Universitat de les Illes Balears

escola politècnica superior

Treball Final de Grau

GRAU DE MATEMÀTIQUES

Una Introducción a las Cadenas de Markov y sus Aplicaciones

JUAN JOSÉ MATAS SOBERÓN

Tutor
Óscar Valero

Escola Politècnica Superior
Universitat de les Illes Balears

Palma, 13 de julio de 2017

Me gustaría agradecer a mi tutor, *Óscar Valero*, el buen trato y ayuda que me ha brindado en el proceso de realización de este trabajo.

En la sociedad que vivimos hoy en día es de agradecer cualquier gesto de bondad, predisposición, respeto y comprensión. Y todo esto, al menos, lo he recibido de su parte.

Por ello, gracias Óscar.

Índice general

Índice general iii

Acrónimos v

Resumen vii

1 Introducción 1

2 Cadenas de Markov 3
- 2.1 Introducción . . . 3
- 2.2 Cadenas de Markov . . . 9
- 2.3 Cadenas de Markov Homogéneas . . . 13
- 2.4 Representación canónica de **P** . . . 14

3 Convergencia de las Matrices regulares estocásticas 23
- 3.1 Existencia de puntos fijos . . . 23
- 3.2 Convergencia de las potencias de **P** . . . 25
  - 3.2.1 La matriz **W** . . . 29
- 3.3 Cálculo del estado de equilibrio . . . 29
  - 3.3.1 Descomposición de Jordan . . . 30

4 Casos reales en los que aplicamos la teoría 33
- 4.1 Caso real: Fútbol . . . 33
  - 4.1.1 Introducción . . . 33
  - 4.1.2 Construcción de la matriz de transición . . . 33
  - 4.1.3 Cálculos . . . 35
- 4.2 Caso real: Póquer . . . 42
  - 4.2.1 Introducción . . . 42
  - 4.2.2 Construcción de la matriz de transición del caso: Rondas de apuestas 43
  - 4.2.3 Construcción de la matriz de transición del caso: Cartas buenas o malas . . . 45
  - 4.2.4 Cálculos . . . 46
- 4.3 Caso real: Google Adwords . . . 49
  - 4.3.1 Introducción . . . 49
  - 4.3.2 Construcción de la matriz de transición . . . 51
  - 4.3.3 Cálculos . . . 53

5 Conclusión 55

A Conocimientos Adicionales **57**
A.1 Conceptos relacionados con Álgebra lineal 57
A.2 Introducción a los Grafos Dirigidos 61
A.3 Probabilidad y Estadística 62
A.4 Conceptos de Topología 65
A.4.1 Conjuntos compactos 68
A.4.2 Espacios métricos completos 69
A.5 Teoría del Orden 69
A.6 Pequeña introducción a los números complejos 70
A.7 Aplicaciones iteradas y puntos fijos 70
A.8 Curiosidades 79
A.9 Más sobre las Cadenas de Markov (CC.M.) 80
A.10 Más sobre las Cadenas de Markov Homogéneas (CC.M.HH.) 85
A.10.1 Parámetros característicos 88
A.11 Ejemplo de una representación canónica 104
A.12 Más sobre la convergencia de las potencias de **P** 106
A.13 Ejemplo de cálculo de un punto fijo de P 107

B Demostraciones pendientes de los capítulos **109**
B.1 Cadenas de Markov 109
B.1.1 Introducción 109
B.1.2 Cadenas de Markov 114
B.1.3 Representación canónica de **P** 116
B.2 Convergencia de las Matrices regulares estocásticas 125

C Códigos de los programas utilizados **139**

Bibliografía **141**

Acrónimos

C.M. Cadena de Markov

CC.M. Cadenas de Markov

C.M.H. Cadena de Markov Homogénea

CC.M.HH. Cadenas de Markov Homogéneas

C.M.A. Cadena de Markov Absorbente

CC.M.AA. Cadenas de Markov Absorbentes

C.M.R. Cadena de Markov Regular

v.a. variable aleatoria

vv.aa. variables aleatorias

sii si y solo si

t.q. tal que

tt.q. tales que

e.e.f.n. espacio de estados finito o numerable

Resumen

Buscamos que el lector entre en profundidad en el mundo de las CC.M.. En particular, nuestro objetivo será encontrar condiciones necesarias y suficientes para obtener CC.M. que converjan a un estado de equilibrio, y mostrar métodos eficientes para poder calcularlo.

Además, veremos qué podemos hacer cuando las condiciones suficientes no se cumplen, y queremos conseguir información relevante de nuestra Cadena de Markov (C.M.).

Por último, como buenos matemáticos que pretendemos ser, ilustraremos con casos reales esta teoría que presentamos para probar cuán fehaciente y práctica es.

CAPÍTULO 1

INTRODUCCIÓN

Las CC.M. deben su nombre al ruso matemático de finales del siglo XIX, *Andrei Andreevich Markov*. Aunque este hombre estudió diversas áreas de las matemáticas, destacó principalmente por su trabajo en la teoría de la probabilidad.

Estos procesos estocásticos[1] aparecieron por primera vez en un estudio que hizo *Markov* sobre una novela en verso llamada *Eugene Onegin*, del escritor *Alexander Pushkin's*. En este análisis, *Markov* buscaba un patrón de aparición de las vocales y las consonantes en el texto de la novela.

A raíz de aquí, la teoría de las CC.M. se ha ido desarrollando con el paso del tiempo, aplicándose a diferentes campos. Algunos de los ámbitos donde han destacado con mayor repercusión son: física, biología y ciencias sociales.

Los físicos las han usado para modelar transformaciones radioactivas y detectores de fisión nuclear. Los astrónomos han estudiado la luminosidad de la *Vía Láctea* y la distribución espacial de las galaxias utilizando CC.M..

Los biólogos han usado esta teoría para saber más sobre el crecimiento de la población, genética molecular, farmacología y expansión de tumores entre otros...

Y, en el ámbito de las ciencias sociales, se aplican para estudiar el comportamiento de los votantes, movilidad geográfica en un país, expansión y disminución de los pueblos, predicción de las matriculaciones en colegios y universidades, etc.

Por todo esto, y mucho más, surge, a día de hoy, la necesidad de profundizar en esta teoría y ver sus diversas aplicaciones a la vida real.

En este trabajo encontraremos, primero de todo, una introducción a las CC.M. y a las CC.M.HH.. Expondremos definiciones, proposiciones y muchos ejemplos para ilustrar y verificar la comprensión de esta teoría.

A continuación, en el Capítulo 3, nos centraremos en el tema principal de este trabajo: la convergencia de las CC.M.HH. a estados de equilibrio.

[1] Un proceso estocástico, en la teoría de la probabilidad, es un concepto matemático que se utiliza para describir una serie de variables aleatorias (vv.aa.) (ver Definición A.3.3 del Apéndice A, para más información sobre las vv.aa.), que evolucionan en función de otra variable aleatoria (v.a.), generalmente, el tiempo.

Por último, pero no menos importante, en el Capítulo 4, estudiaremos tres casos reales en los que aplicamos la teoría que hemos mostrado.

En el Apéndice A, trataremos de introducir y explicar los términos que facilitarán la comprensión del trabajo. Además, encontraremos más información sobre las CC.M.HH. que no hemos podido introducir en el nudo de este trabajo por temas de espacio. Por otro lado, en el Apéndice B, mostraremos las demostraciones de teoremas y proposiciones, por capítulo, que habrán quedado pendientes de consulta para el lector, debido a la falta de espacio, o a la poca relevancia de la demostración con respecto al tema que tratamos.

Finalmente, en el Apéndice C encontraremos códigos programados en *Octave* que hemos usado para hacer algunos cálculos en el Capítulo 4.

Para escribir este trabajo, nos hemos basado principalmente en dos libros de la Bibliografía: [1] y [2]. Y, para ver ejemplos de aplicaciones de la teoría de las CC.M., hemos consultado la Referencia [3]. Además, hemos comprobado algunas demostraciones en la Referencia [4].

CAPÍTULO 2

Cadenas de Markov

2.1 Introducción

Las CC.M. son procesos estocásticos en los que el pasado afecta al futuro solo a través del presente. Aunque esta frase pueda parecernos sorprendente o incomprensible *a priori*, notaremos como, a medida que vayamos leyendo este trabajo, cobrará más y más sentido.

En este capítulo, nos centraremos en explicar la notación que usaremos a lo largo del trabajo, definir las CC.M., y exponer diferentes proposiciones, teoremas y conceptos sobre ellas, que nos servirán para abordar mejor el tema principal que trataremos en el capítulo posterior.

Nos interesa, primero de todo, definir un conjunto S finito o numerable, de tal forma que:

Definición 2.1.1. *El vector fila* $\mathbf{x} = (x_j)_{j\in S}$ *es un* ***vector indexado por*** $\mathbf{S}$ *si sus componentes son reales y se indexan como hemos denotado.*

De igual forma, tendríamos la definición para las matrices:

Definición 2.1.2. *La matriz* $\mathbf{P} = (\mathbf{P}^i_j)_{i,j\in S}$ *es una* ***matriz indexada por*** $\mathbf{S}$ *si sus componentes son reales y se indexan como hemos denotado.*

Denotaremos por $\mathbf{P}_j$, con $j \in S$, a la columna j-ésima de la matriz $\mathbf{P}$. Y, por $\mathbf{P}^i$, con $i \in S$, a la fila i-ésima.

Las matrices indexadas por S que nos interesan, son aquellas que tienen una norma finita:

$$\|\mathbf{P}\| = \|\mathbf{P}\|_\infty = \sup_{i\in S}\{\sum_{j\in S} |\mathbf{P}^i_j|\} < +\infty,$$

donde $\|\mathbf{P}\|_\infty$ es la norma matricial inducida por la norma ℓ_∞ [1]. A partir de ahora, será de estas matrices únicamente de las que hablaremos.

[1] Véase Ejemplo A.1.3 del Apéndice A para más información sobre la norma ℓ_∞.

En lo sucesivo, asumiremos que $|S| = N$, con $N \in \mathbb{N}$ (donde $\mathbb{N}$ denota el conjunto de los enteros no negativos).

Tenemos la necesidad de definir la siguiente aplicación lineal asociada a una matriz $\mathbf{P}$. Debido a que, como veremos en el próximo capítulo, nos hará falta una aplicación que cumple ciertos requisitos para aplicar teoremas de puntos fijos:

Definición 2.1.3. *Sea* $\mathbf{P} \in \mathrm{M}_{N\times N}(\mathbb{R})$ *indexada por S. Definimos la* ***aplicación lineal asociada a la matriz*** $\mathbf{P}$ *como:*

$$\begin{array}{rccc} \mathrm{P}: & \ell_1 & \longrightarrow & \ell_1 \\ & \mathbf{x} & \longrightarrow & \mathbf{xP}. \end{array}$$

(Véase el Ejemplo A.1.1 del Apéndice A para obtener más información sobre ℓ_1*.)*

Hemos comprobado, si le interesa al lector, que esta definición esta bien definida en la Demostración B.1.1 del Apéndice B.

Veamos, con las siguientes definiciones y ejemplos, el tipo de vectores y matrices que trataremos:

Definición 2.1.4. *Sea* $(\mathbb{R}^N, \|\cdot\|_1)$ *un espacio vectorial normado* [2]*, donde* $\|\cdot\|_1 : \mathbb{R}^N \longrightarrow \mathbb{R}$ *viene dada por* $\|\mathbf{x}\|_1 = \sum_{i\in S} x_i$.

El vector fila $\mathbf{x} = (x_i)_{i\in S} \in \mathbb{R}^N$ *es un* ***vector estocástico*** *si cumple las siguientes condiciones:*

a) $x_i \geq 0, \forall i \in S,$

b) $\|\mathbf{x}\|_1 = 1,$

Denotamos por $\mathscr{T}$*, al* ***conjunto de los vectores estocásticos*** *en* $\mathbb{R}^N$.

A continuación, introduciremos un ejemplo que ilustra la definición anterior:

Ejemplo 2.1.1. *Los vectores* $\mathbf{e^1}, \mathbf{e^2}, \ldots, \mathbf{e^N} \in \mathbb{R}^N$*, componentes de la base canónica de* $\mathbb{R}^N$*, son vectores estocásticos.*

Ya que, sea $\mathbf{e^j} = ((e^j)_i) = (0,\ldots,0,\underbrace{1}_{j},0,\ldots,0)$ *un componente de la base canónica arbitrario pero fijado:*

- $(e^j)_i \geq 0, \forall i \in S.$
- $\|\mathbf{e^j}\|_1 = \sum_{i\in S}(e^j)_i = 0 + \cdots + 0 + \underbrace{1}_{j} + 0 + \cdots + 0 = 1.$

Definición 2.1.5. *La matriz* $\mathbf{P} = (\mathbf{P}^i_j)_{i,j\in S} \in \mathrm{M}_{N\times N}$ *es una* ***matriz estocástica*** *si cada una de las filas de la matriz es un vector estocástico. O sea, si cumple las siguientes condiciones:*

a) $\mathbf{P}^i_j \geq 0, \forall i, j \in S.$

[2]Véase la Definición A.1.1 del Apéndice A para más información.

b) $\sum_{j\in S} \mathbf{P}^i_j = 1, \forall i \in S.$

El siguiente es un ejemplo de matrices estocásticas:

Ejemplo 2.1.2.

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}.$$

Cada una de las entradas (o valores) de la matriz $\mathbf{P}$ *son mayores o iguales que* 0 *y, para cada fila, la suma de sus valores es* 1.

Observación 2.1.1. *Si* $\mathbf{P}$ *es una matriz estocástica indexada por* S*, tenemos que*

1) $|\mathbf{P}^i_j| = \mathbf{P}^i_j$ *ya que* $\mathbf{P}^i_j \geq 0, \forall i, j \in S.$

2) $\sum_{j\in S} \mathbf{P}^i_j = 1, \forall i \in S.$

Por tanto:

$$\|\mathbf{P}\| = \sup_{i\in S}\{\sum_{j\in S} |\mathbf{P}^i_j|\} = \sup_{i\in S}\{\sum_{j\in S} \mathbf{P}^i_j\} = \sup_{i\in S}\{1\} = 1.$$

Y, como $\mathbf{P}^i_j \geq 0, \forall i, j \in S$*, tendríamos que*

$$\mathbf{P}^i_j \leq 1, \forall i, j \in S.$$

Una vez contamos con estas definiciones, pasaremos a ver una proposición y un par de propiedades que están relacionadas con las matrices estocásticas y sus respectivas aplicaciones lineales. Nos servirán para aclarar unas propiedades con las que cuentan este tipo de matrices y que nos harán falta en secciones posteriores y en el Capítulo 3:

Proposición 2.1.1. *Sea* $\mathbf{P}$ *una matriz indexada por* S *tal que (t.q.)* $\|\mathbf{P}\| < +\infty$ *y sea* P *su aplicación lineal asociada. Entonces* $\mathbf{P}$ *es una matriz estocástica si y solo si (sii)* $\mathrm{P}(\mathscr{T}) \subset \mathscr{T}$.

Demostración. *Consúltese la Demostración B.1.2 del Apéndice B.*

Una de las consecuencias de esta proposición es que el producto de matrices estocásticas indexadas por S, es también estocástica, y, en particular, las potencias de matrices estocásticas también lo son. Veamos porqué:

Proposición 2.1.2. *Sean* $\mathbf{P}$ *y* $\mathbf{Q}$ *matrices estocásticas indexadas por* S*, y* P *y* Q *las aplicaciones lineales asociadas a las matrices* $\mathbf{P}$ *y* $\mathbf{Q}$ *respectivamente. Sea*

$$\begin{array}{rccc} (\mathrm{Q}\circ\mathrm{P}): & \ell_1 & \longrightarrow & \ell_1 \\ & \mathbf{x} & \longrightarrow & \mathbf{x}(\mathbf{PQ}) \end{array}$$

la composición de las respectivas aplicaciones Q *y* P*. Entonces la matriz* $\mathbf{PQ}$ *es una matriz estocástica.*

(En la Demostración B.1.3 del Apéndice B podemos comprobar porqué esta composición está bien definida).

Demostración. *Queremos ver que* $(\mathrm{Q}\circ\mathrm{P})(\mathscr{T})\subset\mathscr{T}$. *Por ser* **P** *una matriz estocástica, tenemos que, por la Proposición 2.1.1,* $\mathrm{P}(\mathscr{T})\subset\mathscr{T}$. *Por tanto:*

$$(\mathrm{Q}\circ\mathrm{P})(\mathscr{T})=\mathrm{Q}(\mathrm{P}(\mathscr{T}))\subset\mathrm{Q}(\mathscr{T}),$$

y por ser la matriz **Q** *también estocástica, tenemos que* $\mathrm{Q}(\mathscr{T})\subset\mathscr{T}$, *y así nos queda que:*

$$(\mathrm{Q}\circ\mathrm{P})(\mathscr{T})\subset\mathscr{T},$$

tal y como queríamos ver. Ahora, por la Proposición 2.1.1 de nuevo, tendríamos que **PQ** *sería una matriz estocástica.*

□

De aquí deducimos que si **P** es una matriz estocástica, entonces $(\mathbf{P}^n)$ (la potencia n-ésima de la matriz **P**) es una matriz estocástica para cualquier $n\in\mathbb{N}\cup\{0\}$.

La otra propiedad interesante que cumplen las matrices estocásticas es que cualquier combinación convexa de ellas, es también una matriz estocástica. Veamos:

Proposición 2.1.3. *Sea*

$$\mathbf{P}=\sum_{k=1}^{M}t_k(\mathbf{P}_k)$$

una combinación convexa de matrices estocásticas indexadas por S. Donde $(\mathbf{P}_k)$ *denota la matriz k-ésima, con* $k\in\{1,\dots,M\}$, $M\in\mathbb{N}$, *y los coeficientes* t_k *cumplen lo siguiente:*

a) $t_k\geq 0$, $\forall k\in\{1,\dots,M\}$.

b) $\sum_{k=1}^{M}t_k=1$.

