus E) \;=\; \mathbf{P}(F) \;-\; \mathbf{P}(E)\;.$$

(P3) Si $(E_n : n = 1, 2 \ldots)$ *es una sucesión creciente (cumpliendo (1.3.2)) de elementos de* $\mathcal{C}$ *entonces*

$$\lim_{n\to\infty} \mathbf{P}(E_n) \;=\; \mathbf{P}(E)\;, \quad \textit{donde } E := \cup_{n=1}^{\infty} E_n.$$

Cada propiedad (λi), $i = 1, 2, 3$, sobre $\mathcal{C}$, otorga más sentido al respectivo requisito (Pi) pedido de $\mathbf{P}$.

Es bastante claro que $(P1)$ y $(P2)$ son requisitos mínimos para ser probabilidad. $(P3)$ tal vez no tanto. Sin embargo, con $(P1)$ y $(P2)$ sólo alcanza para resolver los problemas de dados, siempre que los tires una cantidad finita de veces. $(P3)$ nos permite determinar la probabilidad de conjuntos complejos a partir de las probabilidades de conjuntos más elementales.

Las propiedades (1.3.3) y (1.3.4) en el siguiente ejercicio son muy importantes para el cálculo de probabilidades. Estas propiedades son consecuencia de nuestra definición.

Ejercicio 1.8. *Si* $\mathbf{P}$ *es una probabilidad definida sobre un* λ*-sistema* $\mathcal{C}$ *vale lo siguiente.*

(a) $\mathbf{P}(\varnothing) = 0$.

(b) $\forall E \in \mathcal{C}$, $\mathbf{P}(E^c) = 1 - \mathbf{P}(E)$.

(c) *Si* $E, F \in \mathcal{C}$ *son disjuntos entonces*

$$\mathbf{P}(E \cup F) \; = \; \mathbf{P}(E) + \mathbf{P}(F) \, . \tag{1.3.3}$$

(d) *Si* $\{E_i : i \in I\}$ *es una colección enumerable de elementos de* $\mathcal{C}$*, disjuntos dos a dos, entonces*

$$\mathbf{P}\left(\cup_{i \in I} E_i\right) \; = \; \sum_{i \in I} \mathbf{P}(E_i) \, . \tag{1.3.4}$$

(e) *Si* $(E_n : n = 1, 2, \dots)$ *es una sucesión decreciente de elementos de* $\mathcal{C}$*, es decir,* $E_n \supseteq E_{n+1}$ $\forall n \geq 1$*, entonces*

$$\lim_{n \to \infty} \mathbf{P}(E_n) \; = \; \mathbf{P}(E) \, , \quad \textit{donde } E := \cap_{n=1}^{\infty} E_n .$$

Ejercicio 1.9. *Pruebe que* $\mathbf{P} : \mathcal{P}(S^k) \to [0, 1]$ *definido en (1.3.1) es una probabilidad según Definición 1.4. Verifique que toda probabilidad definida sobre el conjunto potencia de un conjunto enumerable es generado de ese modo.*

Ejercicio 1.10. *Sea* $\mathbf{P} : \mathcal{P}(S^k) \to [0, 1]$ *una probabilidad para algún* $k \geq 1$ *y sea* $E \in \mathcal{P}(S^k)$*. Pruebe que para todo* $\epsilon > 0$ *existe un subconjunto finito* $A \subseteq E$ *tal que*

$$\mathbf{P}(E) - \epsilon \; \leq \; \mathbf{P}(A) \; \leq \; \mathbf{P}(E) \, .$$

Los más experientes deben haber escuchado hablar de σ-álgebras.

1.3.1 σ-álgebras

Las propiedades de λ-sistema son suficientes para definir probabilidad. Sin embargo, no son suficientes para tratar con conceptos importantes en probabilidad como independencia. Definiremos a continuación un tipo más restricto de colección de subconjuntos llamado σ-álgebra.

Un σ-álgebra es un λ-sistema que cumple con la siguiente propiedad adicional.

Definición 1.5. *Una colección* $\mathcal{C}$ *es llamada* π*-sistema si se cumple la siguiente condición:*

(π) *Si* $E, F \in \mathcal{C}$ *entonces* $E \cap F \in \mathcal{C}$ *.*

Definición 1.6. *Una colección* $\mathcal{C}$ *es llamada* σ*-álgebra si además de ser* λ*-sistema cumple con la propiedad* (π)*.*

Un σ-álgebra $\mathcal{C}$ representa una colección de eventos que pertenecen a un cierto nivel de informacion. $A \in \mathcal{C}$ si y sólo si la ocurrencia o no del evento A es información disponible en el nivel de información asociado a $\mathcal{C}$. Es por eso natural pedir que la ocurrencia simultánea $A \cap B$ sea una información disponible si la ocurrencia tanto de A como de B son informaciones disponibles.

Ejercicio 1.11. $\mathcal{C}$ *es un* σ*-álgebra si y sólo si valen las siguientes propiedades:*

(i) $\varnothing \in \mathcal{C}$*.*

(ii) Si $E \in \mathcal{C}$ *entonces* $E^c \in \mathcal{C}$*.*

(iii) Si $\{E_i : i \in I\}$ *es una familia enumerable de elementos de* $\mathcal{C}$*, entonces* $\cup_{i\in I} E_i \in \mathcal{C}$*.*

Además, si $\mathcal{C}$ *es un* σ*-álgebra vale lo siguiente.*

- *Si* $\{E_i : i \in I\}$ *es una familia enumerable de elementos de* $\mathcal{C}$ *entonces* $\cap_{i\in I} E_i \in \mathcal{C}$*.*

Observación 1.7. *Nosotros hemos definido probabilidad como siendo una función definida sobre un* $\lambda - sistema$ $\mathcal{C}$*, cumpliendo con las propiedades (P1), (P2) y (P3). En otros textos, llaman probabilidad a* $\mathbf{P}$ *sólo en el caso especial de que* $\mathcal{C}$ *cumpla además con la propiedad* (π)*. Pero observe que ninguna propiedad adicional es requerida de la probabilidad* $\mathbf{P}$*.*

En este texto, el dominio de las probabilidades que veremos serán siempre σ-álgebras.

Definición 1.8. *Un espacio medible es una dupla $(\Omega, \mathcal{F})$ donde Ω es un conjunto no vacío y $\mathcal{F}$ es un σ-álgebra. Un espacio de probabilidad es un trio $(\Omega, \mathcal{F}, \mathbf{P})$ donde $(\Omega, \mathcal{F})$ es un espacio medible y $\mathbf{P} : \mathcal{F} \to [0, 1]$ es una probabilidad.*

Ejercicio 1.12. *Sea $(\Omega, \mathcal{F})$ un espacio medible. Pruebe que $\mathbf{P} : \mathcal{F} \to [0, 1]$ es probabilidad si y sólo si ella cumple con las siguientes propiedades:*

(*i*) $\mathbf{P}(\Omega) = 1$.

(*ii*) *Si $\{E_i : i \in I\}$ es una colección enumerable de elementos de $\mathcal{C}$, disjuntos dos a dos, entonces*

$$\mathbf{P}\left(\cup_{i\in I} E_i\right) \;=\; \sum_{i\in I} \mathbf{P}(E_i)\;.$$

Ejercicio 1.13. *Prueba que los cilindros en $S^{\mathbb{Z}_+}$ forman un π-sistema y NO un λ-sistema (por lo tanto tampoco un σ-álgebra).*

Ejercicio 1.14. *Verifique que las colecciones de subconjuntos de Ω,*

$$\big\{\varnothing\,,\;\Omega\big\} \quad y \quad 2^{\Omega} \;:=\; \{\,\textit{todos los subconjuntos de }\Omega\,\}$$

son ambos λ-sistemas y también σ-álgebras.

Ejercicio 1.15. *Prueba que si $\{\mathcal{C}_\alpha : \alpha \in I\}$ es una colección arbitraria de λ-sistemas (resp. σ-álgebras) en Ω entonces $\cap_{\alpha\in I}\mathcal{C}_\alpha$ es también un λ-sistema (resp. σ-álgebra) en Ω.*

Si $\mathcal{C}$ no es ni σ-álgebra ni λ-sistema, podemos agregarle más subconjuntos de Ω para tornarlo uno usando el siguiente procedimiento.

Definición 1.9. *Sea $\mathcal{C}$ una colección de subconjuntos de Ω. Denotemos*

$$\Lambda_{\mathcal{C}} \;:=\; \{\mathcal{D} : \mathcal{D} \textit{ es } \lambda\textit{-sistema en } \Omega \textit{ y } \mathcal{C} \subseteq \mathcal{D}\}$$
$$\textit{y} \quad \Sigma_{\mathcal{C}} \;:=\; \{\mathcal{D} : \mathcal{D} \textit{ es } \sigma\textit{-álgebra en } \Omega \textit{ y } \mathcal{C} \subseteq \mathcal{D}\}\;.$$

Entonces definimos

$$\lambda(\mathcal{C}) \;:=\; \bigcap_{\mathcal{D}\in\Lambda_{\mathcal{C}}} \mathcal{D} \quad y \quad \sigma(\mathcal{C}) \;:=\; \bigcap_{\mathcal{D}\in\Sigma_{\mathcal{C}}} \mathcal{D}\;.$$

Observación 1.10. *Esta definición no tendría sentido si* $\Lambda_{\mathcal{C}}$ *o* $\Sigma_{\mathcal{C}}$ *fueran vacíos. Afortunadamente, no lo son. Por qué?*

Gracias al Ejercicio 1.15, las nuevas colecciones de subconjuntos $\lambda(\mathcal{C})$ y $\sigma(\mathcal{C})$ definidas arriba son, respectivamente, un λ-sistema y un σ-álgebra. Además, por definición, si $\mathcal{C} \subseteq \mathcal{D}$ y $\mathcal{D}$ es un λ-sistema concluimos de inmediato que $\lambda(\mathcal{C}) \subseteq \mathcal{D}$. Igualmente, si $\mathcal{C} \subseteq \mathcal{D}$ y $\mathcal{D}$ es un σ-álgebra concluimos de inmediato que $\sigma(\mathcal{C}) \subseteq \mathcal{D}$. Resumimos estas afirmaciones diciendo

Observación 1.11. $\lambda(\mathcal{C})$ *es el menor* λ*-sistema conteniendo a* $\mathcal{C}$ *y* $\sigma(\mathcal{C})$ *es el menor* σ*-álgebra conteniendo a* $\mathcal{C}$*.*

Teorema 1.12. *Si* $\mathcal{C}$ *es un* π*-sistema entonces* $\sigma(\mathcal{C}) = \lambda(\mathcal{C})$*.*

Prueba: $\sigma(\mathcal{C})$ es un λ-sistema que contiene a $\mathcal{C}$. Pero $\lambda(\mathcal{C})$ es el menor λ-sistema conteniendo a $\mathcal{C}$ (ver la Observación 1.11), por tanto $\lambda(\mathcal{C}) \subseteq \sigma(\mathcal{C})$. Siguiendo el mismo argumento, para concluir la inclusión contraria $\sigma(\mathcal{C}) \subseteq \lambda(\mathcal{C})$ basta probar que $\lambda(\mathcal{C})$ es un σ-álgebra. Es decir, debemos probar que $\lambda(\mathcal{C})$ es un (π)-sistema. En este punto usamos la propiedad (π) de $\mathcal{C}$. Definamos,

$$\mathcal{G} \ := \ \{E \in \lambda(\mathcal{C}) : E \cap B \in \lambda(\mathcal{C}) \, , \text{ para todo } B \in \mathcal{C}\} \ .$$

Se sigue de inmediato que

$$\mathcal{C} \ \subseteq \ \mathcal{G} \ \subseteq \ \lambda(\mathcal{C}) \ .$$

Además, $\mathcal{G}$ resulta un λ-sistema (probar como ejercicio). Por lo tanto (nuevamente Observación 1.11) $\mathcal{G} = \lambda(\mathcal{C})$. Hasta aquí tenemos que $E \cap B \in \lambda(\mathcal{C})$ para todo par $E \in \lambda(\mathcal{C})$, $B \in \mathcal{C}$. Entonces, si definimos

$$\mathcal{H} \ := \ \{E \in \lambda(\mathcal{C}) : E \cap B \in \lambda(\mathcal{C}) \, , \text{ para todo } B \in \lambda(\mathcal{C})\} \ ,$$

podemos ahora afirmar que

$$\mathcal{C} \ \subseteq \ \mathcal{H} \ \subseteq \ \lambda(\mathcal{C}) \ .$$

Pero $\mathcal{H}$ es también λ-sistema, entonces $\mathcal{H} = \lambda(\mathcal{C})$. Eso significa que $\lambda(\mathcal{C})$ es un (π)-sistema y por tanto un σ-álgebra.