Entonces **P** *es una matriz estocástica.*

Demostración. *Consúltese la Demostración B.1.4 del Apéndice B.*

Con el fin de tratar las matrices estocásticas de una forma más intuitiva, podemos representarlas mediante grafos dirigidos.[3]

Las interpretaciones que hagamos de las entradas de **P** serán más precisas si contamos con estas representaciones.

Para obtener dicha representación, lo que debemos hacer es:

Sea **P** una matriz estocástica indexada por S t.q. $\mathbf{P}=(p_{ij})_{i,j\in S}$. Asociamos un grafo dirigido con pesos a la matriz t.q.: si $p_{ij}>0$, dibujamos un arco dirigido del nodo i al nodo j y asociamos el peso p_{ij} al arco, y si $p_{ij}=0$ no dibujamos nada.

Con el siguiente ejemplo lo veremos más sencillo:

[3] Dejamos a interés del lector, una pequeña introducción y notación sobre grafos dirigidos en la Sección A.2 del Apéndice A.

Ejemplo 2.1.3. *Sea*

$$\mathbf{P} = \begin{pmatrix} 2/9 & 0 & 7/9 \\ 1/2 & 0 & 1/2 \\ 1/7 & 6/7 & 0 \end{pmatrix}$$

la matriz estocástica. Notemos que en este caso tenemos 3 *filas y* 3 *columnas, por tanto podemos definir* $S = \{1, 2, 3\}$ *como el conjunto de nodos del grafo para representar cada una de ellas:*

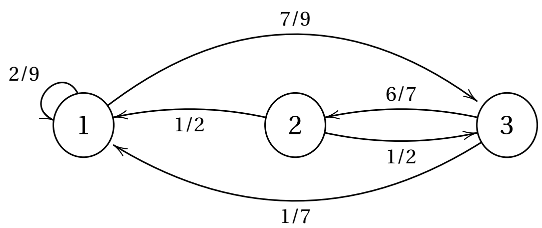

Figura 2.1: Grafo dirigido asociado a **P** del Ejemplo 2.1.3.

Observamos que desde el nodo 1*, podemos ir al nodo* 1 *(él mismo) con probabilidad* $2/9$*, al nodo* 2 *con probabilidad* 0 *(o sea, no podemos ir en un paso o arco) y al nodo* 3 *con probabilidad* $7/9$*. Esto lo hemos podido representar gracias a la información que nos proporciona la primera fila de la matriz* **P**. *Las demás filas nos ayudarían a completar el grafo.*

También vemos que, por ser todos los elementos de la primera columna de la matriz **P** *diferentes de* 0*, tenemos que podemos llegar al nodo* 1 *desde cualquier otro nodo (incluido él mismo) mediante un único arco (en un paso).*

Solo con la matriz **P**, *no hubiéramos visto que hay un camino dirigido del nodo* 1 *al nodo* 2*, ya que* $p_{12} = 0$*. Pero con el grafo dirigido, vemos claramente que sí lo hay, yendo del nodo* 1 *al nodo* 3 *primero, y del nodo* 3 *al nodo* 2 *después.*

Definición 2.1.6. *Sea* **P** *una matriz estocástica indexada por* S *t.q.* $i, j \in S$.

Si $\exists\, n \in \mathbb{N}$ *t.q.* $p_{ij}^{(n)} = (\mathbf{P}^n)^i_j > 0$*, decimos que* **i *conduce a* j**, *o que* **j *es accesible desde* i**, *y lo denotamos como* $i \longrightarrow j$.

Si $i \longrightarrow j$ *y* $j \longrightarrow i$*, entonces* **i *y* j *están comunicados***, *y escribimos* $i \longleftrightarrow j$.

Observación 2.1.2. *Lo que quiere decir la definición anterior es que si existe algún camino del nodo* i *al nodo* j*, entonces* i *conduce a* j*. Y, que si existe dicho camino y además otro que vaya del nodo* j *al* i*, entonces* j *e* i *están comunicados.*

Una vez contamos con esta definición, podemos introducir un concepto muy importante en este trabajo debido a que lo usaremos en el Capítulo 3:

Definición 2.1.7 (Matriz irreducible)**.** *Sea* **P** *una matriz estocástica indexada por* S *t.q.* $i, j \in S$.

Si $i \longleftrightarrow j$*,* $\forall i, j \in S$*, entonces decimos que* **P *es irreducible***.

En el siguiente ejemplo ilustraremos la definición anterior:

Ejemplo 2.1.4. *Sea*

$$\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

una matriz estocástica indexada por $S = \{1, 2\}$*. El grafo dirigido asociado a esta matriz es:*

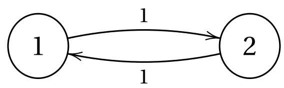

Figura 2.2: Grafo dirigido asociado a **P** del Ejemplo 2.1.4.

Como $p_{12}^{(1)} > 0$*, tenemos que el nodo* 1 *conduce al nodo* 2 *por un camino dirigido compuesto únicamente de un arco (en un paso). De igual manera, como* $p_{21}^{(1)} > 0$*, tendríamos que* 2 *conduce a* 1*. Por tanto,* $1 \longleftrightarrow 2$*, y como* $S = \{1, 2\}$*, tendríamos que* **P** *es irreducible.*

Proposición 2.1.4. *La relación* $i \longleftrightarrow j$ *es una relación de equivalencia del conjunto de índices* S*.*

Demostración. *Consúltese la Demostración B.1.5 del Apéndice B.*

Gracias a esta proposición, podemos dividir el conjunto S en clases de equivalencia de índices que se comunican entre sí. Esta propiedad nos será de gran ayuda en la Sección 2.4 posterior.

La siguiente proposición establece un modo para determinar cuando 2 nodos están comunicados en función de la Definición 2.1.6 anterior:

Proposición 2.1.5. *Sea* **P** *la matriz estocástica indexada por un* S*. Un camino dirigido del nodo* i *al nodo* j *existe sii existe al menos un camino formado por no más de* N *arcos.*

Así, un camino de i *a* j *existe sii*

$$\mathbf{B}^i_j = (\mathbf{P} + \mathbf{P}^2 + \cdots + \mathbf{P}^{N-1} + \mathbf{P}^N)^i_j > 0. \tag{2.1}$$

En particular, para cada nodo $i \in S$*, el conjunto*

$$\{j \in S \,/\, \mathbf{B}^i_j > 0\} \tag{2.2}$$

es el conjunto de nodos que son accesibles, por al menos un camino dirigido, desde i*.*

Demostración. *Consúltese la Demostración B.1.6 del Apéndice B.*

Antes de explicar lo que quiere decir esta proposición, queremos aclarar que: $\mathbf{P}^i_j$ representa los caminos dirigidos de i a j compuestos por un arco únicamente, $(\mathbf{P}^2)^i_j$ representa los de 2 arcos, y así sucesivamente hasta $(\mathbf{P}^N)^i_j$ que representa los de N arcos exactamente.

Ahora bien, al definir $\mathbf{B}^i_j$ como una suma de las diferentes potencias de **P**, lo que nos quiere decir la proposición es que si existe al menos un camino de $1, 2, 3, \ldots,$ o N arcos exactamente, entonces existe al menos un camino de i a j.

Esto es debido a que, como las potencias de las matrices estocásticas son también estocásticas, tenemos que $(\mathbf{P}^k)^i_j \geq 0 \, \forall i, j \in S$ y $k \in \{1, \ldots, N\}$. Ahora, sabemos por la Definición 2.1.6 que si $p^{(k)}_{ij} = (\mathbf{P}^k)^i_j > 0$, entonces hay un camino dirigido de i a j en k pasos (o arcos). Por lo que si existe un $k \in \{1, \ldots, N\}$ t.q. $(\mathbf{P}^k)^i_j > 0$, tendremos que $\mathbf{B}^i_j > 0$ y, por tanto, existirá un camino dirigido de i a j por dicha definición.

El siguiente ejemplo ilustra la Proposición 2.1.5:

Ejemplo 2.1.5. *Sea*

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

la matriz estocástica indexada por $S = \{1, 2, 3\}$. El grafo dirigido asociado a $\mathbf{P}$ es:

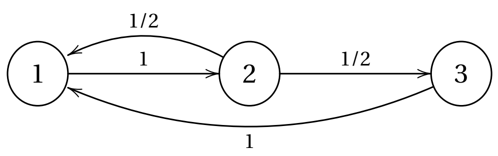

Figura 2.3: Grafo dirigido asociado a $\mathbf{P}$ del Ejemplo 2.1.5.

Calculemos $\mathbf{P}^2$ y $\mathbf{P}^3$:

$$\mathbf{P}^2 = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{P}^3 = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & 1/2 \end{pmatrix}.$$

Ahora calcularemos $\mathbf{B}^i_j$ para detectar los caminos entre los nodos del grafo:

$$\mathbf{B}^i_j = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \\ 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 1 & 3/2 & 1/2 \\ 5/4 & 1 & 5/4 \\ 3/2 & 1 & 1/2 \end{pmatrix}.$$

Por tanto, tenemos que existen caminos de cualquier nodo a cualquier nodo del grafo, ya que todas las entradas de $\mathbf{B}$ son mayores que 0.

Una vez contamos con esta introducción, estamos en disposición de definir y aclarar lo que son las CC.M..

2.2 Cadenas de Markov

A continuación, mostraremos una de las definiciones más importantes de este trabajo e introduciremos gran parte de la notación relacionada con la misma:

Definición 2.2.1 (Cadena de Markov)**.** *Sea* $(\Omega, \mathscr{E}, \mathbb{P})$ *un espacio de probabilidad* [4] *y sean* $1,2,\dots,r \in \mathscr{E}$ *los sucesos posibles del experimento aleatorio* Ω*. A estos sucesos los llamaremos* ***estados*** *y* $S = \{1,2,\dots,r\}$ *será el* ***espacio de estados****.*

Dicho experimento es repetido un número N *de veces. Sea* $\{\mathscr{X}_n\}$*, con* $n \in \{0,1,\dots,N\}$*, la sucesión de vv.aa. en este espacio de probabilidad, de tal forma que:*

$$\begin{array}{rccc} \mathscr{X}_n : & \Omega & \longrightarrow & S \\ & x & \longrightarrow & i. \end{array}$$

Entonces $\{\mathscr{X}_n\}$ *es una C.M. si existe un conjunto* $\{p_{ij}(n+1)\}_{i,j\in S}$ *t.q. la probabilidad condicionada de que ocurra el suceso* j *en el momento* $(n+1) \in \{0,1,\dots,N\}$*, dado que el resultado en la repetición anterior ha sido el suceso* i*, es* $p_{ij}(n+1)$*:*

$$p_{ij}(n+1) = \mathbb{P}(j \textit{ en el experimento } n+1 \;/\; i \textit{ en el experimento } n) = \mathbb{P}(\mathscr{X}_{n+1} = j \;/\; \mathscr{X}_n = i),$$

$$\forall i,j \in S \; y \; n \in \{0,1,\dots,N\}.$$

Los valores $p_{ij}(n+1)$ *son conocidos como* ***probabilidades de transición****.*

Observación 2.2.1. *Fijémonos que para cada* $x \in \Omega$*, la secuencia de estados* $\{\mathscr{X}_n(x)\}_n \subset S$ *define un camino en* S*.*

Si $\mathscr{X}_n(x) = j$*, decimos que* el camino definido por x visita a j en el paso n*. El conjunto* $\{x \in \Omega \;/\; \mathscr{X}_n(x) = j\}$ *es el conjunto de caminos que visitan* j *en el paso* n *y* $\mathbb{P}(\mathscr{X}_n = j \;/\; \mathscr{X}_0 = i)$ *es la probabilidad de que un camino que empieza en* i*, esté en* j *en el paso* n*.*

El conjunto de las probabilidades $\{p_{ij}(n+1)\}$ pueden ser representadas por una matriz:

$$\mathbf{P}(n+1)^i_j = \begin{pmatrix} p_{11}(n+1) & p_{12}(n+1) & \cdots & p_{1r}(n+1) \\ p_{21}(n+1) & p_{22}(n+1) & \cdots & p_{2r}(n+1) \\ \vdots & \vdots & \ddots & \vdots \\ p_{r1}(n+1) & p_{r2}(n+1) & \cdots & p_{rr}(n+1) \end{pmatrix}$$

conocida como la matriz de transición de la C.M. $\{\mathscr{X}_n\}$, del paso n al paso $n+1$. Veamos con más detenimiento este concepto a continuación:

Definición 2.2.2 (Matriz de transición)**.** *Sea* $\{\mathscr{X}_n\}$ *un proceso estocástico con un espacio de estados* S*. Para cada* $n \in \mathbb{N} \cup \{0\}$*, sea*

$$\mathbf{P}(n+1)^i_j = \begin{cases} \mathbb{P}(\mathscr{X}_{n+1} = j \;/\; \mathscr{X}_n = i) & si \; \mathbb{P}(\mathscr{X}_n = i) > 0 \\ \delta^i_j & si \; \mathbb{P}(\mathscr{X}_n = i) = 0 \end{cases}$$

Donde

$$\mathbf{Id} = \delta^i_j = \begin{cases} 1 & si \; i = j \\ 0 & si \; i \neq j \end{cases}$$

Llamamos a $\mathbf{P}(n+1)$ *la* ***matriz de probabilidades de transición*** *o* ***matriz de transición*** *del proceso* $\{\mathscr{X}_n\}$*, del paso* n *al paso* $n+1$*.*

[4] Véase la Definición A.3.2 del Apéndice A para más información sobre lo que es un espacio de probabilidad.

Observación 2.2.2. *Como las entradas de la matriz de transición son probabilidades condicionadas, es necesario que el suceso que condiciona ($\{\mathscr{X}_n = i\}$) tenga probabilidad estrictamente positiva. Es muy extraño que $\mathbb{P}(\mathscr{X}_n = i)$ sea* 0*, de echo, en todos los ejemplos y casos reales que consideraremos en este capítulo y en el Capítulo 4, $\mathbb{P}(\mathscr{X}_n = i)$ será mayor que* 0*.*

Pero en la Definición 2.2.2, que es la definición general, se debe considerar la posibilidad de que $\mathbb{P}(\mathscr{X}_n = i)$ también puede valer 0*. Y para poder realizar la demostración de la Proposición 2.2.1 con éxito, definimos $\mathbf{P}(n+1)^i_j = \delta^i_j$ en el caso en que $\mathbb{P}(\mathscr{X}_n = i) = 0$.*

Los grafos dirigidos que hemos visto en la introducción, tienen nombre propio en las CC.M.:

Definición 2.2.3. *El grafo dirigido asociado a la matriz de transición* **P***, se conoce como* ***diagrama de estados****.*

La siguiente proposición, aunque parece simple, es muy importante, debido a que todos los resultados posteriores de este capítulo y del Capítulo 3 se basarán en ella:

Proposición 2.2.1. *Las matrices de transición son matrices estocásticas.*

Demostración. *Sea $n \in \mathbb{N} \cup \{0\}$ fijado pero arbitrario, y sea $\mathbf{P}(n+1)^i_j = p_{ij}(n+1)$ para cada $i, j \in S$. Para ver que $\mathbf{P}(n+1)$ es una matriz estocástica, veamos que cumple las dos condiciones para serlo:*

a) Por ser $\mathbb{P}$ una medida de probabilidad [5]*, tenemos que $p_{ij}(n+1) \geq 0, \forall i, j \in S$ por la Definición 2.2.2 anterior.*

b) Sea $i \in S$ arbitrario pero fijado, queremos ver que $\sum_{j\in S} p_{ij}(n+1) = 1$. Tenemos dos casos:

- *Si $\mathbb{P}(\mathscr{X}_n = i) = 0$, tenemos que $p_{ij}(n+1) = \delta^i_j$ por la Definición 2.2.2 de nuevo, por lo que:*

$$\sum_{j\in S} p_{ij}(n+1) = \sum_{j\in S} \delta^i_j = 1.$$

- *Si $\mathbb{P}(\mathscr{X}_n = i) > 0$, entonces $p_{ij}(n+1) = \mathbb{P}(\mathscr{X}_{n+1} = j \,/\, \mathscr{X}_n = i)$ y, por la fórmula de las probabilidades condicionadas* [6]*, tendríamos que:*

$$\sum_{j\in S} p_{ij}(n+1) = \sum_{j\in S} \mathbb{P}(\mathscr{X}_{n+1} = j \,/\, \mathscr{X}_n = i) = \frac{\sum_{j\in S} \mathbb{P}(\{\mathscr{X}_{n+1} = j\} \cap \{\mathscr{X}_n = i\})}{\mathbb{P}(\mathscr{X}_n = i)}.$$

Ahora, sea $E_j = \{\mathscr{X}_{n+1} = j\}$ para cada $j \in S$. Notemos que los sucesos E_j forman una partición sobre Ω:

1) Los E_j son disjuntos dos a dos.

[5] Contamos con la Definición A.3.1 en el Apéndice A, por si le interesa al lector.

[6] Consultar la Definición A.3.9 del Apéndice A para ver la fórmula.

2) $\bigcup_{j\in S} E_j = \Omega$.

Por ser $\mathbb{P}$ una medida de probabilidad, la suma de probabilidades de una serie de sucesos es igual a la probabilidad de la unión de dichos sucesos, por tanto:

$$\frac{\sum_{j\in S}\mathbb{P}(E_j\cap\{\mathscr{X}_n=i\})}{\mathbb{P}(\mathscr{X}_n=i)}=\frac{\mathbb{P}\left(\bigcup_{j\in S}(E_j\cap\{\mathscr{X}_n=i\})\right)}{\mathbb{P}(\mathscr{X}_n=i)}.$$

Por teoría de conjuntos, sabemos que la unión de intersecciones de conjuntos es igual a la intersección de uniones de éstos, por lo que:

$$\frac{\mathbb{P}\left(\bigcup_{j\in S}(E_j\cap\{\mathscr{X}_n=i\})\right)}{\mathbb{P}(\mathscr{X}_n=i)}=\frac{\mathbb{P}\left(\{\bigcup_{j\in S}E_j\}\cap\{\mathscr{X}_n=i\}\right)}{\mathbb{P}(\mathscr{X}_n=i)}=\frac{\mathbb{P}(\Omega\cap\{\mathscr{X}_n=i\})}{\mathbb{P}(\mathscr{X}_n=i)}=\frac{\mathbb{P}(\mathscr{X}_n=i)}{\mathbb{P}(\mathscr{X}_n=i)}=1.$$

Esto quiere decir que para cualquiera de los dos casos hemos visto que $\sum_{j\in S} p_{ij}(n+1)=1$. En definitiva:

$$\sum_{j\in S} p_{ij}(n+1)=1,\ \forall i\in S.$$

□

A continuación, expondremos un concepto que nos ayudará a ver la relación que existe entre las distribuciones de las vv.aa. $\mathscr{X}_n$ y $\mathscr{X}_0$ mediante la matriz de transición $\mathbf{P}$. Para ello, necesitamos el concepto de densidad de masa:

Definición 2.2.4. *Sea $n\in\mathbb{N}\cup\{0\}$ arbitrario pero fijado, definimos $\pi(n)$ como el vector fila t.q. cada componente cumple que:*

$$(\pi(n)_i)=\mathbb{P}(\mathscr{X}_n=i),$$

*para cada $i\in S$. Llamaremos a este vector **densidad de masa** de $\mathbb{P}_{\mathscr{X}_n}$.*

Proposición 2.2.2. *La siguiente igualdad se verifica:*

$$\pi(n+k)=\mathbf{P}(n+k)\cdots\mathbf{P}(2+k)\mathbf{P}(1+k)\pi(k),\ \forall n,k\in\mathbb{N}\cup\{0\}.$$

Demostración. *Consúltese en la Demostración B.1.7 del Apéndice B.*

Esta propiedad que acabamos de enunciar establece que dada la C.M. $\{\mathscr{X}_n\}$, la distribución de las vv.aa. depende únicamente de la distribución de la v.a. inicial y de la matriz de transición $\mathbf{P}$.

La siguiente propiedad es de gran relevancia en la teoría de las CC.M.. Por consiguiente, no podía faltar su presencia en este trabajo:

Propiedad 2.2.1 (Propiedad de *Markov*)**.** *Un proceso estocástico* $\{\mathscr{X}_n\}$ *con* S *un conjunto finito o numerable, se dice que tiene la* ***propiedad de Markov*** *si para cada* $k, n \in \mathbb{N} \cup \{0\}$ *tales que (tt.q.)* $k \leq n$*, y para cualquier elección de estados* $i_0, \dots, i_{n+1} \in S$*, tenemos que:*

$$\mathbb{P}(\mathscr{X}_{n+1} = i_{n+1} / \mathscr{X}_n = i_n, \dots, \mathscr{X}_{k+1} = i_{k+1}, \mathscr{X}_k = i_k) = \mathbb{P}(\mathscr{X}_{n+1} = i_{n+1} / \mathscr{X}_n = i_n).$$

La Propiedad 2.2.1 motiva la definición siguiente. Después de anunciarla, veremos porqué es equivalente a la Definición 2.2.1 que hemos introducido al inicio de esta sección:

Definición 2.2.5. *Todo proceso estocástico* $\{\mathscr{X}_n\}$ *que cumple la propiedad de Markov, es una C.M. con* S *su espacio de estados.*

La Definición 2.2.1 nos ha servido para definir las CC.M., para introducir notación y para hablar sobre las matrices de transición.

La Propiedad 2.2.1 y Definición 2.2.5 que acabamos de ver, están totalmente relacionadas con la Definición 2.2.1 inicial. En la Propiedad de *Markov* se resume con una ecuación que, en cada paso $n \in \mathbb{N} \cup \{0\}$ (cada vez que repetimos el experimento), la probabilidad de que ocurra $\mathscr{X}_{n+1}$ solo se ve influida por la de que ocurra $\mathscr{X}_n$, sin importar los valores tomados por las anteriores vv.aa.. Estas probabilidades condicionadas se representan mediante la matriz de transición. De ahí que, en la definición inicial, se muestren las CC.M. como la existencia del conjunto de valores $\{p_{ij}(n+1)\}$ que representan esta misma matriz de transición.

Notemos cómo la frase inicial de la introducción[7] de este capítulo va cobrando sentido...

2.3 Cadenas de Markov Homogéneas

En esta sección veremos un tipo concreto de CC.M. que serán la base de todo lo que mostraremos en los Capítulos 3 y 4. Veamos la definición primero de todo:

Definición 2.3.1. *Una C.M., con* S *un espacio de estados finito o numerable (e.e.f.n.), es* ***homogénea*** *si su matriz de transición cumple que:*

$$\mathbf{P} = \mathbf{P}(n), \quad \forall n \in \mathbb{N} \cup \{0\}.$$

Observación 2.3.1. *En este caso, para cada* $n, k \in \mathbb{N} \cup \{0\}$ *tendríamos que las densidades de masa* $\pi(k)$ *y* $\pi(n+k)$*, de* $\mathscr{X}_k$ *y* $\mathscr{X}_{n+k}$ *respectivamente, estarían relacionadas por la siguiente igualdad:*

$$\pi(n+k) = (\mathbf{P}^n)\pi(k),$$

debido a la Proposición 2.2.2.

[7] Véase el primer parágrafo de la Sección 2.1 para recordar la frase.

2.4 Representación canónica de P

En esta sección vamos a caracterizar un subconjunto de estados, los recurrentes y los transitorios, del espacio de estados S, que resultará de gran interés y necesidad para la representación de **P** que usaremos en el Capítulo 4 (forma canónica de **P**).

Gracias a esta clase de estados, podremos escribir nuestra matriz de transición de forma que se podrá extraer de ella información mucho más relevante, en algunos casos, que de la matriz **P** original.

De ahora en adelante (a no ser que digamos lo contrario en la definición de algún parámetro) supondremos lo siguiente:

- La sucesión de vv.aa. $\{\mathscr{X}_n\}$ es una Cadena de Markov Homogénea (C.M.H.).
- El espacio de probabilidad sobre el que estarán definidas será $(\Omega, \mathscr{E}, \mathbb{P})$.
- Denotamos por S al espacio de estados.
- Sea $C \subset S$, si es un conjunto unitario, o sea, $C = \{j\}$, por abuso de notación usaremos j directamente.
- $\mathbf{P} = (p_{ij})$ es la matriz de transición de la C.M.H..
- Si en algún momento no especificamos el camino definido por el suceso x (véase la Observación 2.2.1), se sobreentiende que las vv.aa. $\mathscr{X}_n$ se aplican a sucesos $x \in \Omega$.

Para alcanzar el objetivo fijado, empecemos con la definición de **clase cerrada** o **clase absorbente** de estados:

Definición 2.4.1. *Sea $C \neq \emptyset$ t.q. $C \subset S$. Decimos que C es una **clase cerrada** o **clase absorbente** para* **P** *si para cualquier $i \in C$ y para cualquier $j \notin C$, j no es accesible desde i ($i \nrightarrow j$).*

*Equivalentemente, tenemos que C es una **clase cerrada** sii $i \in C$ y $i \longrightarrow j$ implica que $j \in C$.*

En otras palabras, C es una clase cerrada si no hay caminos que empiecen desde cualquier estado de C y lleguen a un estado que no esté en C. Notemos que los estados de C no tienen porque estar conectados dos a dos, podrían ser cuatro estados que se conectan entre sí por caminos pero sin conectarse dos a dos.

La manera más simple de encontrar las clases cerradas de S es considerar, para cada $i \in S$, el conjunto de estados accesibles desde i:

$$\{j \in S \,/\, i \longrightarrow j\}.$$

Estos conjuntos serían clases cerradas. Veamos esto mejor con un ejemplo:

Ejemplo 2.4.1. *Sea*

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

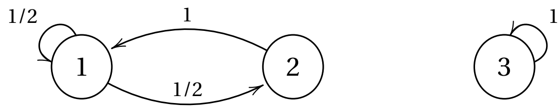

Figura 2.4: Diagrama de estados de la matriz de transición **P** del Ejemplo 2.4.1.

una matriz de transición de una C.M.H. con $S = \{1,2,3\}$ *su conjunto de estados. El diagrama de estados para* **P** *sería el dado en la Figura 2.4.*

Veamos qué clases serían absorbentes:

- *Sea* $C_1 = \{1\}$. *No sería una clase cerrada porque* $1 \in C_1$, $1 \longrightarrow 2$ *y* $2 \notin C_1$.
- *Sea* $C_2 = \{1,2\}$. *Sí sería una clase cerrada porque el estado* $3 \notin C_2$ *no es accesible desde los estados* $1,2 \in C_2$.
- *Sea* $C_3 = \{3\}$. *Sí sería una clase cerrada porque los estados* $1,2 \notin C_3$ *no son accesibles desde* $3 \in C_3$.
- *Sea* $C_4 = S$. *Sería una clase cerrada de forma trivial ya que no existe un estado* $i \notin C_4$.

En general, S siempre será una clase cerrada para **P**. La siguiente proposición es crucial para introducir una clase cerrada de estados especial, la denominada minimal.

Proposición 2.4.1. *Sean C y D clases cerradas. Entonces $C \cup D$ y $C \cap D$ también lo son.*

Esto quiere decir que las clases cerradas son un conjunto parcialmente ordenado[8] *(poset) con la inclusión, denotada por* $\subseteq$, *como orden parcial.*[9]

Demostración. *Consúltese la Demostración B.1.8 del Apéndice B.*

Gracias a que las clases cerradas son un poset, contamos con la siguiente definición:

Definición 2.4.2. *Sea $C \subseteq S$ una clase cerrada. Decimos que C es una **clase cerrada minimal** si*

$$D \subseteq C \Rightarrow C = D, \quad \forall D \subseteq S \text{ clase cerrada.}$$

Como consecuencia de esta definición, tenemos que las clases cerradas minimales de S son disjuntas dos a dos.

Ejemplo 2.4.2. *En el Ejemplo 2.4.1 anterior, las clases cerradas minimales serían* $\{1,2\}$ *y* $\{3\}$.

Para detectar las clases cerradas minimales de una forma más práctica y no tan intuitiva, contamos con la siguiente proposición:

[8] Véase la Definición A.5.4 del Apéndice A , si al lector le interesa.
[9] Véase la Definición A.5.3 del Apéndice A.

Proposición 2.4.2. *Una clase cerrada C es minimal sii todos sus elementos están comunicados dos a dos (o sea, si $i, j \in C$, entonces $i \longleftrightarrow j$, $\forall i, j \in C$).*

Demostración. *Consúltese la Demostración B.1.9 en el Apéndice B.*

Con esta caracterización de las clases cerradas minimales, tenemos que si $\mathbf{P}$ es una matriz irreducible (véase la Definición 2.1.7), entonces S es una clase cerrada minimal.

El siguiente ejemplo ilustra el concepto de clase cerrada minimal:

Ejemplo 2.4.3. *Sea*

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

una matriz de transición de una C.M.H. con $S = \{1, 2, 3\}$ su conjunto de estados. El diagrama de estados para $\mathbf{P}$ sería el dado en la Figura 2.5.

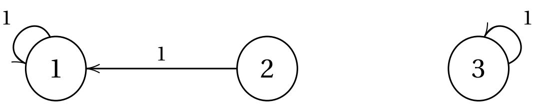

Figura 2.5: Diagrama de estados de la matriz de transición $\mathbf{P}$ del Ejemplo 2.4.3.

Identificaremos las clases cerradas y veremos cuáles son minimales:

- *Por la Proposición 2.4.1, S es una clase cerrada trivialmente.*
- *Sea $i = 1$, el conjunto de estados accesibles desde i es $C_1 = \{1\}$. Por tanto C_1 es una clase cerrada.*
- *Sea $i = 2$, el conjunto de estados accesibles desde i es $C_1 = \{1\}$.*
- *Sea $i = 3$, el conjunto de estados accesibles desde i es $C_2 = \{3\}$. Por tanto C_2 es una clase cerrada.*
- *También tendríamos que $C_3 = \{1, 3\}$ sería una clase cerrada, debido a que $\forall i \in C_3$ y $j \notin C_3$, o sea, $j = 2$, j no es accesible desde i.*

En definitiva, una vez tenemos identificadas las clases cerradas, tendríamos el siguiente árbol con raíz el conjunto S y hojas las que serían las clases cerradas minimales:

$$\begin{tikzcd}
& \{1,2,3\} \arrow[dl] \arrow[dr] & & \\
\{1\} & & \{1,3\} \arrow[dr] & \\
& & & \{3\}
\end{tikzcd}$$

Esto quiere decir que C_1 y C_2 son clases cerradas minimales.

Otra forma de detectar las clases cerradas minimales es a través de la matriz **B** (véase la Ecuación 2.1). Veámoslo:

Proposición 2.4.3. *El estado $i \in S$ pertenece a una clase cerrada minimal sii para cualquier $j \in S$ t.q. $\mathbf{B}^i_j > 0$, tenemos que $\mathbf{B}^j_i > 0$.*

Demostración. *Obvio gracias a la Proposición 2.4.2. Debido a que, por esta proposición, una clase cerrada $C = \{j \in S \,/\, \mathbf{B}^i_j > 0\}$ que contenga a i es minimal sii todos los elementos de C están comunicados dos a dos, en particular para este i:*

$$C = \{j \in S \,/\, \mathbf{B}^i_j > 0\} = \{j \in S \,/\, \mathbf{B}^i_j > 0, \mathbf{B}^j_i > 0\}.$$

□

Una vez tenemos bien claro cómo podemos encontrar una clase cerrada minimal, vayamos con un tipo de C.M. que es sobre la cuál nos interesará profundizar para poder desarrollar una de las aplicaciones en el Capítulo 4:

Definición 2.4.3 (Cadena de Markov absorbente)**.** *Diremos que una C.M. es **absorbente** si cuenta con al menos una clase cerrada minimal y t.q. desde cualquier estado de S es posible llegar a alguna de las clases cerradas minimales de la cadena en un número finito de pasos.*

Ejemplo 2.4.4. *La matriz de transición del Ejemplo 2.4.3 proviene de una Cadena de Markov Absorbente (C.M.A.). Ya que los estados* 1 *y* 3 *forman las dos clases cerradas minimales C_1 y C_2 respectivamente, y desde el estado* 2 *podemos llegar a C_1 en solo* 1 *paso.*

En el resto de esta sección supondremos que $\{\mathcal{X}_n\}$ es una C.M.A.. La notación seguirá como hasta ahora, solo añadimos esta nueva condición sobre la sucesión $\{\mathcal{X}_n\}$.

Al principio de esta sección hemos expuesto que podríamos encontrar una representación de la matriz **P**. Gracias a las definiciones y proposiciones con las que ahora ya contamos, veamos a qué representación nos estábamos refiriendo con el siguiente teorema:

Teorema 2.4.1 (Representación canónica)**.** *Sean $C_1, \ldots, C_r$ las clases cerradas minimales para* **P***. Sea k_l el cardinal de C_l para cada $l \in \{1, \ldots, r\}$, $C = \{C_1 \cup \cdots \cup C_r\}$ y $T = S \setminus C$. Definimos $K = k_1 + \cdots + k_r$.*

Entonces la matriz estocástica **P** *tiene la **forma canónica** o **representación canónica** siguiente:*

$$\mathbf{P} = \begin{pmatrix} (\mathbf{B}_1) & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & (\mathbf{B}_2) & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & (\mathbf{B}_r) & 0 \\ & & \mathbf{R} & & & \mathbf{V} \end{pmatrix}_{N \times N} \tag{2.3}$$

donde $(\mathbf{B}_1), \ldots, (\mathbf{B}_r)$ son matrices estocásticas irreducibles de dimensión $k_1 \times k_1, \ldots, k_r \times k_r$ respectivamente tt.q. $(\mathbf{B}_l) = (p_{ij})_{i,j \in C_l}$, $\forall\, l \in \{1, \ldots, r\}$ y **R** *y* **V** *son las submatrices de* **P** *de dimensiones $(N-K) \times K$ y $(N-K) \times (N-K)$ respectivamente.*

Para obtener la representación canónica de la matriz **P** debemos identificar las clases cerradas minimales. Una vez las tenemos, reetiquetamos (cambiar filas y columnas) los estados para que nos quede una matriz con la forma (2.3). Cada ($\mathbf{B}_l$) representará una clase cerrada minimal C_l, con $l \in \{1, \ldots, r\}$.

La información que transmite la matriz **P** original y su respectiva forma canónica es la misma. La diferencia es que si **P** está expresada en forma canónica, nos es más fácil identificar los tipos de estados y así podremos analizar mejor los estados de **P**.

A continuación, mostraremos las definiciones de los diferentes tipos de estados: los estados recurrentes y los estados transitorios.

Definición 2.4.4. *Sea $j \in S$ un estado:*

- *Diremos que j es un estado **recurrente** si*

$$\sum_{n=1}^{+\infty} p_{jj}^{(n)} = +\infty.$$

Esto nos vendría a decir que habiendo visitado el estado j en $\mathscr{X}_0$, es muy probable que lo volvamos a visitar en posteriores pasos $n \in \mathbb{N}$.