$\square$

Para terminar esta sección, veamos el siguiente resultado útil. Sea $\mathbf{P}$ una probabilidad definida sobre un λ-sistema $\mathcal{L}$. Consideremos una colección $\mathcal{C}$ contenida en $\mathcal{L}$. Nos gustaría saber cuan grande debe ser $\mathcal{C}$ para que los valores de $\mathbf{P}$ sobre $\mathcal{C}$ determinen los valores de $\mathbf{P}$ sobre todo $\mathcal{L}$. Probaremos que basta pedir que $\mathcal{L}$ sea generado por $\mathcal{C}$.

Proposición 1.13. *Sea $\mathcal{C}$ una colección de subconjuntos de Ω. Sean $\mathbf{P}$ y $\mathbf{Q}$ dos probabilidades definidas sobre $\lambda(\mathcal{C})$. Si $\mathbf{P}(C) = \mathbf{Q}(C)$ para todo $C \in \mathcal{C}$ entonces $\mathbf{P}(E) = \mathbf{Q}(E)$ para todo $E \in \lambda(\mathcal{C})$.*

Prueba: Definamos

$$\mathcal{D} \;:=\; \{E \in \lambda(\mathcal{C}) : \mathbf{P}(E) = \mathbf{Q}(E)\}\ .$$

Dos cosas obvias: $\mathcal{C} \subseteq \mathcal{D}$ y $\mathcal{D} \subseteq \lambda(\mathcal{C})$. Pruebe ahora como ejercicio que $\mathcal{D}$ es un λ-sistema (fácil). Pero como $\lambda(\mathcal{C})$ es el menor λ-sistema conteniendo $\mathcal{C}$ y $\mathcal{C} \subseteq \mathcal{D}$ entonces $\lambda(\mathcal{C}) \subseteq \mathcal{D}$. Por la doble inclusión concluimos que $\lambda(\mathcal{C}) = \mathcal{D}$. Es decir que cada elemento de $\lambda(\mathcal{C})$ satisface la propiedad que define a $\mathcal{D}$. Terminamos. □

El siguiente corolario es inmediato gracias al Teorema 1.12.

Corolario 1.14. *Sean $\mathbf{P}$ y $\mathbf{Q}$ dos probabilidades definidas sobre un σ-álgebra $\mathcal{F}$ y sea $\mathcal{C} \subseteq \mathcal{F}$ tal que $\sigma(\mathcal{C}) = \mathcal{F}$. Supongamos que $\mathbf{P}(C) = \mathbf{Q}(C)$, $\forall C \in \mathcal{C}$. Si $\mathcal{C}$ es un π-sistema, entonces $\mathbf{P}(E) = \mathbf{Q}(E)$, $\forall E \in \mathcal{F}$.*

Prueba: Sabemos por la proposición anterior que $\mathbf{P}$ y $\mathbf{Q}$ coinciden en todos los elementos de $\lambda(\mathcal{C})$. Pero por el Teorema 1.12, $\lambda(\mathcal{C}) = \sigma(\mathcal{C}) = \mathcal{F}$. Eso concluye la prueba. □

1.4 Probabilidades sobre $S^{\mathbb{Z}_+}$

Recuerde que S es un conjunto enumerable o finito con al menos dos elementos y $S^{\mathbb{Z}_+}$ denota el espacio de trayectorias sobre S. Denotemos por $\mathcal{C}il$ la colección de cilindros de $S^{\mathbb{Z}_+}$. Como fue visto en un ejercicio, $\mathcal{C}il$ forma un π-sistema. Entonces, por el Teorema 1.12, el menor σ-álgebra y el menor λ-sistema conteniendo $\mathcal{C}il$ coinciden y denotaremos a tal colección de subconjuntos por $\mathcal{B}(S^{\mathbb{Z}_+})$:

$$\mathcal{B}(S^{\mathbb{Z}_+}) \;:=\; \sigma(\mathcal{C}il) \;=\; \lambda(\mathcal{C}il)\ .$$

La cuestión que nos toca resolver es la siguiente. Supongamos que tenemos definida una función $\widetilde{\mathbb{P}} : \mathcal{C}il \to [0,1]$. No tendría sentido preguntarse si $\widetilde{\mathbb{P}}$ es probabilidad ya que $\mathcal{C}il$ no es siquiera λ-sistema. Pero, podría ser $\widetilde{\mathbb{P}}$ la restricción de una probabilidad definida sobre $\mathcal{B}(S^{\mathbb{Z}_+})$. Nos gustaría entonces conocer condiciones sobre $\widetilde{\mathbb{P}}$ que nos garanticen la existencia de dicha probabilidad.

Recuerde de un ejercicio que la unión de dos cilindros es todavía un cilindro. Diremos que $(C_n : n = 1, 2, \dots)$ es una sucesión decreciente de cilindros si $C_n \supset C_{n+1}\ \forall n \geq 1$.

El siguiente teorema es un caso especial del Teorema de Carathéodory.

Teorema 1.15. *Sea $\widetilde{\mathbb{P}} : \mathcal{C}il \to [0,1]$ una función tal que $\widetilde{\mathbb{P}}(\varnothing) = 0$, $\widetilde{\mathbb{P}}(S^{\mathbb{Z}_+}) = 1$. Supongamos que se cumplen las siguientes dos propiedades.*

(i) *Si C y D son cilindros disjuntos entonces $\widetilde{\mathbb{P}}(C \cup D) = \widetilde{\mathbb{P}}(C) + \widetilde{\mathbb{P}}(D)$.*

(ii) *Si $(C_n : n = 1, 2, \dots)$ es una sucesión decreciente de cilindros tales que $\cap_{n=1}^{\infty} C_n = \varnothing$, entonces $\widetilde{\mathbb{P}}(C_n) \to 0$.*

Entonces, existe una única probabilidad $\mathbb{P} : \mathcal{B}(S^{\mathbb{Z}_+}) \to [0,1]$ de modo que $\mathbb{P}(C) = \widetilde{\mathbb{P}}(C)$ para todo cilindro C.

Observe que las condiciones (i) y (ii) son de hecho necesarias para que $\widetilde{\mathbb{P}}$ sea la restricción de una probabilidad. En nuestro problema de definir cadenas de Markov, será más inmediato usar el Teorema 1.17 enunciado más abajo, consecuencia de este último teorema.

Definición 1.16. *Por cada $k \geq 1$ sea $\mu_k : \mathcal{P}(S^k) \to [0,1]$ una probabilidad. Diremos que la sucesión $(\mu_k : k \geq 1)$ es consistente si para todo $k \geq 1$ y subconjuntos $A_1, \dots, A_k \in S$ tenemos que*

$$\mu_{k+1}(A_1 \times \cdots \times A_k \times S) \;=\; \mu_k(A_1 \times \cdots \times A_k)\ .$$

El siguiente ejercicio justifica esta definición.

Ejercicio 1.16. *Sea $\mathbb{P} : \mathcal{B}(S^{\mathbb{Z}_+}) \to [0,1]$ una probabilidad.*

- *Si* $\mu : \mathcal{P}(S) \to [0,1]$ *es definido como* $\mu(A) := \mathbb{P}\{X_0 \in A\}$, $\forall A \subseteq S$, *verifique que* μ *es una probabilidad.*
- *Si por cada* $k \geq 1$, $\mu_k : \mathcal{P}(S^k) \to [0,1]$ *es definido como*

  $$\mu_k(E) \ := \ \mathbb{P}\{(X_0, \dots, X_{k-1}) \in E\} \ , \quad \forall E \subseteq S^k \ ,$$

  entonces μ_k *es una probabilidad.*
- *Si* $(\mu_k : k = 1, 2, \dots)$ *es la sucesión de probabilidades definidas en el item anterior, pruebe que dicha sucesión es consistente.*

Ejercicio 1.17. *Por cada* $k \geq 1$ *sea* $\mu_k : \mathcal{P}(S^k) \to [0,1]$ *una probabilidad. Supongamos que* $(\mu_k : k = 1, 2, \dots)$ *es una sucesión consistente. Si* $1 \leq n < m$ *y* $A \subseteq S^n$ *entonces*

$$\mu_m(A \times \underbrace{S \times \cdots \times S}_{\textit{(m-n) veces}}) \ = \ \mu_n(A) \ .$$

Teorema 1.17. *Por cada* $k \geq 1$ *sea* $\mu_k : \mathcal{P}(S^k) \to [0,1]$ *una probabilidad. Si* $(\mu_k : k = 1, 2, \dots)$ *es una sucesión consistente, entonces existe una única probabilidad* $\mathbb{P} : \mathcal{B}(S^{\mathbb{Z}_+}) \to [0,1]$ *tal que*

$$\mathbb{P}\{(X_0, X_1, \dots, X_{k-1}) \in E\} \ = \ \mu_k(E) \ ,$$

$$\textit{para todo } k \geq 1 \textit{ y todo } E \in S^k \ .$$

Prueba: Definamos una función $\widetilde{\mathbb{P}} : \mathcal{C}il \to [0,1]$ del siguiente modo. Si C es un cilindro entonces existe $k \geq 1$ y $A \subseteq S^k$ tal que $C = \{(X_0, \dots, X_{k-1}) \in A\}$. Luego, hacemos $\widetilde{\mathbb{P}}(C) = \mu_k(A)$. Hay que verificar que $\widetilde{\mathbb{P}}$ está bien definida. Supongamos que existen $n \geq 1$, $A \subseteq S^n$, $m \geq 1$ y $B \subseteq S^m$ de modo que

$$C = \{(X_0, \dots, X_{n-1}) \in A\} = \{(X_0, \dots, X_{m-1}) \in B\} \ .$$

Sin pérdida de generalidad podemos suponer que $k \leq m$. Si $k = m$, entonces $A = B$ y obviamente $\mu_k(A) = \mu_m(B)$. Si $k < m$, por el ejercicio 1.2 tendremos que

$$\mu_m(B) = \mu_m(A \times \underbrace{S \times \cdots \times S}_{\text{(m-n) veces}}) \ = \ \mu_n(A).$$

En la última igualdad hemos usado el Ejercicio 1.17. Eso prueba que $\widetilde{\mathbb{P}}$ está bien definida. Ahora resta probar las condiciones del Teorema 1.15.

Para probar (i) fijemos dos cilindros E y F disjuntos. Por la parte (a) del Ejercicio 1.4 sabemos que existe $k \geq 1$ y subconjuntos $A, B \subseteq S^k$ de modo que $E = \{(X_0, \dots, X_{k-1}) \in A\}$ y $F = \{(X_0, \dots, X_{k-1}) \in B\}$. Ya que

$$E \cup F \;=\; \{(X_0, \dots, X_{k-1}) \in A \cup B\}$$

tendremos

$$\widetilde{\mathbb{P}}(E \cup F) \;=\; \mu_k(A \cup B) \;=\; \mu_k(A) + \mu_k(B) \;=\; \widetilde{\mathbb{P}}(E) + \widetilde{\mathbb{P}}(F)\;.$$

La segunda igualdad sigue de ser μ_k una probabilidad y la última igualdad sigue de la definición de $\widetilde{\mathbb{P}}$.

Para probar (ii), procedemos por contradicción. Sea $(C_n : n = 1, 2, \dots)$ una sucesión de cilindros tal que $C_n \supset C_{n+1}$, $\forall n \geq 1$ y supongamos que existe $\delta > 0$ tal que $\widetilde{\mathbb{P}}(C_n) \geq \delta > 0$, $\forall n \geq 1$. Vamos a probar que $\cap_{n \geq 1} C_n \neq \varnothing$.

Ya que C_1 es un cilindro, existe un entero $k_1 \geq 1$ y un subconjunto $A_1' \subseteq S^{k_1}$ de modo que $C_1 := (X_0, \dots, X_{k_1-1})^{-1}(A_1')$. Por el Ejercicio 1.10, existe un subconjunto finito $A_1 \subseteq A_1'$ de modo que

$$\widetilde{\mathbb{P}}(C_1 \cap K_1^c) \;\leq\; \delta/3\;, \tag{1.4.1}$$

donde $K_1 := (X_0, \dots, X_{k_1-1})^{-1}(A_1) \subseteq C_1$.

Notemos que $K_1 \cap C_2$ es un cilindro. Por la parte (a) del Ejercicio 1.4, existe $k_2 \geq k_1$ y $A_2' \subseteq S^{k_2}$ de modo que $K_1 \cap C_2 = (X_0, \dots, X_{k_2-1})^{-1}(A_2')$. Nuevamente por el Ejercicio 1.10, existe un subconjunto finito $A_2 \subseteq A_2'$ tal que

$$\widetilde{\mathbb{P}}(K_1 \cap C_2 \cap K_2^c) \;\leq\; (\delta/3)^2\;, \tag{1.4.2}$$

donde $K_2 := (X_0, \dots, X_{k_2-1})^{-1}(A_2) \subseteq K_1 \cap C_2$. Siendo $C_2 \subseteq C_1$, tenemos

$$\begin{aligned}\widetilde{\mathbb{P}}(C_2 \cap K_2^c) \;&=\; \widetilde{\mathbb{P}}(K_1 \cap C_2 \cap K_2^c) + \widetilde{\mathbb{P}}(K_1^c \cap C_2)\\ &\leq\; \widetilde{\mathbb{P}}(K_1 \cap C_2 \cap K_2^c) + \widetilde{\mathbb{P}}(K_1^c \cap C_1)\;.\end{aligned}$$

De (1.4.1) y (1.4.2) concluimos que K_2 satisface

$$\widetilde{\mathbb{P}}(C_2 \cap K_2^c) \ \leq \ (\delta/3) + (\delta/3)^2 \ .$$

Procediendo por inducción, probamos que existe una sucesión de enteros $1 \leq k_1 \leq k_2 \leq \dots$, una sucesión de conjuntos finitos $A_n \subseteq S^{k_n}$ y una sucesión decreciente de cilindros $(K_n : n = 1, 2, \dots)$ de modo que $K_n := (X_0, \dots, X_{k_n - 1})^{-1}(A_n)$, $K_n \subseteq C_n$ y

$$\widetilde{\mathbb{P}}(C_n \cap K_n^c) \ \leq \ \sum_{i=1}^{n} (\delta/3)^i \ , \quad \forall n \geq 1 \ .$$

De la última desigualdad y la suposición sobre $(C_n : n = 1, 2, \dots)$ tenemos que para todo $n \geq 1$

$$\begin{aligned} \widetilde{\mathbb{P}}(K_n) \ = \ \widetilde{\mathbb{P}}(C_n) - \widetilde{\mathbb{P}}(C_n \cap K_n^c) \ &\geq \ \widetilde{\mathbb{P}}(C_n) - \sum_{i=1}^{n} (\delta/3)^i \\ &\geq \ \delta - \sum_{i=1}^{\infty} (\delta/3)^i \ = \ \delta/2 \ . \end{aligned}$$

En particular, $K_n \neq \varnothing$ y por tanto $A_n \neq \varnothing$, para todo $n \geq 1$.

Entonces es posible tomar una sucesión $\vec{x}_n \in K_n$, $n \geq 1$. Por cada $n \geq 1$, denotemos

$$\pi_n \equiv (X_0, \dots, X_{k_n - 1}) \ .$$

De este modo $\pi_1(\vec{x}_n) \in A_1$, $\forall n \geq 1$. Ya que A_1 es finito, entonces existe un $a^{(1)} \in A_1$ y una subsucesión $(\vec{x}_n^{(1)} : n \geq 1)$ tal que $\pi_1(\vec{x}_n^{(1)}) = a^{(1)}$ para todo $n \geq 1$. Ahora, $\pi_2(\vec{x}_n^{(1)}) \in A_2$ para todo $n \geq 1$. Siendo A_2 finito, existe un $a^{(2)} \in A_2$ y una subsucesión $(\vec{x}_n^{(2)} : n \geq 1)$ de $(\vec{x}_n^{(1)} : n \geq 1)$ de modo que $\pi_2(\vec{x}_n^{(2)}) = a^{(2)}$ para todo $n \geq 1$. Lo importante es notar que $\pi_1(\vec{x}_n^{(2)}) = a^{(1)}$ y por lo tanto

$$\pi_1'(a^{(2)}) \ = \ a^{(1)} \ ,$$

donde $\pi_1' : S^{k_2} \to S^{k_1}$ es la proyección en las k_1 primeras coordenadas. Tomando progresivamente subsucesiones podemos entonces obtener una sucesión $a^{(n)} \in A_n$ cumpliendo con

$$\pi_n'(a^{(n+1)}) \ = \ a^{(n)} \ , \quad \forall n \geq 1 \ ,$$

donde $\pi_n' : S^{k_{n+1}} \to S^{k_n}$ es la proyección en las k_n primeras coordenadas. Esto es suficiente para garantizar que existe un elemento $\vec{a} \in S^{\mathbb{Z}_+}$ tal que $\pi_n(\vec{a}) = a^{(n)}$ para todo $n \geq 1$. Ahora es fácil verificar que $\vec{a} \in K_n$ para todo $n \geq 1$. Eso termina la prueba. $\square$

Capítulo 2

Cadena de Markov

Fijaremos un conjunto S enumerable o finito con al menos dos elementos. Consideremos una colección de números $p(x, y) \in [0, 1]$, $(x, y) \in S \times S$. En este capítulo, definiremos una cadena de Markov como una familia de probabilidades sobre el espacio de trayectorias en S y asociados a $\{p(x, y) : x, y \in S\}$ del modo que describimos más abajo.

Definición 2.1. *Considere una colección de números* $p(x, y) \in [0, 1]$, $(x, y) \in S \times S$. *Decimos que*

$$\{p(x, y) : x, y \in S\}$$

forma una colección de probabilidades de transición sobre S *si*

$$\sum_{y \in S} p(x, y) \; = \; 1 \; , \quad \forall x \in S \; . \tag{2.0.1}$$

Usamos $\{p(x, y) : x, y \in S\}$ para determinar una dinámica sobre S del siguiente modo. En el instante cero situamos a un individuo en algún lugar en S, digamos $a \in S$. Viéndose en a, el individuo decidirá una nueva posición en S de acuerdo a las probabilidades $p(a, \cdot)$. Desde la nueva posición, digamos $b \in S$, decidirá su siguiente posición (para el instante dos) de acuerdo a las probabilidades $p(b, \cdot)$. El individuo continuará de este modo decidiendo su posición para el instante $k+1$ de acuerdo a su posición x en el instante k y las probabilidades $p(x,\cdot)$. Observe entonces que, luego de fijar un punto inicial, un elemento de $S^{\mathbb{Z}_+}$ es producido aleatoriamente bajo este algoritmo determinado por los números $\{p(x,y) : x,y \in S\}$.

Con esta descripción podemos calcular las siguientes probabilidades.

Ejercicio 2.1. *Supongamos que la posición inicial del individuo es $a \in S$.*

(a) *¿ Cuál es la probabilidad de que el individuo salte al punto $b \in S$ y luego al punto $c \in S$?*
Rpta: $p(a,b)p(b,c)$.

(b) *¿ Cuál es la probabilidad de que el individuo se encuentre en $c \in S$ luego de dos saltos?*
Rpta: $\sum_{y\in S} p(a,y)p(y,c)$.

(c) *¿ Cuál es la probabilidad de que el individuo vuelva al punto $a \in S$ luego de 3 saltos y no antes?*
Rpta: $\sum_{y\in S\setminus\{a\}} \sum_{z\in S\setminus\{a\}} p(a,y)p(y,z)p(z,a)$.

Todos los subconjuntos en el Ejercicio 2.1 son cilindros! La descripción hecha sobre la dinámica nos permite determinar directamente, a partir de las probabilidades de transición $\{p(x,y) : x,y \in S\}$, una familia de funciones sobre los cilindros. Usaremos el Teorema 1.17 para determinar una probabilidad sobre $(S^{\mathbb{Z}_+}, \mathcal{B}(S^{\mathbb{Z}_+}))$ a partir de las probabilidades de transición.

Sea ν una probabilidad sobre S. Para cada $k \geq 2$, determinamos una probabilidad μ_k sobre S^k estableciendo

$$\mu_k(a) \;=\; \nu(a_0)p(a_0,a_1)\dots p(a_{k-2},a_{k-1})\ ,$$
$$\forall a = (a_0,a_1,\dots,a_{k-1}) \in S^k\ .$$

Así tenemos la familia $\{\mu_k : k = 1,2,\dots\}$ donde $\mu_1 \equiv \nu$.