- *Por lo contrario, diremos que j es un estado **transitorio** si*

$$\sum_{n=1}^{+\infty} p_{jj}^{(n)} < +\infty.$$

La siguiente proposición nos ayudará a identificar qué estados son recurrentes y cuáles son transitorios dependiendo de si los estados son de C (definido en el enunciado del Teorema 2.4.1 anterior) o no:

Proposición 2.4.4. *Sean $C_1, \ldots, C_r$ las clases cerradas minimales en S, $T = S \setminus C$ e $i, j \in S$. Entonces se cumple:*

i) Si $j \in C$, entonces j es recurrente.

Y, por la Proposición A.10.5, tendríamos que si $i \longrightarrow j$, entonces $\sum_{k=1}^{+\infty} p_{ij}^{(k)} = +\infty$.

ii) Si $j \in T$, entonces j es transitorio.

Más precisamente, $\forall i, j \in T$, $p_{ij}^{(k)}$ converge a 0 tan rápido como $k \to +\infty$[10]. Por tanto $\sum_{k=1}^{+\infty} p_{ij}^{(k)} < +\infty$.

Demostración. *Consúltese la Demostración B.1.10 del Apéndice B.*

[10]Explicamos esto en detalle en la Nota A.10.2 del Apéndice A.

Gracias a esta proposición, podemos asegurar que los primeros K estados de la forma canónica de **P** son recurrentes, o sea, los correspondientes a las filas y columnas de la matriz diagonal por bloques:

$$\begin{pmatrix} (\mathbf{B}_1) & 0 & \cdots & \cdots & 0 \\ 0 & (\mathbf{B}_2) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & (\mathbf{B}_r) \end{pmatrix}_{K\times K}$$

Y los $N-K$ restantes son transitorios.

Si al lector le interesa, tenemos un ejemplo de cálculo de la representación canónica de una matriz estocástica **P** en la Sección A.11 del Apéndice A.

A la luz de lo expuesto, el siguiente resultado permite establecer el comportamiento a lo largo del tiempo de los distintos estados, tal y como ilustraremos en el Capítulo 4:

Teorema 2.4.2. *Sean $C_1,\ldots,C_r$ las clases cerradas minimales en S para la matriz estocástica **P** y la representación canónica dada por (2.3). Para cada $l \in \{1,\ldots,r\}$, sea k_l el cardinal de C_l. Tenemos que*

$$(\mathbf{P}^n) = \begin{pmatrix} (\mathbf{B}_1)^n & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & (\mathbf{B}_2)^n & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & (\mathbf{B}_r)^n & 0 \\ & & (\mathbf{R}_n) & & & (\mathbf{V}^n) \end{pmatrix}_{N\times N}$$

*donde: $(\mathbf{B}_1)^n,\ldots,(\mathbf{B}_r)^n$ y $(\mathbf{V}^n)$ son las potencias n-ésimas de las matrices $(\mathbf{B}_1),\ldots,(\mathbf{B}_r)$ y **V** del Teorema 2.4.1 respectivamente, $(\mathbf{R}_n)$[11] es una matriz $(N-K)\times K$.*

Demostración. *Consúltese la Demostración B.1.11 del Apéndice B.*

A continuación, calcularemos $\sum_{n=1}^{+\infty}(\mathbf{P}^n)$ para relacionar los valores resultantes con los estados recurrentes y transitorios. Por lo que acabamos de ver, esta serie se corresponde con:

[11] No confundir $(\mathbf{R}^n)$ con $(\mathbf{R}_n)$. Véase la Demostración B.1.11 del Apéndice B.

$$\sum_{n=1}^{+\infty}(\mathbf{P}^n) = \begin{pmatrix} \sum_{n=1}^{+\infty}(\mathbf{B}_1)^n & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \sum_{n=1}^{+\infty}(\mathbf{B}_2)^n & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \sum_{n=1}^{+\infty}(\mathbf{B}_r)^n & 0 \\ & & (\mathbf{R}_\infty) & & & \sum_{n=1}^{+\infty}(\mathbf{V}^n) \end{pmatrix}_{N\times N}$$

donde:

- $$(\mathbf{R}_\infty)^i_j = \begin{cases} +\infty & \text{si } i \longrightarrow j \\ 0 & \text{en otro caso} \end{cases}$$

  Ya que como $j \in C$, por el apartado $i)$ de la Proposición 2.4.4, tenemos que j es recurrente, y si $i \longrightarrow j$, entonces $\sum_{n=1}^{+\infty} p_{ij}^{(n)} = +\infty$. Por otro lado, si $i \nrightarrow j$ entonces el valor de $(\mathbf{R}_\infty)^i_j$ será 0.

- $$\sum_{n=1}^{+\infty}((\mathbf{B}_l)^n) = +\infty$$

  para cada $l \in \{1,\ldots,r\}$.

  Debido a que los estados i, j correspondientes a las filas y columnas de los elementos de las matrices $(\mathbf{B}_l)$, cumplen que $j \in C$ e $i \longrightarrow j$, por lo que $\sum_{n=1}^{+\infty} p_{ij}^{(n)} = +\infty$ por el apartado $i)$ de la Proposición 2.4.4 de nuevo.

- $$\sum_{n=1}^{+\infty}(\mathbf{V}^n) = (\mathbf{Id}-\mathbf{V})^{-1}\mathbf{V}, \tag{2.4}$$

  Demostración. *Consúltese la Demostración B.1.12 del Apéndice B.*

En definitiva:

$$\sum_{n=1}^{+\infty}(\mathbf{P}^n) = \begin{pmatrix} +\infty & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & +\infty & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & +\infty & 0 \\ & & (\mathbf{R}_\infty) & & & (\mathbf{Id}-\mathbf{V})^{-1}\mathbf{V} \end{pmatrix}_{N\times N} . \tag{2.5}$$

Ahora, por la Ecuación (A.26), tenemos que el elemento $\sum_{n=0}^{+\infty} p_{ij}^{(n)}$ (o sea, el elemento en la posición (i,j) de la matriz dada por (2.5)) es el número esperado de veces que, en total, visitaremos el estado j empezando desde el estado i.

Sea $\mathbf{N} = (\mathbf{Id} - \mathbf{V})^{-1}\mathbf{V}$[12], por esto que acabamos de explicar, la suma de los elementos de la fila i-ésima de **N** nos diría el número esperado de veces, en total, que estaremos en un estado transitorio j cualquiera, si empezamos desde el estado transitorio i. En otras palabras, $\sum_{j\in T} \mathbf{N}_j^i$ nos dice el número esperado de pasos totales antes de llegar a un estado recurrente si empezamos en el estado transitorio i.

Con la notación que hemos utilizado hasta ahora, exponemos el siguiente teorema que trata sobre la información que nos proporciona la matriz $\sum_{n=1}^{+\infty}(\mathbf{P}^n)$:

Teorema 2.4.3. *Sea* $\mathbf{P} = (p_{ij})_{i,j\in S}$ *la matriz de transición en su forma canónica.*

Sea a_{ij} *la probabilidad de que el camino definido por x visite el estado recurrente* j *si hemos empezado en el estado transitorio* i*. Entonces* a_{ij} *es el elemento en la posición* (i,j) *de la matriz* $\mathbf{A} = \mathbf{NR}$*.*

Demostración. *Consúltese la Demostración B.1.13 del Apéndice B.*

En definitiva, la forma canónica de la matriz de transición de una C.M.H. nos da mucha información relevante sobre el comportamiento de los diferentes sucesos del experimento que tratamos. Pero, para poder encontrarla, debemos identificar primero las clases cerradas minimales para así saber qué estados son recurrentes y qué estados son transitorios. Para asegurar la existencia de al menos una clase cerrada minimal debemos considerar Cadenas de Markov Absorbentes (CC.M.AA.), de ahí que las hayamos introducido en esta sección.

Esta tarea, la de identificar las clases cerradas minimales, puede ser asequible si la matriz estocástica es finita, debido a que podemos calcular **B** (utilizando un programa como *Octave*) y usar la Proposición 2.4.3. Una vez tenemos identificados qué estados pertenecen a clases cerradas minimales y qué estados no, gracias a la Proposición 2.4.4 averiguamos los que son recurrentes y transitorios para finalmente construir nuestra forma canónica reetiquetando los estados si hace falta. De ahí que en esta sección hayamos trabajado con matrices estocásticas finitas.

Veremos la aplicación de esta teoría en el Caso Real 4.2 del Capítulo 4.

[12] Conocida como la **matriz fundamental de la C.M.**.

CAPÍTULO 3

CONVERGENCIA DE LAS MATRICES REGULARES ESTOCÁSTICAS

En este capítulo nos centraremos en averiguar si las potencias de la matriz de transición **P** convergen a algún vector, y, si lo hacen, a qué convergen exactamente y bajo qué condiciones.

Seguiremos utilizando las suposiciones y notaciones que ya hemos introducido y explicado en el capítulo anterior (a no ser que se exponga lo contrario en algún apartado concreto del capítulo).

Antes de seguir con la lectura de este capítulo, recomendamos al lector que consulte la Sección A.7 del Apéndice A en la que hacemos una breve introducción a la teoría de las aplicaciones iteradas y exponemos una serie de condiciones y teoremas famosos sobre la existencia de puntos fijos. Igualmente, iremos haciendo referencia a cada uno de ellos por si no se consulta dicho apéndice.

3.1 Existencia de puntos fijos

Primero veamos una descripción del conjunto $\mathscr{T} \subset \mathbb{R}^N$ de los vectores estocásticos (ver Definición 2.1.4):

Proposición 3.1.1. *El conjunto $\mathscr{T}$ es convexo y compacto.*

Demostración. *Es convexo debido a la Proposición 2.1.3. La demostración de que es compacto la podemos encontrar en la Sección B.2 del Apéndice B (Demostración B.2.1).*

A partir de ahora, consideraremos $(\mathscr{T}, \|\cdot\|_1)$ como el espacio normado en $\mathbb{R}^N$.

El siguiente teorema nos asegura la existencia de un punto fijo para P:

Teorema 3.1.1 (Teorema de *Perron-Frobenius*)**.** *Existe al menos un vector estocástico* $\mathbf{w}$ *t.q.* $\mathbf{w} = \mathbf{wP}$.

Demostración. *Veamos primero de todo que* P *es continua:*

Recordemos que $\mathrm{P}(\mathbf{x}) = \mathbf{xP}$. *Por ser* $\mathbf{x} \in \mathcal{T}$ *y* $\mathbf{P}$ *una matriz estocástica finita, tenemos que* P *es una aplicación lineal entre espacios normados con* $\mathbf{x}$ *y* $\mathbf{xP}$ *acotados. Esto quiere decir que* P *es una aplicación continua.*

Ahora, como P *es continua y* $\mathcal{T}$ *es compacto y convexo, podemos aplicar el Teorema de* Brouwer[1] *y tenemos que existe un punto fijo* $\mathbf{w}$ *para* P*, o sea:*

$$\mathbf{w} = \mathbf{wP}.$$

□

El siguiente ejemplo ilustra el teorema precedente:

Ejemplo 3.1.1. *Sea*

$$\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

la matriz de transición de una C.M.H.. Queremos encontrar un punto fijo para esta matriz, que ya hemos visto que debe existir. Para hacerlo resolveremos este sistema:

$$\begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a & b \end{pmatrix}.$$

Obtenemos que $a = b$*, y como* $(a,b) \in \mathcal{T}$*, tenemos que* $a = b = 1/2$*. Por tanto, el vector* $\mathbf{w} = (1/2, 1/2)$ *es el punto fijo para la aplicación* P.

Además, como esta matriz es muy simple, podemos calcular $(\mathbf{P}^n)$*, con* $n \in \mathbb{N} \cup \{0\}$*, y tenemos que:*

$$(\mathbf{P}^n) = \begin{cases} \mathbf{Id} & \text{si } n \text{ es par} \\ \mathbf{P} & \text{si } n \text{ es impar.} \end{cases}$$

Esto significa que para $\mathbf{x} \in \mathcal{T}$ *arbitrario pero fijado,* $\mathrm{P}^n(\mathbf{x}) = \mathbf{x}(\mathbf{P}^n)$ *no convergirá nunca a* $\mathbf{w}$ *a no ser que* $\mathbf{x} = \mathbf{w}$*. Por tanto* $\mathbf{w}$ *no es un punto fijo atractivo*[2] *para* P.

A la luz del ejemplo anterior, introducimos la noción de equilibrio:

Definición 3.1.1 (Cadenas de Markov en equilibrio)**.** *Decimos que una C.M. está* ***en equilibrio*** *si existe un punto fijo para la aplicación* P.

Fijémonos que cada vez vamos precisando más y más el tipo de CC.M. que estamos buscando para poder llegar a nuestro principal objetivo. Con la notación y suposiciones que llevamos hechas hasta el momento, podemos asegurar, por el Teorema 3.1.1 que acabamos de ver, que siempre existirá un punto fijo para P, y, por tanto, estamos considerando una C.M. en equilibrio.

[1] Véase el Teorema A.7.1 del Apéndice A.

[2] Véase la Definición A.7.4 del Apéndice A.

3.2 Convergencia de las potencias de P

En esta sección nos centraremos en ver bajo qué condiciones podemos encontrar un punto fijo atractivo para P. Ya hemos visto en el ejemplo de la sección anterior que no siempre tiene porqué existir.

La siguiente definición es de las más importantes de este trabajo:

Definición 3.2.1 (Cadena de Markov Regular)**.** *La matriz estocástica finita* **P** *es* ***regular*** *si existe* $k_0 \in \mathbb{N}$ *t.q.*

$$(\mathbf{P}^{k_0})^i_j \neq 0, \ \forall i,j \in S.$$

Decimos que la sucesión $\{\mathscr{X}_n\}$ *es una Cadena de Markov Regular (C.M.R.) si la matriz de transición* **P** *es regular.*

Observación 3.2.1. *Si la matriz* **P** *es regular, obviamente es irreducible, debido a que si hay una potencia de la matriz para la cual todos los valores son diferentes de* 0*, esto quiere decir que todos los estados están comunicados dos a dos.*

Al final de esta sección veremos que la condición necesaria y suficiente para que P tenga un punto fijo atractivo es que la matriz **P** sea regular. Pero para poder demostrar esto, debemos enunciar y demostrar la siguiente proposición primero:

Proposición 3.2.1. *Sean* $\mathbf{R}^1,\ldots,\mathbf{R}^N$ *las filas de* **P**, $\mathbf{K} \subset \mathbb{R}^N$ *la envoltura convexa* [3] *del conjunto de estos vectores y*

$$C = \frac{1}{2}\max_{i,j\in S}\{\|\mathbf{R}^i - \mathbf{R}^j\|_1\}.$$

Entonces:

i) $C \leq 1$.

ii) Si todos los valores de **P** *son diferentes de* 0*, entonces* $C < 1$.

iii) $diam(\mathbf{K}) = \sup_{\mathbf{x},\mathbf{y}\in\mathbf{K}}\{\|\mathbf{x}-\mathbf{y}\|_1\} = 2C$.

iv) $\|\mathrm{P}(\mathbf{x}) - \mathrm{P}(\mathbf{y})\|_1 \leq C \cdot \|\mathbf{x}-\mathbf{y}\|_1, \ \forall x,y \in \mathscr{T}$.

Demostración. *Esta demostración es muy interesante, pero por temas de espacio, hemos tenido que reubicarla en la Sección B.2 del Apéndice B (Demostración B.2.2).*

Antes de mostrar el teorema principal de este trabajo, debemos mostrar la siguiente proposición, que nos ayudará a asegurar que las componentes del vector **w** son todas positivas:

[3] Véase la Definición A.1.6 del Apéndice A si al lector le interesa.

Proposición 3.2.2. *Sea* **P** *irreducible y* $j \in S$. *Entonces,*

$$\frac{1}{n}\sum_{k=1}^{n} p_{jj}^{(k)} \to \lambda \Rightarrow \lambda > 0.$$

En particular,

$$p_{jj}^{(n)} \to \lambda \Rightarrow \lambda > 0.$$

Demostración. *Consúltese en la Demostración B.2.3 del Apéndice B.*

Ahora sí, vayamos con uno de los teoremas más relevantes de este trabajo debido a que proporciona las condiciones suficiente para asegurar la existencia de un vector estocástico atractivo de la C.M.:

Teorema 3.2.1. *Los siguientes apartados son equivalentes:*

i) **P** *es regular.*

ii) **P** *es irreducible y la aplicación* P *tiene un punto fijo atractivo* $\mathbf{w} \in \mathcal{T}$ *al que llamaremos* ***estado de equilibrio***.

iii) $\exists \overline{n} \in \mathbb{N}$ *t.q.* $\forall n \geq \overline{n}$ *todos los valores de* $\mathbf{P}^n$ *no son cero.*

Además, si se cumple uno de estos apartados, tenemos que:

a) $\mathbf{w}$ *es el único punto fijo de la aplicación* P.

b) Todas las componentes de $\mathbf{w}$ *son diferentes de* 0.

c) Por ser **P** *regular, existe* $k_0 \in \mathbb{N}$ *t.q.* $\mathbf{P}^{k_0}$ *tiene todos sus valores mayores que* 0. *Sean* $\mathbf{R}^1, \ldots, \mathbf{R}^N$ *las filas de esta matriz* $\mathbf{P}^{k_0}$. *Entonces:*

$$C = \frac{1}{2}\max_{i,j \in S}\{\|\mathbf{R}^i - \mathbf{R}^j\|_1\} < 1,$$

y, para cualquier $\mathbf{x} \in \mathcal{T}$*:*

$$\|\mathbf{x}\mathbf{P}^n - \mathbf{w}\|_1 \leq 2C^{\lfloor n/k_0 \rfloor},\ \forall n \geq k_0 \tag{3.1}$$

Esta desigualdad denota la velocidad de convergencia. Notemos que converge a 0 *exponencialmente a medida que* n *crece, ya que* $C < 1$.