Ejercicio 2.2. *La familia $\{\mu_k : k = 1,2,\dots\}$ definida arriba es consistente.*

Usando entonces el Teorema 1.17 podemos concluir que

Teorema 2.2. *Dadas las probabilidades de transición* $\{p(x,y) : x, y \in S\}$ *y la probabilidad* ν *sobre* S*, existe una única probabilidad* $\mathbb{P}_\nu : \mathcal{B}(S^{\mathbb{Z}_+}) \to [0,1]$ *tal que*

- $\mathbb{P}_\nu\{X_0 \in a\} = \nu(a)$ *para todo* $a \in S$.
- $\mathbb{P}_\nu\{X_0{=}a_0; X_1{=}a_1; \dots ; X_k{=}a_k\}{=}\nu(a_0)p(a_0, a_1) \dots p(a_{k-1}, a_k)$ *para todo* $k \geq 1$ *y* $(a_0, \dots , a_k){\in}S^{k+1}$.

Estamos escribiendo $\nu(a) := \nu(\{a\})$.

Definición 2.3. *Llamaremos cadena de Markov a cada probabilidad* $\mathbb{P}_\nu$ *sobre* $(S^{\mathbb{Z}_+}, \mathcal{B}(S^{\mathbb{Z}_+}))$*. Los ingredientes* S*,* $\mathcal{B}(S^{\mathbb{Z}_+})$*,* $\{p(x,y) : x, y \in S\}$ *y* ν *serán llamados, respectivamente, el espacio de estados, el espacio de trayectorias, las probabilidades de transición y la distribución inicial de la cadena de Markov* $\mathbb{P}_\nu$*.*

Cuando existe $a \in S$ tal que la distribución inicial $\nu(a) = 1$, denotaremos a la cadena de Markov correspondiente por $\mathbb{P}_a$.

Ejercicio 2.3. *Pruebe que para todo* $E \in \mathcal{B}(S^{\mathbb{Z}_+})$ *y toda probabilidad* ν *sobre* S *tenemos*

$$\mathbb{P}_\nu(E) \;=\; \sum_{x \in S} \nu(x)\mathbb{P}_x(E) \; .$$

Ejercicio 2.4. *Sea* $n \geq 1$ *y* $A \subseteq S$ *no vacío.*

1. *Pruebe que* $\{T_A = n\}$ *y* $\{T_A > n\}$ *son cilindros.*
2. *Para un* $a \in S$*, calcule* $\mathbb{P}_a\{T_A = n\}$ *y* $\mathbb{P}_a\{T_A > n\}$ *en función de las probabilidades de transición* $\{p(x,y) : x, y \in S\}$*.*

Lema 2.4. *Para todo* $A \subset S$ *no vacío,* $\{T_A = \infty\} \in \mathcal{B}(S^{\mathbb{Z}_+})$*. Además, si* A, B *son subconjuntos disjuntos no vacíos de* S *entonces* $\{T_A < T_B\} \in \mathcal{B}(S^{\mathbb{Z}_+})$*.*

Prueba: Primero, podemos escribir $\{T_A < \infty\} = \cup_{n=1}^{\infty}\{T_A = n\}$. Entonces, por el Ejercicio 2.4, concluimos que $\{T_A < \infty\} \in \mathcal{B}(S^{\mathbb{Z}_+})$. Eso prueba que $\{T_A = \infty\} \in \mathcal{B}(S^{\mathbb{Z}_+})$. Ahora, el conjunto $\{T_A < T_B\}$ se puede escribir como

$$\cup_{n=1}^{\infty}\big(\{T_A < T_B\} \cap \{T_A = n\}\big) \;=\; \cup_{n=1}^{\infty}\big(\{n < T_B\} \cap \{T_A = n\}\big) \; .$$

Eso acaba la prueba ya que $\{n < T_B\}$ y $\{T_A = n\}$ son cilindros para todo $n \geq 1$. $\square$

Podemos calcular las siguiente probabilidades? Supongamos que $a \in S$ es la posición inicial.

(A) La probabilidad de que el individuo llegue algún día a $b \in S$ y que suceda antes de llegar a $c \in S$.

(B) La probabilidad de salir de a y nunca más regresar.

Las probabilidades en (A) y (B) se pueden escribir, respectivamente, como $\mathbb{P}_a\{T_b < T_c\}$ y $\mathbb{P}_a\{T_a = \infty\}$.

2.1 Propiedad de Markov

Para calcular la probabilidad de eventos más complicados como los enunciados al final de la sección anterior, enunciaremos los Teoremas 2.6, 2.7, 2.12 y 2.13. Estos teoremas son asociados a la propiedades markovianas de la dinámica que hemos definido. Entre otras aplicaciones, veremos como estos teoremas nos permiten escribir las probabilidades que queremos calcular como soluciones de un sistema de ecuaciones.

Vamos a denotar

$$p^{(n)}(a,x) \ := \ \mathbb{P}_a\{X_n = x\}\ ,$$

de modo que $p^{(1)}(a,x) = p(a,x)$.

Ejercicio 2.5. *Exprese* $p^{(n)}(a,x)$ *en términos de* $\{p(x,y) : x,y \in S\}$.

Definición 2.5. *Para cada* $n \geq 1$*, el operador* $\Theta^{(n)} : S^{\mathbb{Z}_+} \to S^{\mathbb{Z}_+}$ *es definido como*

$$\Theta^{(n)}(\vec{x}) \ = \ (x_n, x_{n+1}, \dots)\ , \quad \forall \vec{x} = (x_0, x_1, x_2, \dots) \in S^{\mathbb{Z}_+}\ .$$

Además escribimos $\Theta := \Theta^{(1)}$.

Teorema 2.6. *Sea* $\mathbb{P}_\nu$ *una cadena de Markov. Para todo* $E \in \mathcal{B}(S^{\mathbb{Z}_+})$ *vale*

$$\mathbb{P}_\nu\{\Theta^{(n)} \in E\} \ = \ \sum_{x \in S} \mathbb{P}_\nu(X_n = x)\, \mathbb{P}_x(E)\ , \tag{2.1.1}$$

En particular, para todo $a, b \in S$ tenemos

$$\mathbb{P}_a\{X_n = b\,;\, \Theta^{(n)} \in E\} \;=\; p^{(n)}(a,b)\,\mathbb{P}_b(E)\;. \tag{2.1.2}$$

Ejercicio 2.6. • *Pruebe que $\mathbb{Q}(E) := \mathbb{P}_\nu\{\Theta^{(n)} \in E\}$, $\forall E \in \mathcal{B}(S^{\mathbb{Z}_+})$ define una probabilidad*

- *Pruebe que $\mathbb{Q}'(E) := \sum_{x\in S}\mathbb{P}_\nu(X_n = x)\,\mathbb{P}_x(E)$, $\forall E \in \mathcal{B}(S^{\mathbb{Z}_+})$ define una probabilidad.*
- *Pruebe que $\mathbb{Q}(C) = \mathbb{Q}'(C)$ para todo C cilindro. Concluya el resultado del Teorema 2.6.*

Podemos mejorar el teorema anterior del siguiente modo. Para simplificar la notación, denotemos por

$$\pi_k \equiv (X_0, X_1, \dots, X_k)\;, \quad \forall k \geq 0\;.$$

El siguiente teorema nos dice que, fijada la posición del presente, el pasado y el futuro resultan independientes.

Teorema 2.7. *Sea $k \geq 1$, $E \in \mathcal{B}(S^{\mathbb{Z}_+})$ y $A \subseteq S^{k+1}$. Tenemos*

$$\mathbb{P}_\nu\big\{\pi_k \in A\,;\, X_k = x\,;\, \Theta^{(k)} \in E\big\} \;=\; \mathbb{P}_\nu\big\{\pi_k \in A\,;\, X_k = x\big\}\,\mathbb{P}_x(E)\;.$$

Prueba: Por el Ejercicio 2.3, basta mostrar el enunciado para $\mathbb{P}_\nu = \mathbb{P}_a$. Dividiendo A como unión enumerable de sus elementos, podemos suponer que $A = \{(b_0, \dots, b_k)\}$. Para descartar casos triviales, podemos suponer que $b_k = x$ y $b_0 = a$. Basta probar entonces que

$$\begin{aligned} &\mathbb{P}_{b_0}\big\{X_1 = b_1, \dots, X_k = b_k\,;\, \Theta^{(k)} \in E\big\} \\ &=\; \mathbb{P}_{b_0}\big\{X_1 = b_1, \dots, X_k = b_k\big\}\,\mathbb{P}_{b_k}(E)\;. \end{aligned}$$

Si $\mathbb{P}_{b_0}\big\{X_1 = b_1, \dots, X_k = b_k\big\} = 0$ no hay nada más que hacer. Caso contrario

$$E \mapsto \mathbb{P}_{b_0}\big\{X_1{=}b_1, \dots, X_k{=}b_k\,;\, \Theta^{(k)}{\in}E\big\}/\mathbb{P}_{b_0}\big\{X_1{=}b_1, \dots, X_k{=}b_k\big\}\;,$$

define una probabilidad sobre $\mathcal{B}(S^{\mathbb{Z}_+})$. Es fácil verificar que esta probabilidad coincide con $\mathbb{P}_{b_k}$ sobre los cilindros. Por tanto ellos deben ser iguales, probando lo que queríamos. $\square$

Ejercicio 2.7. *Sean A, B dos subconjuntos disjuntos no vacíos de S. Si $E = \{T_B < T_A\}$ pruebe que*

- $E \cap \{X_1 = b\} = \{X_1 = b\}$ *para todo* $b \in B$.
- $E \cap \{X_1 = a\} = \varnothing$ *para todo* $a \in A$.
- $E \cap \{X_1 = y\} = \{\Theta \in E\} \cap \{X_1 = y\}$, *para todo* $y \notin A \cup B$.

Lema 2.8. *Si definimos $f : S \to [0,1]$ como $f(x) = \mathbb{P}_x\{T_B < T_A\}$ para cada $x \notin B \cup A$ y $1 - f(a) = f(b) = 1$ para todo $a \in A$ y $b \in B$, entonces*

$$\mathbb{P}_x\{T_B < T_A\} \;=\; \sum_{y \in S} p(x,y) f(y) \;, \quad \forall x \in S \;.$$

En particular, $\sum_{y \in S} p(x,y) f(y) = f(x)$ para todo $x \notin A \cup B$.

Prueba: Si $E = \{T_B < T_A\}$ observemos que

$$\begin{aligned} \mathbb{P}_x(E) \;&=\; \sum_{y \in S} \mathbb{P}_x(E \cap \{X_1 = y\}) \\ &=\; \sum_{y \in B} \mathbb{P}_x\{X_1 = y\} + \sum_{y \in \Delta} \mathbb{P}_x(\{\Theta \in E\} \cap \{X_1 = y\}) \;, \end{aligned}$$

donde $\Delta = S \setminus (A \cup B)$. La última expresión se escribe como

$$\sum_{y \in B} p(x,y) f(y) + \sum_{y \in \Delta} \mathbb{P}_x(\Theta \in E \cap \{X_0 = y\}) \;.$$

Usando el Teorema 2.7, esta última expresión coincide con

$$\begin{aligned} &\sum_{y \in B} p(x,y) f(y) + \sum_{y \in \Delta} \sum_{a \in S} p(x,a) \mathbb{P}_a(E \cap \{X_0 = y\}) \\ &\qquad\qquad = \sum_{y \in B} p(x,y) f(y) + \sum_{y \in \Delta} p(x,y) \mathbb{P}_y(E) \;. \end{aligned}$$

Eso concluye la prueba. $\square$

Este resultado inspira la siguiente definición

Definición 2.9. *Asociamos a las probabilidades de transición* $\{p(x, y) : x, y \in S\}$ *el siguiente operador. A cada* $f : S \to \mathbb{R}$ *limitado,* $Mf : S \to \mathbb{R}$ *es definido como*

$$Mf(x) \ := \ \sum_{y \in S} p(x, y) f(y) \ .$$

Usaremos el Lema 2.8 en el próximo capítulo, para expresar en términos de cadenas de Markov la solución de un sistema de ecuaciones.