Demostración. *Primero demostraremos las equivalencias entre los tres apartados, y después veremos que se cumplen las tres propiedades sobre* $\mathbf{w}$*:*

- *Vamos primero a demostrar que* $i) \Rightarrow ii)$*:*

  Ya hemos explicado anteriormente (en la Observación 3.2.1), que si una matriz es regular, entonces es irreducible. Sea $(\mathbf{P}^{k_0})$ *la potencia de* **P** *con todos sus valores mayores que* 0. *Queremos ver que* P *tiene un único punto fijo atractivo.*

Por ser **P** *estocástica,* $(\mathbf{P}^{k_0})$ *es también estocástica.*[4] *Ahora, por los apartados* $ii)$ *y* $iv)$ *de la Proposición 3.2.1 anterior, tenemos que* P^{k_0} *es una contracción. Además, sabemos que* $(\mathbb{R}^N, d_1)$[5] *es un espacio métrico completo,*[6] *y como* $\mathcal{T} \subset \mathbb{R}^N$ *es cerrado, porque es compacto, tenemos que* $(\mathcal{T}, d_1)$ *es también un espacio métrico completo.*

Sabiendo esto, podemos aplicar el teorema del punto fijo de Banach,[7] *y así tenemos que* P^{k_0} *tiene un único punto fijo atractivo,* **w***, en* $\mathcal{T}$*. Ahora, por el apartado* $vi)$ *de la Proposición A.7.1, tenemos que* **w** *es el único punto fijo atractivo para* P.

- *Veamos que* $ii) \Rightarrow iii)$*:*

  Al ser **w** *el punto fijo atractivo de* P*, tenemos que* $\mathbf{x}(\mathbf{P}^n) \to \mathbf{w}$, $\forall \mathbf{x} \in \mathcal{T}$*. En particular, para los vectores* $\mathbf{e^i} \in \mathcal{T}$ *que componen la base canónica de* $\mathbb{R}^N$*:*

  $$\mathbf{e}^i(\mathbf{P}^n) \to \mathbf{w},\ \forall i \in S \Rightarrow (\mathbf{P}^n)^i \to \mathbf{w},\ \forall i \in S.$$

  Esto quiere decir que cada una de las filas de $(\mathbf{P}^n)$ *converge a* **w***. Ahora, por la propiedad b) de este teorema, tenemos que las componentes de* **w** *son todas mayores que* 0*. Esto quiere decir que para un* n *suficientemente grande, las filas de* $(\mathbf{P}^n)$ *son* **w***, y como las componentes de* **w** *son todas mayores que* 0*, tenemos que las componentes de las filas de* $(\mathbf{P}^n)$ *son todas mayores que* 0*. Esto significa que*

  $$\exists \overline{n} \in \mathbb{N},\ \textit{t.q.}\ (\mathbf{P}^n)^i_j > 0,\ \forall i, j \in S \text{ y } \forall n \geq \overline{n}.$$

- *Por último, queremos ver que* $iii) \Rightarrow i)$*, pero esta implicación es obvia por propia definición de matriz regular.*

Veamos ahora que si se cumple uno de los tres apartados anteriores, entonces se cumplen las tres propiedades sobre **w** *que mostraremos a continuación:*

a) Si se cumple el apartado $ii)$ *de este teorema, tenemos que* **w** *es un punto fijo atractivo para* P*, y por el apartado* $iv)$ *de la Proposición A.7.1, tendríamos que* **w** *es el único punto fijo de* P.

b) En la demostración de que el apartado $ii)$ *implica el apartado* $iii)$ *de este teorema, hemos visto que las filas de la matriz* $(\mathbf{P}^n)$ *convergen todas al vector* **w***. Esto quiere decir que* $(\mathbf{P}^n) \to \mathbf{W}$*.*[8] *Entonces tenemos, en particular, que* $p^{(n)}_{jj} \to w_j$, $\forall j \in S$*. Ahora, si se cumple el apartado* $ii)$ *de este teorema, tenemos que* **P** *es irreducible, y por la Proposición 3.2.2, tendríamos que* $w_j > 0$, $\forall j \in S$*. De esta forma quedaría demostrado que el vector* **w** *tiene todas sus componentes estrictamente positivas.*

c) Si se cumple el apartado $i)$ *del teorema, tenemos que* **P** *es regular y por tanto* $\exists k_0 \in \mathbb{N}$ *t.q.* $(\mathbf{P}^{k_0})^i_j > 0$, $\forall i, j \in S$*. Como* $\mathcal{T}$ *es un conjunto convexo,* $\mathcal{T}$ *es su propia*

[4] Por la Proposición 2.1.2 ya vista.
[5] Véase el Ejemplo A.1.4 del Apéndice A para más información sobre la distancia d_1.
[6] Véase el Ejemplo A.4.2 del Apéndice A.
[7] Véase el Teorema A.7.2 del Apéndice A para ver el enunciado.
[8] En la Subsección 3.2.1 hablaremos de la matriz **W** en profundidad.

envoltura convexa. Sean $\mathbf{R}^1, \ldots, \mathbf{R}^N$ *las filas de la matriz* $(\mathbf{P}^{k_0})$*, por el apartado* $ii)$ *de la Proposición 3.2.1, tenemos que* $C < 1$*, donde* C *es:*

$$C = \frac{1}{2} \max_{i,j \in S} \{\|\mathbf{R}^i - \mathbf{R}^j\|_1\}.$$

Por otro lado, de la demostración de que el apartado $i)$ *implica el apartado* $ii)$ *de este teorema, hemos visto que* $(\mathcal{T}, d_1)$ *es un espacio métrico completo y que* P *es una contracción. Además,* P *es continua. Por todo esto, podemos aplicar el Corolario A.7.1, y como* $\mathcal{T}$ *está también acotado, para* $n \geq k_0$*, se cumple que:*

$$\|\mathrm{P}^n(x) - \mathbf{w}\|_1 \leq diam(\mathcal{T}) \cdot C^{\lfloor n/k_0 \rfloor}.$$

Ahora, como $\mathcal{T}$ *es compacto:*

$$diam(\mathcal{T}) = \max_{x,y \in \mathcal{T}} \{\|\mathbf{x} - \mathbf{y}\|_1\} \leq \max_{x,y \in \mathcal{T}} \{\|\mathbf{x}\|_1 + \|\mathbf{y}\|_1\} = \max_{x,y \in \mathcal{T}} 2 = 2.$$

Aplicando esto a la desigualdad anterior obtenemos que:

$$diam(\mathcal{T}) \cdot C^{\lfloor n/k_0 \rfloor} \leq 2 \cdot C^{\lfloor n/k_0 \rfloor}.$$

En definitiva:

$$\|\mathrm{P}^n(x) - \mathbf{w}\|_1 \leq 2 \cdot C^{\lfloor n/k_0 \rfloor},$$

tal y como nos quedaba por ver.

□

En la Sección A.13 del Apéndice A encontraremos un ejemplo de cálculo de un punto fijo para una aplicación P y su respectiva interpretación posterior. Como en el siguiente capítulo vamos a mostrar varios casos reales, dejamos este ejemplo para consulta al lector interesado.

Una vez hemos visto este último teorema se nos puede plantear la siguiente pregunta: ¿ de qué nos sirve asegurar la existencia de un estado de equilibrio **w** para una aplicación P de una C.M.H. $\{\mathcal{X}_n\}$?

Pues bien, sea $j \in S$ arbitrario pero fijado y $n \in \mathbb{N} \cup \{0\}$ cualquiera, aplicamos el teorema de la Probabilidad Total[9] a $\mathbb{P}(\mathcal{X}_n = j)$ y obtenemos que:

$$\mathbb{P}(\mathcal{X}_n = j) = \sum_{i \in S} \mathbb{P}(\mathcal{X}_0 = i) \mathbb{P}(\mathcal{X}_n = j \,/\, \mathcal{X}_0 = i) = \sum_{i \in S} \mathbb{P}(\mathcal{X}_0 = i) (\mathbf{P}^n)^i_j \to w_j. \Rightarrow$$

$$\mathbb{P}(\mathcal{X}_n = j) \to w_j.$$

Esto significa que, sea cual sea la distribución de la v.a. inicial $\mathcal{X}_0$, la densidad de masa (véase la Definición 2.2.4) de los $\mathcal{X}_n$ converge al vector estocástico **w**. De ahí la relevancia de asegurar su existencia.

[9] Véase el Teorema A.3.1 del Apéndice A.

3.2.1 La matriz W

En lo que resta de este capítulo supondremos que la matriz $\mathbf{P}$ es regular para asegurar que $\mathbf{P}^n$ converge a un estado de equilibrio.

Como ya habíamos planteado en la demostración de $ii) \Rightarrow iii)$ en el Teorema 3.2.1, cada una de las filas de $(\mathbf{P}^n)$ convergen al vector $\mathbf{w}$. Por ello, en esta sección, vamos a introducir la matriz $\mathbf{W}$ definida como sigue:

$$\mathbf{W}^i = \mathbf{w}\begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}, \text{ para cada } i \in S.$$

O sea, todas las filas de $\mathbf{W}$ son iguales al vector $\mathbf{w}$. Esto quiere decir que $\mathbf{W}$ es una matriz $N \times N$ estocástica.

Con esta nueva notación, nos será más sencillo explicar las proposiciones y teoremas que veremos a continuación. Para empezar, podemos reescribir algunas afirmaciones del Teorema 3.2.1 como sigue:

i) $\mathbf{x}(\mathbf{P}^n) \to \mathbf{w}$, $\forall \mathbf{x} \in \mathscr{T} \Leftrightarrow (\mathbf{P}^n) \to \mathbf{W}$.

ii) La Desigualdad (3.1) quedaría así:

$$\|(\mathbf{P}^n) - \mathbf{W}\| \leq 2C^{\lfloor n/k_0 \rfloor}, \ \forall n \geq k_0.$$

iii) $\mathbf{wP} = \mathbf{w} \Leftrightarrow \mathbf{WP} = \mathbf{W}$.

Y, como todas las filas de $\mathbf{W}$ son iguales a $\mathbf{w}$ y $\mathbf{WP} = \mathbf{W}$, se cumple lo siguiente:

$$\mathbf{W} = (\mathbf{W}^2) = \mathbf{WP} = \mathbf{PW}. \tag{3.2}$$

La demostración de que estas igualdades se mantienen, la podemos encontrar en la Sección B.2 del Apéndice B (Demostración B.2.4).

3.3 Cálculo del estado de equilibrio

Aunque ya hemos adelantado un par de formas de calcular el vector $\mathbf{w}$,[10] en este apartado vamos a averiguar un nuevo método para saber a qué convergen las potencias de $\mathbf{P}$.

La necesidad de explicar este método surge del hecho de que si $\mathbf{P}$ es de dimensión elevada, calcular el punto fijo o las potencias de $\mathbf{P}$ para ver a qué convergen, puede ser una ardua tarea.

Con esta nueva técnica, basada en la descomposición de *Jordan*[11], podremos facilitar el cálculo de las potencias de $\mathbf{P}$. Veámosla:

[10] Una forma es calculando las potencias de $\mathbf{P}$ para un n suficientemente grande, y la otra manera sería calculando el punto fijo para la aplicación P.

[11] Véase la Definición A.1.11 del Apéndice A.

3.3.1 Descomposición de Jordan

Para poder encontrar la descomposición de Jordan, primero debemos intentar averiguar algo sobre los vectores propios[12] de $\mathbf{P}$. Veamos:

Proposición 3.3.1. *El vector* $\mathbf{v} = (1, \ldots, 1)$ *es un vector propio para la matriz* $\mathbf{P}$*, con valor propio* $\lambda_v = 1$.

Además, para cualquier valor propio $\lambda \in \mathbb{C}$ *de* $\mathbf{P}$*, tenemos que* $\lambda \leq 1$*, donde* $|\lambda|$ *es el módulo del número complejo* λ.[13]

Demostración. *Consúltese la Demostración B.2.5 del Apéndice B.*

Teorema 3.3.1. *El único valor propio de* $\mathbf{P}$ *con módulo* 1 *es* 1*. Sus multiplicidades algebraicas y geométricas*[14] *son* 1 *ambas. Consecuentemente, todos los valores propios de* $\mathbf{P}$*, excepto el propio* 1*, tienen módulo estrictamente menor que* 1.

Demostración. *Consúltese la Demostración B.2.6 del Apéndice B.*

Una vez contamos con estos dos resultados, podemos calcular las potencias de $\mathbf{P}$ y su límite fácilmente. Veamos por qué:

El Teorema 3.3.1 que acabamos de ver, implica que $\mathbf{P}$ tiene una descomposición de Jordan, $\mathbf{P} = \mathbf{SJS}^{-1}$, donde la primera columna de $\mathbf{S}$, es el vector propio de $\mathbf{P}$ asociado al valor propio 1, o sea, $(1, \ldots, 1)$. La primera columna es exactamente este vector propio debido a que $\mathbf{S}$ es la matriz de cambio de base y, en este caso, cambiamos de la base canónica de $\mathbf{R}^N$ a una base de vectores propios linealmente independientes y de la misma dimensión que $\mathbf{R}^N$.

Si los vectores propios no generaran todo $\mathbf{R}^N$, buscaríamos vectores linealmente independientes a ellos que lo generen y que cumplan además que $\mathbf{P} = \mathbf{SJS}^{-1}$.

Así pues,

$$\mathbf{S} = \begin{pmatrix} 1 & s_{12} & \cdots & s_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & s_{N2} & \cdots & s_{NN} \end{pmatrix} \quad \text{y} \quad \mathbf{J} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & (\mathbf{J}_1) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & (\mathbf{J}_{p-1}) & 0 \\ 0 & \cdots & \cdots & 0 & (\mathbf{J}_p) \end{pmatrix}$$

con $(\mathbf{J}_k)$, $k \in \{1, \ldots, p\}$, un bloque de Jordan t.q.

$$(\mathbf{J}_k) = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda_k & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_k \end{pmatrix}_{n_k \times n_k}$$

donde n_k es la multiplicidad algebraica del valor propio λ_k.

[12] Véase la Definición A.1.8 del Apéndice A para más información sobre los vectores propios.

[13] Véase la Definición A.6.2 del Apéndice A para más información.

[14] Véase la Definición A.1.9 del Apéndice A.

Ahora, por cumplirse que $\mathbf{P} = \mathbf{S}\mathbf{J}\mathbf{S}^{-1}$, se tiene que:

$$\begin{gathered}(\mathbf{P}^2) = \mathbf{S}\mathbf{J}\mathbf{S}^{-1}\mathbf{S}\mathbf{J}\mathbf{S}^{-1} = \mathbf{S}\mathbf{J}\mathbf{Id}\mathbf{J}\mathbf{S}^{-1} = \mathbf{S}(\mathbf{J}^2)\mathbf{S}^{-1}\\ \vdots \\ (\mathbf{P}^n) = \mathbf{S}(\mathbf{J}^n)\mathbf{S}^{-1}.\end{gathered}$$

(Podríamos comprobar fácilmente, con un argumento de inducción, que esta igualdad se cumple $\forall n \in \mathbb{N}$). Y, la matriz $(\mathbf{J}^n)$, como $\mathbf{J}$ es diagonal por bloques, tiene la forma:

$$(\mathbf{J}^n) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & (\mathbf{J}_1)^n & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & (\mathbf{J}_{p-1})^n & 0 \\ 0 & \cdots & \cdots & 0 & (\mathbf{J}_p)^n \end{pmatrix}.$$

En el Teorema 3.3.1 hemos visto que todos los valores propios tienen módulo menor estrictamente que 1, esto quiere decir que la matriz $(\mathbf{J}^n)$ converge a la matriz

$$\mathbf{K} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix},$$

cuando $n \to +\infty$.[15]

De aquí tenemos que:

$$(\mathbf{P}^n) \xrightarrow[n\to\infty]{} \mathbf{S}\mathbf{K}\mathbf{S}^{-1} =$$

$$\begin{pmatrix} 1 & s_{12} & \cdots & s_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & s_{N2} & \cdots & s_{NN} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} \begin{pmatrix} s^{-1}_{11} & s^{-1}_{12} & \cdots & s^{-1}_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ s^{-1}_{N1} & s^{-1}_{N2} & \cdots & s^{-1}_{NN} \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} \begin{pmatrix} s^{-1}_{11} & s^{-1}_{12} & \cdots & s^{-1}_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ s^{-1}_{N1} & s^{-1}_{N2} & \cdots & s^{-1}_{NN} \end{pmatrix} =$$

$$\begin{pmatrix} s^{-1}_{11} & s^{-1}_{12} & \cdots & s^{-1}_{1N} \\ \vdots & \vdots & \vdots & \vdots \\ s^{-1}_{11} & s^{-1}_{12} & \cdots & s^{-1}_{1N} \end{pmatrix}.$$

Por otro lado, de la Subsección 3.2.1 sabemos que:

[15] Consúltese la Demostración B.2.7 del Apéndice B.

$$(\mathbf{P}^n) \xrightarrow[n\to\infty]{} \mathbf{W}.$$

Esto quiere decir que:

$$\begin{pmatrix} s_{11}^{-1} & s_{12}^{-1} & \cdots & s_{1N}^{-1} \\ \vdots & \vdots & \vdots & \vdots \\ s_{11}^{-1} & s_{12}^{-1} & \cdots & s_{1N}^{-1} \end{pmatrix} = \begin{pmatrix} w_1 & w_2 & \cdots & w_N \\ \vdots & \vdots & \vdots & \vdots \\ w_1 & w_2 & \cdots & w_N \end{pmatrix}.$$

O sea:

$$w_j = s_{1j}^{-1}, \forall j \in S. \tag{3.3}$$

Además, podemos asegurar que $(\mathbf{P}^n)$ converge a $\mathbf{W}$ con una velocidad exponencial.[16]

[16] Véase la Demostración B.2.8 del Apéndice B.

CAPÍTULO 4

Casos reales en los que aplicamos la teoría

4.1 Caso real: Fútbol

4.1.1 Introducción

En la *LFP* (Liga de Fútbol Profesional) participan 20 equipos de fútbol. Cada temporada, que va de Agosto del año X a Mayo del año $X+1$[1], está compuesta por 38 jornadas, correspondientes a los 38 partidos que deberá disputar cada equipo.

En cada jornada se juegan 10 partidos, y cada partido tiene un resultado por el cuál se conoce si un equipo ha: ganado (G), empatado (E) o perdido (P). A mitad de temporada, cada equipo ha conseguido 19 resultados.

4.1.2 Construcción de la matriz de transición

En esta Subsección explicaremos cómo hemos obtenido la matriz de transición **P**:

$$\begin{pmatrix} 0{,}3 & 0{,}3 & 0{,}4 \\ 0{,}4 & 0{,}3 & 0{,}3 \\ 0{,}4 & 0{,}3 & 0{,}3 \end{pmatrix},$$

que empleamos en la Subsección 4.1.3 posterior.

La matriz **P** se corresponde con el redondeo al primer decimal de los valores que se encuentran en las columnas BD, BE y BF, y filas $32, 33$ y 34 de la hoja *Total Equipos* del archivo *Excel* adjunto: *Estadisticas futbol*.

Para obtenerla, hemos considerado el acumulado, de entre todos los equipos, de probabilidades condicionadas en las 19 primeras jornadas (mitad de temporada).

A continuación, explicaremos paso a paso cómo hemos obtenido la matriz de transición **P**. Pero primero de todo, debemos aclarar que en esta subsección cuando nos refiramos a

[1] Donde $1983 \leq X \leq 2017$ t.q. $X \in \mathbb{N}$.

filas y columnas, nos estaremos refiriendo a las filas y columnas de la hoja *Total Equipos* del archivo *Excel* adjunto: *Estadisticas futbol*, a no ser que se especifique que estamos hablando de la matriz de transición.

- Cada uno de los 20 equipos de la *LFP* 2016/2017 está representado en la primera columna entre las filas 3 y 22. Para cada una de estas filas, tenemos $19 \cdot 3 = 57$ columnas que se corresponden con los 3 estados posibles que pueden obtenerse en cada una de las 19 jornadas.

- Sea $\{P, E, G\} \longrightarrow \{1, 2, 3\}$ una biyección entre los posibles resultados de un partido para un equipo, y los tres primeros números naturales.

- Para cada fila $k \in \{3, \dots, 22\}$, si en la jornada $n \in \{1, \dots, 19\}$ ha obtenido el estado $i \in \{P, E, G\}$[2], entonces escribimos un 1 en la columna correspondiente al estado i, y dejamos en blanco las casillas correspondientes a los otros dos estados de la jornada n de la fila k.

  Repetimos este proceso para cada jornada n y cada equipo k, y así obtenemos las casillas en blanco o con un 1 que están comprendidas entre las filas 3 y 22, y las columnas B y BF.

- En la fila 23 sumamos las veces que se ha dado cada estado en cada jornada. De tal forma que, por ejemplo, en la jornada 1 vemos que ha habido 7 equipos que han perdido (P), 6 que han empatado (E) y 7 que han ganado (G). Notemos que, en cada jornada, habrá el mismo número de ganados que de perdidos.

  De las filas 28 a la 30, lo que hacemos es calcular el acumulado de P, E o G respectivamente, hasta la jornada $n-1$. Así pues, en la jornada 2, solo habrá los datos de la jornada 1, pero, a partir de la jornada 3, debemos sumar los datos de la fila 23 de la jornada anterior, con los valores de las filas $28, 29$ y 30 (también de la jornada anterior) correspondientes a cada estado, con tal de tener siempre el número total de P, E o G hasta la jornada $n-1$.

- Para cada jornada n y para cada fila comprendida entre la 35 y la 54, si un equipo ha perdido (P) en la jornada $n-1$, escribimos un 1 en la columna correspondiente al estado P, E o G obtenido en la jornada n, y un 0 en las demás columnas.

  Para cada jornada n y para cada fila comprendida entre la 55 y la 74, si un equipo ha empatado (E) en la jornada $n-1$, escribimos un 1 en la columna correspondiente al estado P, E o G obtenido en la jornada n, y un 0 en las demás columnas.

  Y, de las filas 75 a la 94, hacemos lo mismo pero para los que han ganado (G) en la jornada $n-1$.

- En la fila 96, sumamos las casillas con un 1 de las filas 35 a la 54, junto con el valor correspondiente en la jornada $n-1$ de la misma fila 96. De esta forma, obtenemos el acumulado de equipos que han perdido (P) en la jornada $n-1$ y han perdido (P), empatado (E) o ganado (G) en la jornada n, hasta la jornada n.

  En la fila 97, sumamos las casillas con un 1 de las filas 55 a la 74, junto con el valor correspondiente en la jornada $n-1$ de la misma fila 97. De esta forma, obtenemos

[2] Los resultados han sido extraídos de la Referencia [5].

el acumulado de equipos que han empatado (E) en la jornada $n-1$ y han perdido (P), empatado (E) o ganado (G) en la jornada n, hasta la jornada n.

Y, en la fila 98, sumamos las casillas con un 1 de las filas 75 a la 94, junto con el valor correspondiente en la jornada $n-1$ de la misma fila 98. De esta forma, obtenemos el acumulado de equipos que han ganado (G) en la jornada $n-1$ y han perdido (P), empatado (E) o ganado (G) en la jornada n, hasta la jornada n.

- Por último, denotamos por P_n al acumulado de equipos que han perdido, hasta la jornada n incluida. De forma semejante, denotamos E_n y G_n. Entonces, para cada jornada n, en la fila 32 calculamos:

$$\frac{P_n \cap P_{n-1}}{P_{n-1}}, \quad \frac{E_n \cap P_{n-1}}{P_{n-1}}, \quad \frac{G_n \cap P_{n-1}}{P_{n-1}},$$

que se correspondería con:

$$\mathbb{P}(P_n / P_{n-1}), \quad \mathbb{P}(E_n / P_{n-1}), \quad \mathbb{P}(G_n / P_{n-1}).$$

De esta forma, obtendríamos los valores correspondientes a la primera fila de nuestra matriz de transición hasta la jornada n. De manera análoga, en la fila 33 obtenemos los valores de la segunda fila de la matriz de transición hasta la jornada n, y en la 34, los de la tercera fila de la matriz de transición.

- Las demás filas que no hemos explicado, las hemos usado simplemente para automatizar los cálculos y no tener que repetirlos en cada jornada. Así ha bastado con hacerlos en la jornada 2 y 3 únicamente.

- Finalmente, en las columnas BD, BE y BF y filas $32, 33$ y 34, se encuentran los valores de la matriz de transición que estábamos buscando.

4.1.3 Cálculos

Caso general

Sean: Ω el conjunto de los 20 equipos de la *LFP* $2016/2017$ y $S = \{1, 2, 3\}$ el espacio de estados. Definimos $\{\mathscr{X}_n\}$ como la sucesión de vv.aa. que nos dicen el estado de cada equipo en la jornada $n \in \{1, \dots, 38\}$.

Como ya hemos explicado en la Subsección 4.1.2 anterior, la matriz de transición de $\{\mathscr{X}_n\}$ es:

$$\mathbf{P} = \begin{pmatrix} 0{,}3 & 0{,}3 & 0{,}4 \\ 0{,}4 & 0{,}3 & 0{,}3 \\ 0{,}4 & 0{,}3 & 0{,}3 \end{pmatrix}.$$

Queremos averiguar, si es posible, las probabilidades que tendría cualquier equipo de ganar, empatar o perder, en el tramo final de la liga.

Supondremos que $\{\mathscr{X}_n\}$ es una C.M.H.. Al ser $\mathbf{P}$ obviamente regular, por el Teorema 3.2.1, podemos encontrar un estado de equilibrio.

Para encontrarlo, vamos a ver si podemos diagonalizar la matriz $\mathbf{P}$:

- Primero vamos a calcular el polinomio característico [3]:

$$det(\mathbf{P}-\mathbf{xId}) = \begin{vmatrix} 0{,}3-x & 0{,}3 & 0{,}4 \\ 0{,}4 & 0{,}3-x & 0{,}3 \\ 0{,}4 & 0{,}3 & 0{,}3-x \end{vmatrix} = 0{,}1\cdot x + 0{,}9\cdot x^2 - x^3.$$

  Ahora, por la Proposición 3.3.1 sabemos que $\lambda_1 = 1$ es un valor propio de $\mathbf{P}$, así que para encontrar los otros dos valores propios, dividiremos este polinomio por $(x-1)$:

$$\frac{0{,}1\cdot x + 0{,}9\cdot x^2 - x^3}{(x-1)} = -x^2 - 0{,}1\cdot x.$$

  Los otros dos valores propios son las raíces del polinomio $-x^2 - 0{,}1\cdot x$, que son: 0 y $-0{,}1$.

  Esto quiere decir que el polinomio característico de $\mathbf{P}$, que es de grado 3, tiene 3 valores propios diferentes entre sí, por lo tanto $\mathbf{P}$ es diagonalizable.

- Vamos a encontrar los vectores propios asociados a cada uno de los 3 valores propios usando la ecuación que sabemos que cumplen:[4] $\mathbf{P}\mathbf{v}^t = \lambda\mathbf{v}^t$.

  - Sea $\lambda_1 = 1$ el valor propio, por la misma Proposición 3.3.1 ya sabemos que el vector propio asociado es $\mathbf{v}_1 = (1,1,1)$.
  - Sea $\lambda_2 = 0$ el valor propio,

$$(\mathbf{P}-\lambda_2 Id)\mathbf{x}^t = \mathbf{0}^t \Rightarrow \begin{pmatrix} 0{,}3 & 0{,}3 & 0{,}4 \\ 0{,}4 & 0{,}3 & 0{,}3 \\ 0{,}4 & 0{,}3 & 0{,}3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow$$

$$\begin{cases} 0{,}3x_1 + 0{,}3x_2 + 0{,}4x_3 = 0 \\ 0{,}4x_1 + 0{,}3x_2 + 0{,}3x_3 = 0 \\ 0{,}4x_1 + 0{,}3x_2 + 0{,}3x_3 = 0 \end{cases} \Rightarrow \begin{cases} 0{,}3x_1 + 0{,}3x_2 + 0{,}4x_3 = 0 \\ 0{,}4x_1 + 0{,}3x_2 + 0{,}3x_3 = 0 \end{cases} \Rightarrow$$

    Claramente es un sistema compatible indeterminado. Sea $x_3 = \alpha$, entonces:

$$\begin{cases} 0{,}3x_1 + 0{,}3x_2 = -0{,}4\alpha \\ 0{,}4x_1 + 0{,}3x_2 = -0{,}3\alpha \end{cases} \Rightarrow$$

    Le restamos la primera fila a la segunda fila y nos queda:

$$0{,}1x_1 = 0{,}1\alpha \Rightarrow x_1 = \alpha.$$

    Solo nos falta por saber qué vale x_2. Por sustitución en la primera ecuación:

$$0{,}3\alpha + 0{,}3x_2 = -0{,}4\alpha \Rightarrow 0{,}3x_2 = -0{,}7\alpha \Rightarrow x_2 = -\frac{0{,}7}{0{,}3}\alpha = -\frac{7}{3}\alpha.$$

    En definitiva, el vector propio asociado al valor propio λ_2 es $v_2 = (1, -\frac{7}{3}, 1)$.

[3] Véase la Definición A.1.10 del Apéndice A.
[4] Véase la Definición A.1.8 del Apéndice A.

- Por último, sea $\lambda_3 = -0{,}1$ el valor propio:

$$(\mathbf{P} - \lambda_2 Id)\mathbf{x}^t = \mathbf{0}^t \Rightarrow \begin{pmatrix} 0{,}4 & 0{,}3 & 0{,}4 \\ 0{,}4 & 0{,}4 & 0{,}3 \\ 0{,}4 & 0{,}3 & 0{,}4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow$$

$$\begin{cases} 0{,}4x_1 + 0{,}3x_2 + 0{,}4x_3 = 0 \\ 0{,}4x_1 + 0{,}4x_2 + 0{,}3x_3 = 0 \\ 0{,}4x_1 + 0{,}3x_2 + 0{,}4x_3 = 0 \end{cases} \Rightarrow \begin{cases} 0{,}4x_1 + 0{,}3x_2 + 0{,}4x_3 = 0 \\ 0{,}4x_1 + 0{,}4x_2 + 0{,}3x_3 = 0 \end{cases} \Rightarrow$$

Este es otro sistema compatible indeterminado. Sea $x_3 = \beta$, entonces:

$$\begin{cases} 0{,}4x_1 + 0{,}3x_2 = -0{,}4\beta \\ 0{,}4x_1 + 0{,}4x_2 = -0{,}3\beta \end{cases} \Rightarrow$$

Le restamos la primera fila a la segunda fila y nos queda:

$$0{,}1x_2 = 0{,}1\beta \Rightarrow x_2 = \beta.$$

Solo nos falta por saber qué vale x_1. Por sustitución en la primera ecuación:

$$0{,}4x_1 + 0{,}3\beta = -0{,}4\beta \Rightarrow 0{,}4x_1 = -0{,}7\beta \Rightarrow x_1 = -\frac{0{,}7}{0{,}4}\beta = -\frac{7}{4}\beta.$$

En definitiva, el vector propio asociado al valor propio λ_3 es $v_3 = (-\frac{7}{4}, 1, 1)$.

- Una vez conocemos los vectores propios asociados a cada valor propio, ya podemos escribir nuestra matriz **P** como sigue:

$$\mathbf{P} = \mathbf{U}\mathbf{D}\mathbf{U}^{-1} \Rightarrow$$

$$\begin{pmatrix} 0{,}3 & 0{,}3 & 0{,}4 \\ 0{,}4 & 0{,}3 & 0{,}3 \\ 0{,}4 & 0{,}3 & 0{,}3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -7/4 \\ 1 & -7/3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -0{,}1 \end{pmatrix} \begin{pmatrix} 0{,}36364 & 0{,}3 & 0{,}33636 \\ 0 & -0{,}3 & 0{,}3 \\ -0{,}36364 & 0 & 0{,}36364 \end{pmatrix}$$

Finalmente, por la técnica expuesta en la Subsección 3.3.1, tenemos que el estado de equilibrio **w** que buscábamos es el vector estocástico:

$$(u_{11}^{-1}, u_{12}^{-1}, u_{13}^{-1}) = (0{,}36364, 0{,}3, 0{,}33636).$$

Esto quiere decir que, suponiendo que $\{\mathscr{X}_n\}$ es una C.M.H., en el tramo final de la temporada, cualquier equipo tiene un 36,36 % de probabilidad de perder, un 30 % de empatar y un 33,63 % de ganar.

Como comprobación de que la teoría que hemos mostrado es verídica, vamos a calcular el estado de equilibrio mediante el resto de técnicas que hemos expuesto en la Sección 3.3.

A continuación, vamos a calcular el punto fijo de **P**:

$$(w_1, w_2, w_3)\begin{pmatrix} 0{,}3 & 0{,}3 & 0{,}4 \\ 0{,}4 & 0{,}3 & 0{,}3 \\ 0{,}4 & 0{,}3 & 0{,}3 \end{pmatrix} = (w_1, w_2, w_3) \Rightarrow \begin{cases} 0{,}3w_1 + 0{,}4w_2 + 0{,}4w_3 = w_1 \\ 0{,}3w_1 + 0{,}3w_2 + 0{,}3w_3 = w_2 \\ 0{,}4w_1 + 0{,}3w_2 + 0{,}3w_3 = w_3 \end{cases} \Rightarrow$$

$$\begin{cases} -0{,}7w_1 + 0{,}4w_2 + 0{,}4w_3 = 0 \\ 0{,}3w_1 - 0{,}7w_2 + 0{,}3w_3 = 0 \\ 0{,}4w_1 + 0{,}3w_2 - 0{,}7w_3 = 0 \end{cases}$$

El determinante de la matriz de este sistema es 0. Como existen determinantes 2×2 de la matriz que son diferentes a 0, sabemos con certeza que este sistema tiene una solución compatible indeterminada. Sea $w_3 = \gamma$, por el *método de Gauss*[5] nos resulta el siguiente sistema[6] a resolver:

$$\begin{cases} -0{,}7w_1 + 0{,}4w_2 + 0{,}4\gamma = 0 \\ -0{,}52857w_2 + 0{,}47143\gamma = 0 \end{cases}$$

De la segunda ecuación, nos queda que $w_2 = \frac{0{,}47143}{0{,}52857}\gamma = 0{,}8919\gamma$. Y, sustituyendo en la primera ecuación, nos resulta:

$$-0{,}7w_1 + 0{,}4 \cdot 0{,}8919\gamma + 0{,}4\gamma = 0 \Rightarrow -0{,}7w_1 = -0{,}75676\gamma \Rightarrow w_1 = 1{,}0811\gamma.$$

Ahora bien, sabemos que el estado de equilibrio es un vector estocástico, así que $w_1 + w_2 + w_3 = 1$. Añadiendo esta condición a lo que hemos obtenido:

$$w_1 + w_2 + w_3 = 1 \Rightarrow 1{,}0811\gamma + 0{,}8919\gamma + \gamma = 1 \Rightarrow \gamma = 0{,}33636.$$

Por tanto, el punto fijo (habiendo redondeado algunos valores) es:

$$\mathbf{w} = (w_1, w_2, w_3) = (1{,}0811\gamma, 0{,}8919\gamma, \gamma) = (0{,}36364, 0{,}3, 0{,}33636),$$

que es tal cual lo que nos debía salir.