La probabilidad $\mathbb{P}_a\{T_a < \infty\}$ de volver algún día resulta ser uno cuando el conjunto de estados S es finito y la familia de probabilidades de transición es irreductible. Si S no es finito, esto no es más verdad en general, es decir, hay probabilidad positiva de nunca volver (incluso en el caso irreductible). Definimos a continuación irreductibilidad.

Ejercicio 2.8. *Sea* $\{p(x, y) : x, y \in S\}$ *una familia de probabilidades de transición. Fijemos* $x, y \in S$ *tal que* $x \neq y$*. Son equivalentes:*

(a) Existe $n \geq 1$ *tal que* $p^{(n)}(x, y) > 0$*.*

(b) $\mathbb{P}_x(T_y < \infty) > 0$*.*

Notación: Dada una cadena de Markov y dos elementos $x, y \in S$ distintos, escribiremos $x \to y$ para decir que x, y cumplen con el item (a) (o equivalentemente item (b)) del Ejercicio 2.8.

Definición 2.10. *Diremos que una familia de probabilidades de transición* $\{p(x, y) : x, y \in S\}$ *es irreductible si* $x \to y$ *para todo par* $x, y \in S$*,* $x \neq y$*.*

Ejercicio 2.9. *Sea* A *un subconjunto no vacío de* S*. Pruebe que si* $g : S \to [0, 1]$ *es definido como* $g(x) = \mathbb{P}_x\{T_A < \infty\}$ *para* $x \notin A$ *y* $g(a) = 1$ *para todo* $a \in A$*, entonces*

$$\mathbb{P}_x\{T_A < \infty\} \ = \ Mg(x) \ , \quad \forall x \in S \ .$$

Usando este ejercicio podemos resolver el siguiente que prueba en particular que $\mathbb{P}_a\{T_a < \infty\} = 1$ para S finito en el caso irreductible.

Ejercicio 2.10. *Sea* $\{p(x, y) : x, y \in S\}$ *una familia de probabilidades de transición irreductible. Pruebe que si* S *es finito entonces* $\mathbb{P}_x\{T_A < \infty\} = 1$ *para todo* $x \in S$ *y todo subconjunto no vacío* $A \subseteq S$*.*

2.2 Propiedad fuerte de Markov

Extenderemos ahora el Teorema 2.7 al caso en que el futuro y pasado es determinado por tiempos aleatorios.

Fijado un subconjunto $A \subseteq S$ no vacío, definamos el siguiente operador sobre el subconjunto $\{T_A < \infty\} \subset S^{\mathbb{Z}_+}$.

Definición 2.11. *Definimos* $\Theta_{T_A} : \{T_A < \infty\} \to S^{\mathbb{Z}_+}$ *como*

$$\Theta_{T_A}(\vec{x}) \;:=\; \Theta_{T_A(\vec{x})}(\vec{x}) \; .$$

El siguiente resultado es el análogo al Teorema 2.6 para tiempos aleatorios del tipo T_a.

Teorema 2.12. *Para todo* $a \in S$ *y para toda* ν *probabilidad sobre* S *tenemos*

$$\mathbb{P}_\nu(\{T_a < \infty\} \cap \{\Theta_{T_a} \in E\}) \;=\; \mathbb{P}_\nu\{T_a < \infty\}\mathbb{P}_a(E)$$

para todo $E \in \mathcal{B}(S^{\mathbb{Z}_+})$.

Prueba: Ejercicio. Divida $\{T_a < \infty\} = \cup_{n\geq 1}\{T_a = n\}$ y use el Teorema 2.7. □

Podemos extender el resultado anterior del siguiente modo. Para cada $k \geq 0$, denotemos por $\mathcal{F}_k \subset \mathcal{B}(S^{\mathbb{Z}_+})$ la colección de todos los conjuntos de la forma $\{(X_0, \ldots, X_k) \in A\}$ para algún $A \in S^{k+1}$. Observe que $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots$

Teorema 2.13. *Sea* $a \in S$ *y sea* ν *probabilidad sobre* S*. Supongamos que* $L \in \mathcal{B}(S^{\mathbb{Z}_+})$ *es tal que* $L \cap \{T_a = k\} \in \mathcal{F}_k$ *para todo* $k \geq 1$ *. Entonces*

$$\mathbb{P}_\nu(L \cap \{T_a < \infty\} \cap \{\Theta_{T_a} \in E\}) \;=\; \mathbb{P}_\nu(L \cap \{T_a < \infty\})\mathbb{P}_a(E)$$

para todo $E \in \mathcal{B}(S^{\mathbb{Z}_+})$.

Prueba El lado izquierdo se escribe como

$$\sum_{k=1}^{\infty} \mathbb{P}_\nu(L \cap \{T_a = k\} \cap \{\Theta_k \in E\}) \; .$$

Como $L \cap \{T_a = k\} \in \mathcal{F}_k$ entonces, usando el Teorema 2.7, esta expresión es igual a

$$\sum_{k=1}^{\infty} \mathbb{P}_\nu(L \cap \{T_a = k\})\, \mathbb{P}_a(E)\ .$$

Eso prueba el enunciado. □

Como ejemplo de L en este teorema, pruebe el siguiente ejercicio.

Ejercicio 2.11. *Sea $A \subseteq S$ un subconjunto no vacío y $a \in A$. Considere $L = \{T_A = T_a\}$. Pruebe que $L \cap \{T_a = k\} \in \mathcal{F}_k$ para todo $k \geq 1$.*

Usaremos el caso de este ejercicio en la siguiente sección.

2.3 La traza

Una familia de probabilidades de transición $\{p(x,y) : x, y \in S\}$ determina entonces una dinámica sobre S. Fijado un subconjunto no vacío $A \subset S$, vamos a restringir la dinámica al conjunto A. Esto será la traza de $\{p(x,y) : x, y \in S\}$ sobre A.

Para esta sección vamos a suponer que S es finito y que las probabilidades de transición $\{p(x,y) : x, y \in S\}$ es irreductible como fue definido en la sección anterior. Definiremos la traza de manera inductiva.

Fijemos un punto $z \in S$.

Definición 2.14. *Llamaremos traza de $\{p(x,y) : x, y \in S\}$ sobre $S \setminus \{z\}$ a la familia de probabilidades de transición $\{p'(x,y) : x, y \in S \setminus \{z\}\}$ definido por*

$$p'(x,y) \;=\; p(x,y) + \frac{p(x,z)p(z,y)}{1 - p(z,z)}\ , \quad \forall x, y \in S \setminus \{z\}\ . \tag{2.3.1}$$

Ejercicio 2.12. *Pruebe que $\{p'(x,y) : x, y \in S \setminus \{z\}\}$ es una familia irreductible de probabilidades de transición.*

Ahora, fijado un nuevo punto $w \in S \setminus \{z\}$, podemos tomar la traza de $\{p'(x,y) : x,y \in S \setminus \{z\}\}$ sobre $S \setminus \{z,w\}$. De ese modo obtenemos otra familia irreductible de probabilidades de transición $\{p''(x,y) : x,y \in S \setminus \{z,w\}\}$ definida ahora sobre el conjunto reducido de estados $S \setminus \{z,w\}$.

Iterando este método podemos retirar de la dinámica un número finito de puntos en S.

Definición 2.15. *Dado un subconjunto $A \subset S$ no vacío y una familia irreductible $\{p(x,y) : x,y \in S\}$ de probabilidades de transición, la traza de $\{p(x,y) : x,y \in S\}$ sobre A es la familia de probabilidades de transición $\{\tilde{p}(x,y) : x,y \in A\}$ obtenido después de retirar cada punto en $S \setminus A$ vía ecuación (2.3.1).*

Ya hemos observado que la traza $\{\tilde{p}(x,y) : x,y \in A\}$ resulta una familia irreductible de probabilidades de transición como consecuencia de la irreductibilidad de $\{p(x,y) : x,y \in A\}$.

Siendo $\{\tilde{p}(x,y) : x,y \in A\}$ una nueva familia de probabilidades de transición ella define una familia de cadenas de Markov con espacio de estados A. Vamos a mostrar la relación entre las cadenas asociadas a $\{p(x,y) : x,y \in S\}$ y las cadenas asociadas a su traza sobre A.

El siguiente ejercicio es un paso inicial. Este ejercicio además explica la intención que hubo en la forma de escribir la ecuación (2.3.1).

Ejercicio 2.13. *Pruebe que si $\{\tilde{p}(x,y) : x,y \in A\}$ es la traza de $\{p(x,y) : x,y \in S\}$ sobre A, entonces*

$$\mathbb{P}_a\{T_A = T_x\} \;=\; \tilde{p}(a,x) \;, \quad \forall x,a \in A \;,$$

donde $\mathbb{P}_a$ es la cadena de Markov asociada a $\{p(x,y) : x,y \in S\}$.

Sea $A \subseteq S$ no vacío. Recordemos que

$$T_A \;:=\; \inf\{n \geq 1 : X_n \in A\} \;.$$

Vamos a definir inductivamente la sucesión $\{T_A^{(k)} : k \geq 1\}$. Primero $T_A^{(1)} = T_A$. Luego, definido $T_A^{(k)}$, definimos

$$T_A^{(k+1)} \;:=\; T_A \circ \Theta_{T_A^{(k)}} \;, \quad \text{sobre } \{T_A^{(k)} < \infty\}$$

y $T_A^{(k+1)} \equiv \infty$ sobre $\{T_A^{(k)} = \infty\}$.

En el caso particular de esta sección, en que el espacio de estados S es finito y $p(\cdot,\cdot)$ es irreductible tenemos que todos estos tiempos son finitos casi ciertamente, para todas las cadenas asociadas a $p(\cdot,\cdot)$. Es el siguiente ejercicio.

Ejercicio 2.14. (*a*) *Pruebe que* $\{T_A^{(k)} < \infty$ *para todo* $k \geq 1\} \in \mathcal{B}(S^{\mathbb{Z}_+})$.

(*b*) *Pruebe que* $\forall x \in S$ *y para todo subconjunto* $A \subseteq S$,

$$\mathbb{P}_x\{T_A^{(k)} < \infty \textit{ para todo } k \geq 1\} \;=\; 1 \;.$$

Denotemos

$$H \;:=\; \{T_A^{(k)} < \infty \text{ para todo } k \geq 1\} \;.$$

sobre H podemos entonces definir la sucesión

$$Y_k(\vec{x}) \;:=\; X_{T_A^{(k)}(\vec{x})}(\vec{x}) \;, \quad \forall k \geq 1 \;.$$

Es decir $(Y_k : k \geq 1)$ es definida como la traza de $(X_n : n \geq 1)$ sobre A. Hagamos $Y_0 \equiv X_0$. Tenemos definida así una función $Y : H \to S^{\mathbb{Z}_+}$,

$$Y(\vec{x}) \;:=\; (Y_0(\vec{x}), Y_1(\vec{x}), \dots) \;, \quad \forall \vec{x} \in H \;.$$

Ejercicio 2.15. *Pruebe que para todo* $E \in \mathcal{B}(S^{\mathbb{Z}_+})$ *tenemos* $Y^{-1}(E) \in \mathcal{B}(S^{\mathbb{Z}_+})$.

Fijemos un punto $a \in A$ arbitrario. Definamos

$$\mathbb{Q}_a(E) \;:=\; \mathbb{P}_a(H \cap \{Y \in E\}) \;, \quad \forall E \in \mathcal{B}(S^{\mathbb{Z}_+}) \;. \tag{2.3.2}$$

donde $\mathbb{P}_a$ es la cadena de Markov partiendo de $a \in S$ y de probabilidades de transición $p(\cdot,\cdot)$.

Ejercicio 2.16. *Verifique que* $\mathbb{Q}_a$ *es una probabilidad para todo* $a \in A$.

Procedamos ahora a relacionar $\mathbb{Q}_a$ con la familia de probabilidades de transición $\{\tilde{p}(x,y) : x, y \in A\}$. Si fijamos $x \in A$, observe que

$$\mathbb{Q}_a\{X_1 = x\} \;=\; \mathbb{P}_a(H \cap \{Y_1 = x\}) \;.$$

Ya que H tiene probabilidad uno, esta probabilidad en la ecuación de arriba coincide con

$$\mathbb{P}_a\{T_A = T_x\,;\, T_x < \infty\} \;=\; \mathbb{P}_a\{T_A = T_x\} \;.$$

Por el Ejercicio 2.13, concluimos que

$$\mathbb{Q}_a\{X_1 = x\} \;=\; \tilde{p}(a,x) \;, \quad \forall a, x \in A. \tag{2.3.3}$$

Continuando, tomemos ahora un par de puntos $x, y \in A$. Tendremos

$$\mathbb{Q}_a\{X_1 = x, X_2 = y\} \;=\; \mathbb{P}_a(H \cap \{Y_1 = x, Y_2 = y\}). \tag{2.3.4}$$

Denotando $E := \{Y_0 = x, Y_1 = y\} \cap H$ observamos que

$$H \cap \{Y_1 = x, Y_2 = y\} \;=\; \{\Theta_{T_A} \in E\} \cap \{T_A < \infty\} \;.$$

Esta última expresión coincide con

$$\{\Theta_{T_x} \in E\} \cap \{T_x < \infty\} \cap \{T_A = T_x\} \;.$$

Entonces la probabilidad en (2.3.4) se puede escribir como

$$\mathbb{P}_a(\{\Theta_{T_x} \in E\} \cap \{T_x < \infty\} \cap \{T_A = T_x\}) \;.$$

Por el Teorema 2.13 y el Ejercicio 2.11, esta probabilidad es igual a

$$\mathbb{P}_a(\{T_x < \infty\} \cap \{T_A = T_x\})\, \mathbb{P}_x(E) \;=\; \mathbb{P}_a\{T_A = T_x\}\, \mathbb{P}_x\{Y_1 = y\} \;.$$

En la última igualdad hemos usado que $\mathbb{P}_a(\{T_x < \infty\} = 1$. Hemos probado así que

$$\mathbb{Q}_a\{X_1 = x\,;\, X_2 = y\} \;=\; \tilde{p}(a,x)\, \tilde{p}(x,y) \;. \tag{2.3.5}$$

Así, repitiendo la prueba de (2.3.5), y usando (2.3.3), podemos concluir por inducción que

Lema 2.16. *Para todo $n \geq 2$ y $x_1, \ldots, x_n \in A$ tenemos*

$$\begin{aligned} \mathbb{Q}_a\{X_1 = x_1\,;\,\ldots\,;\,X_n = x_n\} & \\ := \mathbb{P}_a\{Y_1 = x_1\,;\,\ldots\,;\,Y_n = x_n\} & \\ = \tilde{p}(a, x_1) \cdots \tilde{p}(x_{n-1}, x_n)\ , & \end{aligned}$$

cualquiera que sea el punto inicial $a \in A$.

En el lema anterior hemos escrito

$$\mathbb{P}_a\{Y_1 = x_1\,;\,\ldots\,;\,Y_n = x_n\} := \mathbb{P}_a\big(H \cap \{Y_1 = x_1\,;\,\ldots\,;\,Y_n = x_n\}\big)$$

para simplificar la notación. El hecho de que H tiene probabilidad uno, nos permite hacerlo.

El Lema 2.16 nos permite entonces decir que las probabilidades de transición

$$\{\tilde{p}(x, y) : x, y \in A\}$$

están determinando la dinámica de Y, quien es la traza que va dejando la dinámica de X sobre A, cuando X se está moviendo de acuerdo a las probabilidades de transición $\{p(x, y) : x, y \in S\}$.

Con un poco más de esfuerzo, podemos concluir la siguiente proposición a partir del Lema 2.16.

Proposición 2.17. *Sea $\{\tilde{p}(x, y) : x, y \in S\}$ la traza de $\{p(x, y) : x, y \in S\}$ sobre A. Para cada $a \in A$, la probabilidad $\mathbb{Q}_a$, inducida por Y como en la ecuación (2.3.2), es la cadena de Markov partiendo de $a \in A$ y de probabilidades de transición $\{\tilde{p}(x, y) : x, y \in S\}$.*

Prueba: Si $\mathbb{P}_a^A$ denota la cadena de Markov asociada a la traza

$$\{\tilde{p}(x, y) : x, y \in S\}\ ,$$

debemos probar que $\mathbb{P}_a^A \equiv \mathbb{Q}_a$. El Lema 2.3.2 nos permite concluir que $\mathbb{P}_a^A$ y $\mathbb{Q}_a$ coinciden sobre los cilindros. Eso implica que coinciden sobre todo $\mathcal{B}(S^{\mathbb{Z}_+})$. □

Usaremos el concepto de traza en el siguiente capítulo para resolver un problema interesante sobre la minimización del funcional de energía para funciones definidas en grafos finitos. Luego de haber asociado el funcional de energía con cadenas de Markov sobre los vértices del grafo, podremos reducir el número de vértices, tomando la traza de las cadenas, sin alterar el funcional de energía.

Capítulo 3

Minimización de la energía

En este capítulo resolveremos un problema sobre la minimización de la energía sobre grafos finitos, usando el concepto de cadena de Markov. Probaremos que el problema se puede simplificar, cambiando el grafo por otro que resulta equivalente para el problema. Esta simplificación es hecha vía la idea de traza de una cadena de Markov, introducida al final del capítulo anterior.

3.1 Energía y resultado previo

Fijemos un grafo G simple y conexo cuyo conjunto de vértices S es finito. Escribamos $x \sim y$ si los vértices $x, y \in S$ son unidos por alguna arista. Vamos a denotar por $\deg(x)$ el grado del vértice $x \in S$, es decir

$$\deg(x) \;=\; \sum_{y \in S} \mathbf{1}_{\{y \sim x\}} \; .$$

Para cada $f : S \to \mathbb{R}$ definamos la energía de f como

$$\mathcal{E}(f) \;=\; \frac{1}{2} \sum_{x \in S} \sum_{y \in S} \{f(y) - f(x)\}^2 \mathbf{1}_{\{x \sim y\}} \; .$$

Ejercicio 3.1. *Pruebe que* $\mathcal{E}(f) = 0$ *si y sólo si* f *es constante.*

Fijados dos subconjuntos $A, B \subseteq S$ disjuntos denotemos

$$F(A,B) \ := \ \{f : S \to \mathbb{R} : f \equiv 1 \text{ sobre } A \text{ y } f \equiv 0 \text{ sobre } B\}$$

Nosotros estamos interesados en el siguiente número

Definición 3.1. *Para cada par de subconjuntos disjuntos* $A, B \subseteq S$ *definimos la capacidad de* A *y* B *como*

$$\operatorname{cap}(A,B) \ := \ \inf\{\mathcal{E}(f) : f \in F(A,B)\} \ .$$

Obviamente $0 \le \operatorname{cap}(A,B) < \infty$.

Ejercicio 3.2. *Probar que* $\operatorname{cap}(A,B) = \operatorname{cap}(B,A)$.

Definamos,

$$\mathcal{E}(f,g) \ := \ \frac{1}{2} \sum_{x \in S} \sum_{y \in S} \{f(y) - f(x)\}\{g(y) - g(x)\} \mathbf{1}_{\{x \sim y\}} \ .$$

De modo que $\mathcal{E}(f) = \mathcal{E}(f,f)$.

Ejercicio 3.3. *Verifique que* $\mathcal{E}(\cdot,\cdot)$ *es bilineal, simétrica y definida positiva.*

El Ejercicio 3.4 abajo nos permite dar una expresión bastante útil para la energía. Para ello introducimos primero los siguientes dos ingredientes, un operador sobre las funciones y un producto interno.

Definimos M como el operador que a cada $f : S \to \mathbb{R}$ asocia la función $Mf : S \to \mathbb{R}$ definida por

$$Mf(x) \ := \ \frac{1}{\deg(x)} \sum_{y \in S} f(y) \mathbf{1}_{\{y \sim x\}} \ .$$

Además, escribamos I para denotar el operador identidad, es decir $If \equiv f$.

Por otro lado, para cada par de funciones $f, g : S \to \mathbb{R}$, definamos

$$\langle f, g \rangle \ = \ \sum_{x \in S} f(x) g(x) \deg(x) \ .$$

El siguiente ejercicio nos expresa la energía usando M y el producto $\langle\cdot,\cdot\rangle$.

Ejercicio 3.4. *Pruebe que para todo $f, g : S \to \mathbb{R}$ tenemos*

$$\mathcal{E}(f,g) \;=\; \langle (I-P)f, g\rangle \;=\; \langle f, (I-P)g\rangle\;.$$

Queremos probar que existe una única función en $F(A,B)$ tal que $\mathcal{E}(f) = \mathrm{cap}(A,B)$. Denotaremos $\Delta := S \setminus (A \cup B)$. Probemos el siguiente lema.

Lema 3.2. *Si $f \in F(A,B)$ es solución de la ecuación*

$$(I-M)f \equiv 0 \;\; \textit{sobre}\; \Delta\;, \tag{3.1.1}$$

entonces f es la única función en $F(A,B)$ tal que $\mathcal{E}(f) = \mathrm{cap}(A,B)$.