Por último, como $\mathbf{P}$ es de pequeña dimensión, vamos a computar las potencias de $\mathbf{P}$, usando *Octave*, para ver cómo las filas convergen al vector estocástico $\mathbf{w}$ encontrado:

$$\mathbf{P} = \begin{pmatrix} 0{,}3 & 0{,}3 & 0{,}4 \\ 0{,}4 & 0{,}3 & 0{,}3 \\ 0{,}4 & 0{,}3 & 0{,}3 \end{pmatrix} \quad (\mathbf{P}^2) = \begin{pmatrix} 0{,}37 & 0{,}3 & 0{,}33 \\ 0{,}36 & 0{,}3 & 0{,}34 \\ 0{,}36 & 0{,}3 & 0{,}34 \end{pmatrix} \quad (\mathbf{P}^3) = \begin{pmatrix} 0{,}363 & 0{,}3 & 0{,}337 \\ 0{,}364 & 0{,}3 & 0{,}336 \\ 0{,}364 & 0{,}3 & 0{,}336 \end{pmatrix}$$

$$(\mathbf{P}^4) = \begin{pmatrix} 0{,}36370 & 0{,}3 & 0{,}3363 \\ 0{,}36360 & 0{,}3 & 0{,}3364 \\ 0{,}36360 & 0{,}3 & 0{,}3364 \end{pmatrix} \quad (\mathbf{P}^5) = \begin{pmatrix} 0{,}36363 & 0{,}3 & 0{,}33637 \\ 0{,}36364 & 0{,}3 & 0{,}33636 \\ 0{,}36364 & 0{,}3 & 0{,}33636 \end{pmatrix} \approx \mathbf{W}$$

Notemos que, en $(\mathbf{P}^5)$, cada fila se aproxima al valor de $\mathbf{w}$ hasta casi el quinto decimal. Es increíble ver lo rápido que converge $(\mathbf{P}^n)$ a $\mathbf{W}$. Aunque, por otro lado, ya habíamos adelantado en el Capítulo 3 que $(\mathbf{P}^n)$ converge a $\mathbf{W}$ con velocidad exponencial.

[5] Véase el enunciado del método en la Definición A.1.7 del Apéndice A.

[6] Para averiguar la matriz de este sistema, hemos utilizado una función programada con *Octave*, que el lector podrá consultar en el Apéndice C.

Equipos ganadores

En este caso vamos a realizar un experimento similar al anterior pero con un subconjunto de Ω al que llamaremos Ω_1.

Este nuevo subconjunto estará formado únicamente por equipos que nosotros hemos considerado, según estadísticas pasadas, ganadores. A saber: *Real Madrid*, *Barcelona* y *Sevilla*.

Para esta aplicación utilizaremos: el mismo espacio de probabilidad que en el caso general anterior (lo único que ahora tendremos Ω_1 en lugar de Ω), el mismo espacio de estados y la misma sucesión de vv.aa..

En este caso, el método de calculo de la matriz de transición ha sido similar al del caso general anterior, pero considerando los equipos de Ω_1 en lugar de los de Ω. Así pues, la matriz de transición[7] es:

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & 1 \\ 0{,}1 & 0{,}3 & 0{,}6 \\ 0{,}2 & 0{,}2 & 0{,}6 \end{pmatrix}.$$

De nuevo, supondremos que $\{\mathscr{X}_n\}$ es una C.M.H..

El procedimiento para encontrar el estado de equilibrio será semejante al del apartado anterior. Por ello, no explicaremos porqué hacemos lo que hacemos en cada paso como antes.

Primero de todo debemos ver si $\mathbf{P}$ es regular:

$$(\mathbf{P}^2) = \begin{pmatrix} 0{,}2 & 0{,}2 & 0{,}6 \\ 0{,}15 & 0{,}21 & 0{,}64 \\ 0{,}14 & 0{,}18 & 0{,}68 \end{pmatrix}.$$

Como $(\mathbf{P}^2)$ tiene todos sus valores mayores que 0, tenemos que $\mathbf{P}$ es regular.

Ahora obtendremos el estado de equilibrio calculando el punto fijo para $\mathbf{P}$:

$$(w_1, w_2, w_3) \begin{pmatrix} 0 & 0 & 1 \\ 0{,}1 & 0{,}3 & 0{,}6 \\ 0{,}2 & 0{,}2 & 0{,}6 \end{pmatrix} = (w_1, w_2, w_3) \Rightarrow \begin{cases} 0{,}1 w_2 + 0{,}2 w_3 = w_1 \\ 0{,}3 w_2 + 0{,}2 w_3 = w_2 \\ w_1 + 0{,}6 w_2 + 0{,}6 w_3 = w_3 \end{cases} \Rightarrow$$

$$\begin{cases} -w_1 + 0{,}1 w_2 + 0{,}2 w_3 = 0 \\ -0{,}7 w_2 + 0{,}2 w_3 = 0 \\ w_1 + 0{,}6 w_2 - 0{,}4 w_3 = 0 \end{cases}$$

El determinante de la matriz de este sistema es 0. Como en el caso anterior, consideramos $w_3 = \gamma$. Por el *método de Gauss*, obtenemos el siguiente sistema a resolver:

$$\begin{cases} -w_1 + 0{,}1 w_2 + 0{,}2\gamma = 0 \\ 0{,}7 w_2 - 0{,}2\gamma = 0 \end{cases}$$

De la segunda ecuación nos resulta que: $w_2 = \frac{0{,}2}{0{,}7}\gamma = \frac{2}{7}\gamma$.

Sustituimos este valor en la primera ecuación y el resultado es:

[7] Esta matriz es el resultado de redondear al primer decimal los valores que se encuentran en las columnas BD, BE y BF y filas $15, 16$ y 17 de la hoja *Equipos ganadores* del archivo *Excel* adjunto: *Estadisticas futbol*.

$$-w_1 + 0{,}1\frac{2}{7}\gamma + 0{,}2\gamma = 0 \Rightarrow w_1 = \frac{16}{70}\gamma.$$

Finalmente, como sabemos que **w** es un vector estocástico, tenemos que $w_1 + w_2 + w_3 = 1$. Así pues:

$$w_1 + w_2 + w_3 = 1 \Rightarrow \frac{16}{70}\gamma + \frac{2}{7}\gamma + \gamma = 1 \Rightarrow \frac{53}{35}\gamma = 1 \Rightarrow \gamma = \frac{35}{53} = 0{,}6604.$$

En definitiva, el estado de equilibrio **w** es:

$$\mathbf{w} = (w_1, w_2, w_3) = (\frac{16}{70}\cdot 0{,}6604, \frac{2}{7}\cdot 0{,}6604, 0{,}6604) = (0{,}1510, 0{,}1886, 0{,}6604),$$

donde hemos tenido que redondear al cuarto decimal (teniendo en cuenta que el vector es estocástico).

Este resultado implica que, en el tramo final de la liga, es muy probable que los equipos ganadores ganen, como era de esperar.

Equipos perdedores

En este otro caso vamos a definir el experimento $\Omega_2 \subset \Omega$ como los equipos que nosotros hemos considerado, según estadísticas pasadas, perdedores. A saber: *Sporting*, *Granada* y *Osasuna*.

Como en la Subsubsección 4.1.3 anterior, utilizaremos: el mismo espacio de probabilidad que en el caso general (lo único que ahora tendremos Ω_2 en lugar de Ω), el mismo espacio de estados y la misma sucesión de vv.aa..

En este caso, el método de calculo de la matriz de transición ha sido similar al del caso general anterior (al igual que en el caso de los *Equipos ganadores*), pero considerando los equipos de Ω_2 en lugar de los de Ω. Así pues, la matriz de transición[8] es:

$$\mathbf{P} = \begin{pmatrix} 0{,}6 & 0{,}3 & 0{,}1 \\ 0{,}7 & 0{,}2 & 0{,}1 \\ 0{,}6 & 0{,}4 & 0 \end{pmatrix}.$$

Supondremos, de nuevo, que $\{\mathscr{X}_n\}$ es una C.M.H..

Primero de todo, debemos ver si **P** es regular:

$$(\mathbf{P}^2) = \begin{pmatrix} 0{,}63 & 0{,}28 & 0{,}09 \\ 0{,}62 & 0{,}29 & 0{,}09 \\ 0{,}64 & 0{,}26 & 0{,}1 \end{pmatrix}.$$

Como $(\mathbf{P}^2)$ tiene todos sus valores mayores que 0, tenemos que **P** es regular.

Esta vez, como la dimensión de **P** es pequeña, vamos a encontrar una aproximación del vector estocástico **w** mediante el cálculo de las potencias de **P**. Ya hemos visto, en el caso general, que para $n = 5$ la aproximación ya era muy buena. Como ya sabemos el valor de $(\mathbf{P}^2)$, vamos a calcular, únicamente, $(\mathbf{P}^5)$ y $(\mathbf{P}^{10})$:

[8] Esta matriz es el resultado de redondear al primer decimal los valores que se encuentran en las columnas BD, BE y BF y filas $15, 16$ y 17 de la hoja *Equipos perdedores* del archivo *Excel* adjunto: *Estadisticas futbol*.

$$(\mathbf{P}^5) = \begin{pmatrix} 0{,}6281 & 0{,}28099 & 0{,}09091 \\ 0{,}62811 & 0{,}28098 & 0{,}09091 \\ 0{,}62806 & 0{,}28104 & 0{,}0909 \end{pmatrix} \quad (\mathbf{P}^{10}) = \begin{pmatrix} 0{,}628099173 & 0{,}280991735 & 0{,}0909090908 \\ 0{,}628099173 & 0{,}280991735 & 0{,}0909090908 \\ 0{,}628099174 & 0{,}280991734 & 0{,}0909090909 \end{pmatrix} \approx \mathbf{W}.$$

Notemos que, hasta el octavo decimal, los valores de cada fila son iguales. De nuevo, hemos comprobado lo rápido que converge la matriz de transición al estado de equilibrio.

En definitiva, podemos afirmar, redondeando hasta el cuarto decimal y teniendo en cuenta que **w** es estocástico, que:

$$\mathbf{w} = (w_1, w_2, w_3) = (0{,}6281, 0{,}2810, 0{,}0909).$$

Este resultado nos asegura que, en el tramo final de la liga, es muy probable que los equipos perdedores pierdan, como era también de esperar.

Concluimos que, si nos interesan las apuestas deportivas (que ahora están muy de moda), en el último tramo de liga deberíamos apostar fuerte porque los equipos ganadores ganen y porque los equipos perdedores pierdan. Debido a que hemos comprobado matemáticamente (aunque hemos debido hacer algunas suposiciones y aproximaciones), que las probabilidades de que ocurran estos sucesos son altas.

En este ejercicio, además, podemos comprobar cuán cerca estaban nuestras expectativas con respecto a la realidad. Debido a que, a día de hoy (11/05/2017), se han jugado 36 jornadas ya. Esto implica que, el conjunto de los equipos ganadores y el de perdedores, han obtenido 51 resultados nuevos cada uno, con respecto a la primera mitad de la temporada. Veamos, por separado, los resultados para los equipos ganadores y para los equipos perdedores:

- Del conjunto de equipos ganadores, se han obtenido 34 nuevas victorias, 9 empates y 8 derrotas.[9] Esto da como resultado un vector de proporciones, que equivaldría a las probabilidades que han tenido estos equipos de perder, empatar o ganar en estas 36 jornadas de liga:

  $$\mathbf{z} = (\frac{8}{51}, \frac{9}{51}, \frac{34}{51}) = (0{,}155, 0{,}175, 0{,}670).$$

  Recordemos que en el caso de los equipos ganadores, el estado de equilibrio era

  $$\mathbf{w} = (0{,}1510, 0{,}1886, 0{,}6604).$$

  Notemos cómo ambos vectores son prácticamente iguales. Difieren, únicamente, en un par de centésimas unas componentes de otras.

- Del conjunto de equipos perdedores, se han obtenido 8 nuevas victorias, 10 empates y 33 derrotas.[10] Esto da como resultado un vector de proporciones:

  $$\mathbf{t} = (\frac{33}{51}, \frac{10}{51}, \frac{8}{51}) = (0{,}647, 0{,}196, 0{,}157).$$

[9] Estos datos pueden ser consultados en la Referencia [5].

[10] Estos datos también pueden ser consultados en la Referencia [5].

Recordemos que en el caso de los equipos perdedores, el estado de equilibrio era

$$\mathbf{w} = (0{,}6281, 0{,}2810, 0{,}0909).$$

En este caso, los vectores no son tan semejantes. Aún así, son muy próximos los valores de un vector respecto a los del otro.

Suponemos que no son tan próximos como en el otro caso, debido a que en la segunda mitad de la temporada, por miedo a no descender de categoría, los equipos perdedores tienden a mejorar su % de partidos ganados, y eso provoca que las demás probabilidades se reajusten en consecuencia.

4.2 Caso real: Póquer

4.2.1 Introducción

El *Texas hold'em* es la modalidad de póquer más conocida y practicada del mundo. Aunque muchas personas lo traten de juego de azar, el *Texas hold'em* es un juego en el que influyen las probabilidades y la estadística más de lo que muchos suponen.

Normalmente, juegan entre 9 y 10 personas en una mesa. Se reparten 2 cartas por jugador. Cada jugador mira sus cartas sin dejar que los compañeros miren sus cartas.

Una vez se han repartido las cartas a cada jugador, nos encontramos en el *pre-flop*. Se trata de la ronda anterior al *flop* (de ahí el nombre). Lo que hacen ahora los jugadores es, por orden, apostar fichas o no, dependiendo de las cartas que les hayan tocado. Cada jugador tiene tres posibles acciones: apostar (*raise* en inglés: R), pasar (*check* en inglés: C) o abandonar (*fold* en inglés: F).

Cuando se acaban las rondas de apuestas en el *pre-flop*, si 2 o más jugadores no han abandonado, se ponen 3 cartas boca arriba en medio de la mesa donde todos puedan verlas. En este momento nos encontramos en el *flop*. Cada jugador, de nuevo por orden, habla diciendo si R, C o F. Cuando se acaban las rondas de apuestas y siguen jugando un mínimo de 2 jugadores, se muestra otra carta boca arriba y se coloca al lado de las 3 anteriores. Ahora nos encontramos en el *turn*.

Vuelve a hablar cada jugador, elige la acción que prefiera y al acabar las rondas de apuestas, si siguen 2 jugadores como mínimo jugando, se muestra la última carta boca arriba. Esta fase se conoce como *river*.

Una vez que están las 5 cartas boca arriba sobre la mesa, cada jugador debe realizar, mentalmente, la combinación de cartas que más le convenga, entre las dos que posee y las que se encuentran encima de la mesa. De tal forma que, de entre las 7 cartas en total que puede combinar, debe elegir las 5 que mejor resultado le de para intentar llevarse el bote contra los oponentes. Se producen de nuevo una serie de apuestas, o no, entre los jugadores que siguen jugando. Cuando se acaban las apuestas, si siguen jugando un mínimo de 2 personas, se muestran las 2 cartas que posee cada uno y se ve quién ha conseguido la mejor combinación de cartas.

El que tiene la mejor combinación, gana el bote total apostado.

Durante una partida de póquer normal, se juegan muchas manos. Lo que hemos descrito ahora es una única mano. Se juegan tantas como fichas les quede por apostar a los jugadores. Y, en cada mano, se producen muchas rondas de apuestas.

Cada una de las dos cartas que se reparten al inicio de la mano, por orden de mejor valorada a peor, puede ser: $A, K, Q, J, 10, 9, 8, 7, 6, 5, 4, 3$ o 2. Además, para cada uno de estos valores posibles, hay 4 palos. A saber: corazones, diamantes, tréboles y picas. Por tanto, en total, la baraja cuenta con 52 cartas distintas.

Gracias a un compañero que es profesional del póquer, hemos podido recopilar muchos datos precisos sobre sus partidas y saber más sobre este mundo. Para ello, las dos herramientas que él más utiliza son: el programa *Hold'em Manager* y la *app Poker Equity Calculator*:

- El *Hold'em Manager* es un programa que utilizan los jugadores profesionales de póquer online. Es una base de datos donde se almacenan todas las manos que has jugado. De cada mano, tiene un registro de cada una de las acciones que has tomado tú y tus adversarios, además de los beneficios obtenidos al final de la mano.

  Cuando se tiene una muestra suficiente de manos, este programa te proporciona una gran cantidad de porcentajes relevantes. Desde qué probabilidad hay de que si tu adversario hace R en la ronda x, haga C en la siguiente ronda, hasta la probabilidad que tiene un adversario de hacer F, si yo he hecho previamente R en la misma ronda de apuestas. Además, estas probabilidades las conoces mientras juegas la partida con ellos, de forma que tienes una cierta ventaja mientras juegas.

  También lleva un control de cuánto dinero has ganado o perdido, para saber si tienes beneficios o pérdidas desde que empezaste a jugar.

  Estas estadísticas también las tiene sobre tu propio juego, así podrás evaluarte y mejorar con el tiempo.

  Este programa puede descargarse libremente sin ningún cargo. Una vez descargado, puedes acceder a los datos de algún compañero que te pase su *Id*, pero no puedes jugar partidas con él tu mismo hasta que pagues una licencia. Por tanto, yo puedo consultar los datos de mi compañero gratis desde mi ordenador, pero si quisiera jugar usando este programa, debería pagar.