Prueba: Primero observemos que si f es como en el enunciado entonces $g \mapsto \mathcal{E}(f,g)$ es constante sobre todo $F(A,B)$. En particular $\mathcal{E}(f,g) = \mathcal{E}(f,f)$ para todo $f \in F(A,B)$. Eso prueba que para todo $g \in F(A,B)$ tenemos

$$0 \;\le\; \mathcal{E}(g-f, g-f) \;=\; \mathcal{E}(g,g) - \mathcal{E}(f,f)\;.$$

Entonces $\mathcal{E}(f) \le \mathcal{E}(g)$, $\forall g \in F(A,B)$ y si $f' \in F(A,B)$ es tal que $\mathcal{E}(f') = \mathcal{E}(f)$ entonces $\mathcal{E}(f'-f) = 0$ y por lo tanto $f \equiv f'$. □

Pero ya conocemos una solución para (3.1.1). Consideremos la familia de probabilidades de transición $\{p(x,y) : x, y \in S\}$ definida por

$$p(x,y) \;=\; \frac{1}{\deg(x)}\mathbf{1}_{\{y\sim x\}}\;. \tag{3.1.2}$$

Entonces $Mf(x) = \sum_{y\in S} p(x,y)f(y)$. Si definimos $f_{AB} \equiv 1$ sobre A, $f_{AB} \equiv 0$ sobre B y

$$f_{AB}(x) \;=\; \mathbb{P}_x\{T_A < T_B\}\;, \quad \forall x \in \Delta\;, \tag{3.1.3}$$

entonces $f_{AB} \in F(A,B)$ es solución de (3.1.1). Eso prueba que

Proposición 3.3. *La función $f_{A,B}$ definida en (3.1.3) es la única función en $F(A,B)$ que minimiza la energía, es decir, $\mathcal{E}(f_{AB}) = \mathrm{cap}(A,B)$.*

De manera un poco más general, dado un subconjunto no vacío $S_0 \subseteq S$ y una función $f : S_0 \to \mathbb{R}$ podemos definir

$$F_f(S_0) \ := \ \{g : S \to \mathbb{R} : \ g \equiv f \text{ sobre } S_0\} \ .$$

Observe que $F_f(S_0) = F(A, B)$ si $S_0 = A \cup B$, con A, B disjuntos y $f = \mathbf{1}_A$. Ahora podemos definir

$$\operatorname{cap}(f, S_0) \ := \ \inf\{\mathcal{E}(g) : g \in F_f(S_0)\} \ ,$$

de modo que $\operatorname{cap}(f, S_0) = \operatorname{cap}(A, B)$ en el caso particular $S_0 = A \cup B$, con A, B disjuntos y $f = \mathbf{1}_A$.

Ejercicio 3.5. *Probar que si $\hat{f} \in F_f(S_0)$ es solución de*

$$(I - M)\hat{f} \equiv 0 \ , \quad \textit{sobre } S \setminus S_0 \ , \tag{3.1.4}$$

entonces $\hat{f}$ es la única función en $F_f(S_0)$ que minimiza la energía.

Recuerde las probabilidades de transición definidas en (3.1.2).

Ejercicio 3.6. *Sea $\hat{f} \in F_f(S_0)$ la función definida por*

$$\hat{f}(x) \ = \ \sum_{a \in S_0} f(a) \mathbb{P}_x\{T_{S_0} = T_a\} \ , \quad \forall x \in S \setminus S_0 \ , \tag{3.1.5}$$

donde $\mathbb{P}_x$ es la cadena de Markov asociada a $\{p(x, y) : x, y \in S\}$. Pruebe que $\hat{f}$ es la única solución de (3.1.4) en $F_f(S_0)$.

El último ejercicio prueba que

Proposición 3.4. *La función $\hat{f} \in F_f(S_0)$ definida en (3.1.5) es la única función en $F_f(S_0)$ que minimiza la energía $\mathcal{E}$, es decir*

$$\mathcal{E}(\hat{f}) \ = \ \operatorname{cap}(f, S_0) \ .$$

3.2 Problema

Para esta sección, hemos fijado un subconjunto no vacío de vértices S_0 y una función $f : S_0 \to \mathbb{R}$. Continuaremos denotando por $\hat{f} : S \to \mathbb{R}$, la solución del problema en la Proposición 3.4. Es decir $\hat{f}$ es una extensión de S_0 de modo que la energía haya sido minimizada.

Consideremos ahora un subconjunto no vacío $D \subseteq S \setminus S_0$ y denotemos por $\mathfrak{C}_D$ el conjunto de funciones que son constantes sobre D. Por cada $g \in \mathfrak{C}_D$ denotemos por $g(D)$ dicha constante.

Queremos ahora minimizar la energía, pero sujetos a que las funciones deben ser constantes sobre un subconjunto $D \subseteq S \setminus S_0$. Definamos

$$\operatorname{cap}_f^D(S_0) \;:=\; \inf\{\mathcal{E}(g) : g \in \mathfrak{C}_D \cap F_f(S_0)\}\ .$$

Es claro que $\operatorname{cap}_f(S_0) \leq \operatorname{cap}_f^D(S_0)$.

Intente probar como ejercicio la siguiente proposición. Sugiero intente primero el caso $D = S \setminus S_0$.

Proposición 3.5. *Dentro de $\mathfrak{C}_D \cap F_f(S_0)$, existe una única función $f^* \in \mathfrak{C}_D \cap F_f(S_0)$ tal que*

$$\mathcal{E}(f^*) \;=\; \operatorname{cap}_f^D(S_0)\ .$$

Más aún, si $S \setminus (S_0 \cup D) \neq \varnothing$ entonces f^* debe satisfacer

$$(I - M)f^* \equiv 0\ , \quad \textit{sobre } S \setminus (S_0 \cup D)\ .$$

Esta última proposición nos permite extender la función f de un segundo modo f^*. En este modo, hemos obligado a la extensión f^* ser constante sobre un subconjunto D.

Es natural pensar que si f^*, al igual que $\hat{f}$, minimiza la energía, entonces f^* no puede ser demasiado diferente a $\hat{f}$ sobre D. Nuestro problema es mostrar lo siguiente.

Conjetura:

$$\min\{\hat{f}(x) : x \in D\} \;\leq\; f^*(D) \;\leq\; \max\{\hat{f}(x) : x \in D\}\ .$$

Vamos a generalizar un poco más lo que tenemos hasta ahora asociando la energía con un elemento definido a partir de la cadena de Markov.

3.3 Forma de Dirichlet

Vamos a fijar una familia irreductible de probabilidades de transición

$$\{p(x,y) : x, y \in S\} \ .$$

Vamos a suponer que tenemos el siguiente caso especial.

Definición 3.6. *Diremos que $\{p(x,y) : x, y \in S\}$ es reversible si existe una función $m : S \to (0,\infty)$ tal que*

$$m(x)p(x,y) \ = \ m(y)p(y,x) \ , \quad \forall x, y \in S \ . \qquad (3.3.1)$$

Observe que si m satisface (3.3.1), los múltiplos de m también lo hacen.

Ejercicio 3.7. *Siendo $\{p(x,y) : x, y \in S\}$ irreductible pruebe que si m, m' satisfacen (3.3.1) entonces $m = \alpha m'$ para algún $\alpha > 0$.*

Vamos a suponer en esta sección que $\{p(x,y) : x, y \in S\}$ además de ser irreductible es reversible. Llamaremos a cada m que satisface las ecuaciones en (3.3.1) de medida reversible para $\{p(x,y) : x, y \in S\}$.

Fijada un medida reversible m podemos definir

$$r(x,y) \ := \ m(x)p(x,y) \ , \quad \forall x, y \in S \ , \qquad (3.3.2)$$

de modo que $r(\cdot,\cdot)$ resulta simétrico. Podemos ahora usar $\{r(x,y) : x, y \in S\}$ como pesos sobre las aristas para generalizar la idea de energía. La energía calculada usando $\{r(x,y) : x, y \in S\}$ se llama la forma de Dirichlet asociada a las probabilidades de transición $\{p(x,y) : x, y \in S\}$ y una medida reversible $m : S \to (0,\infty)$.

Definición 3.7. *Sea $\{p(x,y) : x, y \in S\}$ una familia de probabilidades de transición reversible y sea $m : S \to (0,\infty)$ una medida reversible. La forma de Dirichlet asociada a $p(\cdot,\cdot)$ y m es la energía calculada con los pesos $r(\cdot,\cdot)$, es decir*

$$\mathcal{E}_r(f,g) \ := \ \frac{1}{2}\sum_{x\in S}\sum_{y\in S}\{f(y)-f(x)\}\{g(y)-g(x)\}r(x,y) \ .$$

Escribiremos $\mathcal{E}(f) := \mathcal{E}(f,f)$.

Como antes, definimos el operador

$$Mf(x) \;:=\; \sum_{y\in S} p(x,y)f(y)\ ,$$

para todo $f : S \to \mathbb{R}$. Además, definamos

$$\langle f, g\rangle_m \;:=\; \sum_{x\in S} f(x)g(x)m(x)\ .$$

En el caso particular de la sección anterior teníamos $m(x) = \deg(x)$ y $r(x,y)$ resultaba $\mathbf{1}_{\{x\sim y\}}$.

Ejercicio 3.8. *Con la notación que acabamos de introducir pruebe que*

$$\mathcal{E}_r(f,g) \;=\; \langle f, (I-M)g\rangle_m \;=\; \langle (I-M)f, g\rangle_m\ .$$

Denotemos por $\mathbb{P}_x$, $x \in S$ la cadena de Markov partiendo en x asociada a la familia de probabilidades de transición irreductible y reversible $\{p(x,y) : x, y \in S\}$. Sean $\{r(x,y) : x, y \in S\}$ los pesos sobre las aristas definidos como en (3.3.2) a partir de $p(\cdot,\cdot)$ y una medida reversible m. La prueba del siguiente teorema consiste en adaptar los argumentos usados en el contexto particular de la sección anterior.

Teorema 3.8. *Sea $S_0 \subseteq S$ no vacío y fijemos una función $f : S_0 \to \mathbb{R}$. Existe una única función $\hat{f}$ en $F_f(S_0)$ tal que*

$$\mathcal{E}_r(\hat{f}) \;=\; \inf\{\mathcal{E}_r(g) : g \in F_f(S_0)\}\ .$$

Esta función es definida como $\hat{f} \equiv f$ sobre S_0 y

$$\hat{f}(x) \;=\; \sum_{a\in S_0} f(a)\mathbb{P}_x\{T_{S_0} = T_a\}\ , \quad \forall x \in S \setminus S_0\ .$$

Además, si $S \setminus S_0 \neq \varnothing$ entonces

$$(I-M)\hat{f} \equiv 0\ , \quad \textit{sobre } S \setminus S_0\ . \tag{3.3.3}$$

3.4 Conservación de la forma de Dirichlet

Nuestro plan es usar la traza de $\{p(x,y) : x,y \in S\}$ sobre $S_0 \cup D$, para descartar el resto de vértices y simplificar el problema. Para ello, es necesario verificar que al retirar dichos vértices, el problema no se altera, es decir $\hat{f}$ y f^* continuan siendo los mismos. Eso es consecuencia de que la forma de Dirichlet de la traza coincide con la forma de Dirichlet en el grafo original para esas funciones.

Fijemos un punto $z \in S$. Recordemos que las probabilidades de transición de la traza de $p(\cdot,\cdot)$ sobre $S \setminus \{z\}$ son

$$p'(x,y) \;:=\; p(x,y) + \frac{p(x,z)p(z,y)}{1-p(z,z)}\;, \quad \forall x,y \in S \setminus \{z\}\;.$$

Ejercicio 3.9. *Pruebe que si $\{p'(x,y) : x,y \in S \setminus \{z\}\}$ es la traza de $\{p(x,y) : x,y \in S\}$ entonces*

$$m(x)p'(x,y) \;=\; m(y)p'(y,x)\;, \quad \forall x,y \in S \setminus \{z\}\;.$$

Es decir, $p'(x,y)$ es reversible y la restricción de m a $S \setminus \{z\}$ es una medida reversible para p'.

Entonces, podemos todavía definir nuevos pesos $\{r'(x,y) : x,y \in S \setminus \{z\}\}$, como $r'(x,y) = m(x)p'(x,y)$, sobre las aristas que restan después de retirar $z \in S$. Denotemos por $\mathcal{E}_{r'}$ la forma de Dirichlet asociada a la traza sobre $S \setminus \{z\}$.

El siguiente resultado nos permite retirar algunos puntos de S sin afectar la energía de una función g, siempre que el valor de $(I-M)(g)$ sobre dichos puntos resulte cero.

Lema 3.9. *Si $(I-M)g(z) = 0$ entonces, $\mathcal{E}_r(g) = \mathcal{E}_{r'}(g)$. En el lado derecho de la igualdad se toma la restricción de g sobre $S \setminus \{z\}$.*

Prueba: Ejercicio. □

Aplicando sucesivamente el lema anterior podemos concluir que

Lema 3.10. *Si $(I-M)f(z) = 0$ para todo $z \in A$ y algún subconjunto propio $A \subset S$ entonces*

$$\mathcal{E}_r(f) \;=\; \mathcal{E}_{\tilde{r}}(f)\;, \tag{3.4.1}$$

donde $\{\tilde{r}(x,y) : x,y \in S \setminus A\}$ *es definida como* $\tilde{r}(x,y) = m(x)\tilde{p}(x,y)$ *y*

$$\{\tilde{p}(x,y) : x,y \in S \setminus A\}$$

es la traza de $\{p(x,y) : x,y \in S\}$ *sobre* $S \setminus A$.

En el lado derecho de la igualdad (3.4.1) se toma la restricción de f sobre $S \setminus A$.

Como ejemplo de aplicación, definamos para cada familia irreductible y reversible de probabilidades de transición $\{p(x,y) : x,y \in S\}$ y m medida reversible, la capacidad de dos subconjunto disjuntos $A, B \subset S$ como

$$\mathrm{cap}_r(A,B) \ := \ \inf\{\mathcal{E}_r(g) : g \in F(A,B)\} \ .$$

Si fijamos un $C \subseteq S\setminus(A\cup B)$, podemos considerar $\{\tilde{p}(x,y) : x,y \in S\}$ la traza de $\{p(x,y) : x,y \in S\}$ sobre $S \setminus C$, $\tilde{r}(x,y) = m(x)\tilde{p}(x,y)$ y la capacidad $\mathrm{cap}_{\tilde{r}}(A,B)$ calculada con r' sobre $S \setminus C$.

Ejercicio 3.10. *Probar que* $\mathrm{cap}_r(A,B) = \mathrm{cap}_{\tilde{r}}(A,B)$.

3.5 Prueba de la conjetura

Fijemos una familia de probabilidades de transición irreductible y reversible $\{p(x,y) : x,y \in S\}$. Sean $\{r(x,y) : x,y \in S\}$ los pesos sobre las aristas definidos como en (3.3.2) a partir de $p(\cdot,\cdot)$ y una medida reversible m.

Fijemos un subconjunto no vacío $S_0 \subseteq S$ y una función $f : S_0 \to \mathbb{R}$. Sabemos que existe una única función $\hat{f} : S \to \mathbb{R}$ en $F_f(S_0)$ tal que

$$\mathcal{E}_r(\hat{f}) \ = \ \inf\{\mathcal{E}_r(g) : g \in F_f(S_0)\} \ , \tag{3.5.1}$$

donde $\mathcal{E}_r$ es la forma de Dirichlet de $p(\cdot,\cdot)$ y m. Pero además sabemos que $(I - M)\hat{f} \equiv 0$ sobre $S \setminus S_0$. Por lo visto en la sección anterior tenemos entonces lo siguiente.

Proposición 3.11. *Sea* $A \subseteq S \setminus S_0$. *Denotemos por* $\tilde{F}_f(S_0)$ *el conjunto de funciones definidas sobre* $S \setminus A$ *que coinciden con* f *sobre*

S_0. Entonces, la restricción de $\hat{f}$ sobre $S \setminus A$ es la única solución en $\tilde{F}_f(S_0)$ de

$$\mathcal{E}_{\tilde{r}}(\hat{f}) \ := \ \inf\{\mathcal{E}_{\tilde{r}}(g) : g \in \tilde{F}_f(S_0)\} \ ,$$

donde $\tilde{r}(x, y) = m(x)\tilde{p}(x, y)$ y $\tilde{p}(\cdot, \cdot)$ es la traza de $p(\cdot, \cdot)$ sobre $S \setminus A$.

Es decir, modificando convenientemente los pesos sobre las aristas, podemos reducir el problema a un grafo menor.

Ahora fijemos un subconjunto no vacío $D \subseteq S \setminus S_0$ y recordemos que $\mathfrak{C}_D$ es el conjunto de funciones que son constantes sobre D y para cada $g \in \mathfrak{C}_D$, $g(D)$ denota dicha constante. Estamos interesados en

$$\inf\{\mathcal{E}_r(g) : g \in \mathfrak{C}_D \cap F_f(S_0)\} \ .$$

Ejercicio 3.11. *Pruebe que* $\inf\{\mathcal{E}_r(g) : g \in \mathfrak{C}_D \cap F_f(S_0)\}$ *coincide con buscar el ínfimo sobre funciones que además cumplen con $(I - M)g \equiv 0$ sobre $S \setminus (S_0 \cup D)$:*

$$\begin{aligned} &\inf\{\mathcal{E}_r(g) : g \in \mathfrak{C}_D \cap F_f(S_0)\} \ = \\ &\inf\{\mathcal{E}_r(g) : g \in \mathfrak{C}_D \cap F_f(S_0) \text{ y } (I - M)g \equiv 0 \text{ sobre } S \setminus (S_0 \cup D)\} \ . \end{aligned}$$

Sugerencia: Use el Teorema 3.8.

Esta última observación y lo discutido en la sección anterior nos permite reducir el conjunto de vértices S a sólo $S_0 \cup D$ luego de redefinir los pesos de las aristas. Use esta técnica para probar la siguiente proposición.

Proposición 3.12. *Existe una única función f^* en $F_f(S_0) \cap \mathfrak{C}_D$ tal que*

$$\mathcal{E}_r(f^*) \ = \ \inf\{\mathcal{E}_r(g) : g \in \mathfrak{C}_D \cap F_f(S_0)\} \ .$$

Más aún, si denotamos por $\tilde{F}_f(S_0)$ el conjunto de funciones definidas sobre $S_0 \cup D$ que coinciden con f sobre S_0 entonces la restricción de f^* sobre $S_0 \cup D$ es también la única solución de

$$\mathcal{E}_{\tilde{r}}(f^*) \ = \ \inf\{\mathcal{E}_{\tilde{r}}(g) : g \in \mathfrak{C}_D \cap \tilde{F}_f(S_0)\} \ ,$$

donde $\tilde{r}(x, y) = m(x)\tilde{p}(x, y)$ y $\tilde{p}(\cdot, \cdot)$ es la traza de $p(\cdot, \cdot)$ sobre $S_0 \cup D$.

Recordemos que queremos probar que

$$\min\{\hat{f}(x) : x \in D\} \;\leq\; f^*(D) \;\leq\; \max\{\hat{f}(x) : x \in D\} \;.$$

Podemos ahora suponer sin pérdida de generalidad que $D = S \setminus S_0$, porque caso contrario podemos reducir el número de vértices y modificar $r(x,y)$ usando la traza de $p(\cdot,\cdot)$ sobre $S_0 \cup D$, de modo que ni la función $\hat{f}$ ni la función f^* se alteren.

Ahora es más fácil calcular $f^*(D)$. Si $S = S_0 \cup D$, es fácil ver que la forma de Dirichlet $\mathcal{E}_r(f^*)$ se puede escribir como

$$\sum_{y \in S_0} \{f(y) - f^*(D)\}^2 r(y,D) + \sum_{x \in S_0} \sum_{y \in S_0} \{f(y) - f(x)\}^2 r(x,y) \;, \quad (3.5.2)$$

Donde definimos, para todo par de subconjuntos disjuntos $A, B \subseteq S$,

$$r(A,B) \;:=\; \sum_{x \in A} \sum_{y \in B} r(x,y) \;. \qquad (3.5.3)$$

En (3.5.2) estamos escribiendo $r(y,D) := r(\{y\},D)$.

Es claro que $r(A,B) = r(B,A)$ para todo A, B subconjuntos disjuntos de S. Además tenemos la siguiente relación importante.

Ejercicio 3.12. *Verifique que para todo A, B subconjuntos disjuntos de S tales que $A \cup B = S$ se tiene*

$$r(A,B) \;=\; \mathcal{E}_r(\mathbf{1}_B) \;=\; \mathcal{E}_r(\mathbf{1}_A) \;=\; -\mathcal{E}(\mathbf{1}_A, \mathbf{1}_B) \;.$$

Ejercicio 3.13. *El valor de $f^*(D)$ que minimiza la expresión en (3.5.2) es*

$$f^*(D) \;=\; \frac{1}{r(S_0,D)} \sum_{y \in S_0} f(y) r(y,D) \;.$$

La expresión en (3.5.2) es fácil de minimizar.

Ejercicio 3.14. *Verifique que el valor $f^*(D)$ que minimiza la expresión en (3.5.2) se puede escribir como*

$$\mathcal{E}_r(f^* \mathbf{1}_{S_0}, \mathbf{1}_{S_0}) / \mathcal{E}_r(\mathbf{1}_{S_0}) \;=\; \mathcal{E}_r(f^* \mathbf{1}_{S_0}, \mathbf{1}_D) / \mathcal{E}_r(\mathbf{1}_{S_0}, \mathbf{1}_D) \;.$$

En este ejercicio note que la función $f^*\mathbf{1}_{S_0} : S \to \mathbb{R}$ que aparece en la expresión de $f^*(D)$, sólo depende de f.

Ahora ya estamos en posición de dar respuesta final al problema.

Prueba de la conjetura: Recuerde que $\hat{f}$ satisface $(I - M)\hat{f} \equiv 0$ sobre D. Eso significa que $\mathcal{E}_r(\hat{f}, \mathbf{1}_D) = 0$ y entonces,

$$\mathcal{E}_r(f^*\mathbf{1}_{S_0}, \mathbf{1}_D) \;=\; \mathcal{E}_r(\hat{f}\mathbf{1}_{S_0}, \mathbf{1}_D) \;=\; -\mathcal{E}_r(\hat{f}\mathbf{1}_D, \mathbf{1}_D)$$

Por el Ejercicio 3.14 y esta identidad podemos concluir que

$$f^*(D) \;=\; \mathcal{E}_r(\hat{f}\mathbf{1}_D, \mathbf{1}_D)/\mathcal{E}_r(\mathbf{1}_D)\ .$$

Resta observar que este cálculo es un promedio de los valores de $\hat{f}$ sobre D. □