- La *app Poker Equity Calculator* nos ayuda a saber las probabilidades que tenemos de ganar o perder en una mano, dependiendo de la posición que ocupemos en la mesa, los jugadores a los que nos enfrentemos en las rondas de apuestas y, por supuesto, de las 2 cartas iniciales que nos hayan tocado.

  La ventaja de esta *app* respecto al programa *Hold'em Manager*, es que podemos consultar las probabilidades en nuestro móvil en cualquier momento, incluso mientras estamos jugando una partida en vivo en un casino. Aunque, lo normal, no es utilizarla mientras juegas, debido a que las manos se juegan muy rápido, y no da tiempo a consultar estas probabilidades en cada mano.

  Los profesionales, como mi amigo, se suelen saber estas probabilidades de memoria. Para lo que les sirve esta aplicación en el día a día, principalmente, es para comprobar si lo que pensaban era cierto o no, y así ir repasando las probabilidades para seguir memorizándolas.

4.2.2 Construcción de la matriz de transición del caso: Rondas de apuestas

En esta sección explicaremos cómo hemos obtenido la matriz de transición **P**:

$$\begin{pmatrix} 0{,}3 & 0{,}6 & 0{,}1 \\ 0 & 0{,}7 & 0{,}3 \\ 0 & 0 & 1 \end{pmatrix},$$

que usamos en la Subsubsección 4.2.4 de este caso real.

Para cada estado, hemos obtenido una muestra de manos gracias al programa *Hold'em Manager*[11] de un amigo. No hemos descargado el programa en nuestro ordenador, nos ha bastado con consultar los datos en su propio ordenador, y apuntar los resultados en una hoja a mano, debido a que queríamos una muestra relevante pero no muy grande. Si hubiéramos querido, podríamos haber afinado las probabilidades todo lo que hubiéramos deseado, ya que mi amigo contaba con una base de datos de más de 3608 manos.

Otra causa por la que lo hemos hecho así, es porque nosotros necesitábamos unos valores específicos para nuestra matriz, y hemos tenido que ir consultando a ojo cada mano para poder contabilizar los datos que realmente nos interesaban.

En definitiva, para obtener la matriz, lo que hemos hecho ha sido:

- En las manos que mi amigo ha llegado al *flop*, he apuntado las veces que mi amigo ha hecho R, C o F en la ronda de apuestas n, si en la ronda $n-1$ ha hecho R, C o F.

- Sea $\{1,2,3\} \longrightarrow \{R,C,F\}$ una biyección entre los tres primeros números naturales y las diferentes acciones que puede tomar mi amigo en cada ronda de apuestas.

- En total, he apuntado 32 manos en las que se ha hecho R en la ronda $n-1$ y R, C o F en la ronda n. Las probabilidades han sido:

$$\frac{9}{32} = 0{,}28, \quad \frac{19}{32} = 0{,}59, \quad \frac{4}{32} = 0{,}13,$$

  para R, C y F respectivamente. Por ello, redondeando, tenemos que la primera fila de la matriz es $(0{,}3, 0{,}6, 0{,}1)$.

- En total, he apuntado 39 manos en las que se ha hecho C en la ronda $n-1$ y R, C o F en la ronda n. Las probabilidades han sido:

$$\frac{1}{39} = 0{,}026, \quad \frac{28}{39} = 0{,}72, \quad \frac{10}{39} = 0{,}26,$$

  para R, C y F respectivamente. Por ello, redondeando, tenemos que la segunda fila de la matriz es $(0, 0{,}7, 0{,}3)$.

- Por último, la tercera fila de la matriz es $(0,0,1)$. Esto es debido a que si en la ronda de apuestas $n-1$ mi amigo ha abandonado (F), entonces en la ronda n y en posteriores también abandonará seguro. Esto se entiende como que si haces F, entonces harás F el resto de rondas seguro (porque ya no las juegas).

[11] El código de su licencia es: HM2-E4J8W6-0V7TQ-4854N-127H3-2GUJH-M5LH3.

4.2.3 Construcción de la matriz de transición del caso: Cartas buenas o malas

En esta sección explicaremos cómo hemos obtenido la matriz de transición **P**

$$\begin{pmatrix} 0{,}36 & 0{,}64 \\ 0{,}24 & 0{,}76 \end{pmatrix},$$

que hemos usado en la Subsubsección 4.2.4 de este caso real.

Sea

$$T = \{A_{\clubsuit}A_{\spadesuit}, A_{\clubsuit}K_{\diamondsuit}, K_{\heartsuit}A_{\clubsuit}, A_{\clubsuit}Q_{\blacklozenge}, Q_{\heartsuit}A_{\clubsuit}, \ldots, 2_{\heartsuit}2_{\clubsuit}\},$$

(donde los subíndices $\clubsuit, \spadesuit, \diamondsuit$ y $\heartsuit$ representan los diferentes palos que puede tener una carta de la baraja) el conjunto formado por las combinaciones posibles de 2 cartas que podemos obtener inicialmente en una mano. Este conjunto tiene un total de $(52 \cdot 51)/2 = 1326$ elementos (ya que queremos saber el numero de subconjuntos de 2 elementos escogidos de un conjunto de 52 elementos), o sea $|T| = 1326$. Dividiremos este conjunto de cartas en dos: buenas (B) y malas (M). Así:

$$T = B \cup M,$$

con:

$$B = \{AA_{\circ}, KK_{\circ}, \ldots, 22_{\circ}, AK_{\circ}, AQ_{\circ}, \ldots, A2_{\circ}, KQ_{\circ}, KJ_{\circ}, \ldots, K5_{\circ}\},$$

donde $XX_{\circ}$ representa las diferentes combinaciones de XX con los 4 palos existentes.

Consultaremos, una por una, las manos de mi amigo en la base de datos del *Hold'em Manager* y las clasificaremos dependiendo de si son buenas o malas.

Sea $\{B, M\} \longrightarrow \{1, 2\}$ una biyección entre los 2 tipos de cartas que hemos considerado y los dos primeros números naturales.

Hemos clasificado si eran buenas o malas en la mano n, sabiendo que son buenas en la mano $n-1$, un total de 67 manos. Las probabilidades resultantes han sido:

$$\frac{24}{67} = 0{,}36, \quad \frac{43}{67} = 0{,}64,$$

para cartas buenas y malas respectivamente. Por ello, la primera fila de la matriz de transición es $(0{,}36, 0{,}64)$.

Y, por otro lado, hemos seleccionado si eran buenas o malas en la mano n, sabiendo que son malas en la mano $n-1$, un total de 34 manos. Las probabilidades resultantes han sido:

$$\frac{8}{34} = 0{,}24, \quad \frac{26}{34} = 0{,}76,$$

para cartas buenas y malas respectivamente. Por ello, la segunda fila de la matriz de transición es $(0{,}24, 0{,}76)$.

4.2.4 Cálculos

Rondas de apuestas

Sea Ω el conjunto formado por cada una de las manos que ha jugado mi compañero durante una partida de póquer. Sea $S = \{1,2,3\}$ el espacio de estados y $\{\mathscr{X}_n\}$ la sucesión de vv.aa. que nos dice, una vez en el *flop*, la acción que ha tomado mi compañero en la ronda de apuestas $n \in \mathbb{N}$.

Como ya hemos visto en la Subsección 4.2.2, la matriz de transición es:

$$\mathbf{P} = \begin{pmatrix} 0{,}3 & 0{,}6 & 0{,}1 \\ 0 & 0{,}7 & 0{,}3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Lo que nos gustaría averiguar ahora, son las probabilidades que tendría mi compañero de apostar, pasar o abandonar, en las últimas rondas de apuesta de una mano. De esta forma, podrá hacerse una idea de cuánto puede llegar a apostar o si va a abandonar, por ejemplo. Supondremos que $\{\mathscr{X}_n\}$ es una C.M.H. con matriz de transición $\mathbf{P}$.

Primero de todo, veremos que esta matriz no es irreducible. Con el diagrama de estados asociado a $\mathbf{P}$ en la Figura 4.1 lo comprobaremos mejor.

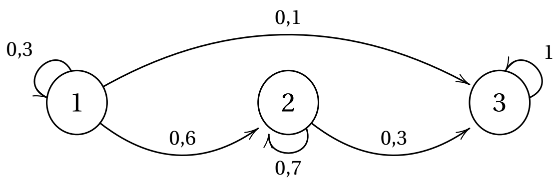

Figura 4.1: Diagrama de estados asociado a $\mathbf{P}$

Como podemos ver, los estados 1 y 3, por ejemplo, no están comunicados (véase la Figura 4.1). El estado 1 conduce a 3, pero 3 no conduce al 1. Notemos, además, que es claramente una C.M.A..

Por tanto, no podemos asegurar que exista un estado de equilibrio. Pero no nos quedaremos de brazos cruzados, vamos a encontrar la representación canónica de $\mathbf{P}$ para intentar extraer algo de información relevante de esta matriz.

Para encontrarla, primero debemos identificar las clases cerradas minimales. Por ello, vamos a calcular la matriz $\mathbf{B}$ con la ayuda de *Octave*:

$$\mathbf{B} = \sum_{n=1}^{3} (\mathbf{P}^n) = \begin{pmatrix} 0{,}42 & 1{,}67 & 0{,}9 \\ 0 & 1{,}53 & 1{,}47 \\ 0 & 0 & 3 \end{pmatrix}$$

Por la Proposición 2.4.3, los estados 1 y 2 no pertenecen a ninguna clase cerrada minimal. Debido a que $\mathbf{B}_2^1 > 0$ pero $\mathbf{B}_2^1 = 0$, y $\mathbf{B}_3^2 > 0$ pero $\mathbf{B}_2^3 = 0$. En cambio, el estado 3 sí pertenece a una clase cerrada minimal que llamaremos $C_1 = \{3\}$. La Proposición 2.4.4 nos asegura que 3 es un estado recurrente y los estados 1 y 2 son transitorios.

Para obtener la representación canónica, debemos reetiquetar los estados. Haremos los siguientes cambios de filas y columnas:

$$\mathbf{P}^1 \leftrightarrow \mathbf{P}^3 \text{ y } \mathbf{P}_1 \leftrightarrow \mathbf{P}_3.$$

Así, la matriz **P** nos quedará de la siguiente forma:

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0{,}3 & 0{,}7 & 0 \\ 0{,}1 & 0{,}6 & 0{,}3 \end{pmatrix}.$$

Y, gracias al Teorema 2.4.1, podemos asegurar que:

$$(\mathbf{B}_1) = 1, \quad \mathbf{R} = \begin{pmatrix} 0{,}3 \\ 0{,}1 \end{pmatrix}, \text{ y } \mathbf{V} = \begin{pmatrix} 0{,}7 & 0 \\ 0{,}6 & 0{,}3 \end{pmatrix}.$$

Ahora, con la ayuda de *Octave*, calcularemos la matriz fundamental $\mathbf{N} = (Id - \mathbf{V})^{-1} V$ para posteriormente poder calcular la matriz $\mathbf{A} = \mathbf{NR}$:

$$\mathbf{N} = \begin{pmatrix} 2{,}33 & 0 \\ 2{,}86 & 0{,}43 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 3{,}33 & 0 \\ 2{,}86 & 1{,}43 \end{pmatrix} \begin{pmatrix} 0{,}3 \\ 0{,}1 \end{pmatrix} = \begin{pmatrix} 0{,}7 \\ 0{,}9 \end{pmatrix}.$$

El Teorema 2.4.3 aplicado a nuestro caso, nos dice que $\mathbf{A}^i_j$ representa la probabilidad de hacer F, si he empezado haciendo C o R. Entonces, mi amigo tiene una probabilidad de 0,9 de hacer F si ha empezado haciendo R, y tiene una probabilidad de 0,7 de hacer F si ha empezado haciendo C.

En otras palabras, si mi amigo empieza haciendo C es menos probable que acabe haciendo F que si empieza haciendo R.

Por otro lado, la matriz **N** también puede darnos información muy valiosa. Recordemos que después de la demostración del Teorema 2.4.2, hablamos sobre esta matriz. La suma de los valores de cada fila de la matriz **N** por separado, será una estimación de las rondas de apuestas que faltarán para hacer F si hemos empezado haciendo C o R.

Hay que tener en cuenta que en la representación canónica hemos intercambiado las etiquetas de los estados R y F. Por ello, la primera fila de **N** representa el estado C (este estado no ha sido reetiquetado, sigue igual), y la segunda fila representa el estado R.

Calculemos las sumas:

$$\sum_{j \in T} \mathbf{N}^1_j = 2{,}33, \quad \sum_{j \in T} \mathbf{N}^2_j = 2{,}86 + 0{,}43 = 3{,}29.$$

Esto quiere decir que la esperanza de rondas que habrá desde el principio hasta llegar a F, es mayor si empezamos haciendo R que si empezamos haciendo C.

Este resultado es interesante y lógico a la vez en este juego. Debido a que, por lo general, si en la primera ronda de apuestas (recordemos que hemos considerado la primera ronda de apuestas empezando desde el *flop*) haces R, suele implicar que tienes una buena combinación de cartas, y, por lo tanto, es más probable que llegues hasta el final de la mano para ver si puedes llevarte el bote. En cambio, si haces C en la primera ronda, entonces tu mano no es tan buena, y estás expectante de ver cómo evolucionan las rondas de apuestas para ver si sigues jugando la mano o abandonas (F).

Cartas buenas o malas

Sea ahora Ω el experimento formado por todas las manos que ha jugado mi amigo en una partida. Sea $\{\mathscr{X}_n\}$ la sucesión de vv.aa. que asigna a cada mano el estado B (cartas buenas), o el estado M (cartas malas). Definimos el espacio de estados S como $S = \{1, 2\}$.

Como ya hemos visto en la Subsección 4.2.3, la matriz de transición para este caso es:

$$\mathbf{P} = \begin{pmatrix} 0{,}36 & 0{,}64 \\ 0{,}24 & 0{,}76 \end{pmatrix}.$$

Supondremos, como en las demás ocasiones, que $\{\mathscr{X}_n\}$ es una C.M.H. con matriz de transición $\mathbf{P}$.

Como la matriz es obviamente regular, vamos a calcular el estado de equilibrio. Para calcularlo, vamos a ver si es diagonalizable:

- Primero calculamos el polinomio característico:

$$det(\mathbf{P} - \mathbf{xId}) = \begin{vmatrix} 0{,}36 - x & 0{,}64 \\ 0{,}24 & 0{,}76 - x \end{vmatrix} = x^2 - 1{,}12x + 0{,}12$$

  Ahora, por la Proposición 3.3.1 de nuevo, sabemos que $\lambda_1 = 1$ es un valor propio de $\mathbf{P}$ (sustituyendo $x = 1$ podemos comprobar fácilmente que el resultado es 0), así que para encontrar el otro valor propio, dividimos este polinomio por $(x - 1)$ y encontramos la raíz del polinomio característico restante:

$$\frac{x^2 - 1{,}12x + 0{,}12}{x - 1} = x - 0{,}12.$$

  Por tanto, el otro valor propio de $\mathbf{P}$ es $0{,}12$.

  Esto quiere decir que el polinomio característico de $\mathbf{P}$, que es de grado 2, tiene 2 valores propios diferentes entre sí, por lo tanto $\mathbf{P}$ es diagonalizable.

- Hallaremos los vectores propios asociados a cada uno de los 2 valores propios:

  - Sea $\lambda_1 = 1$ el valor propio, por la misma Proposición 3.3.1 ya sabemos que el vector propio asociado es $\mathbf{v}_1 = (1, 1)$.
  - Sea $\lambda_2 = 0{,}12$ el valor propio,

$$(\mathbf{P} - \lambda_2 Id)\mathbf{x}^t = \mathbf{0}^t \Rightarrow \begin{pmatrix} 0{,}24 & 0{,}64 \\ 0{,}24 & 0{,}64 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow 0{,}24x_1 + 0{,}64x_2 = 0$$

    Claramente es un sistema compatible indeterminado. Sea $x_2 = \lambda$, entonces:

$$0{,}24x_1 + 0{,}64\lambda = 0 \Rightarrow x_1 = -\frac{0{,}64}{0{,}24}\lambda \Rightarrow x_1 = -\frac{8}{3}\lambda.$$

    Por tanto, el vector propio asociado al valor propio λ_2 es $v_2 = (-\frac{8}{3}, 1)$.

- Una vez conocemos los vectores propios asociados a cada valor propio, ya podemos escribir nuestra matriz $\mathbf{P}$ como sigue:

$$\mathbf{P} = \mathbf{UDU}^{-1} \Rightarrow$$

$$\begin{pmatrix} 0{,}36 & 0{,}64 \\ 0{,}24 & 0{,}76 \end{pmatrix} = \begin{pmatrix} 1 & -8/3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0{,}12 \end{pmatrix} \begin{pmatrix} 0{,}2727272727 & 0{,}7272727272 \\ -0{,}2727272727 & 0{,}2727272727 \end{pmatrix}$$

Finalmente, el estado de equilibrio $\mathbf{w}$ que buscábamos es el vector estocástico:

$$(u_{11}^{-1}, u_{12}^{-1}) = (0{,}272727272727, 0{,}727272727272).$$

Notemos, suponiendo que $\{\mathscr{X}_n\}$ es una C.M.H., que estos valores se acercan a un 30 % de cartas buenas y un 70 % de cartas malas. Gracias a la *app Poker Equity Calculator* sabemos que el subconjunto de cartas buenas (B) elegido al inicio, representa un 30 % del total (T) de combinaciones posibles, y el subconjunto de cartas malas (M) representa un 70 %. Por tanto, el resultado obtenido tiene sentido.

Veamos ahora, al igual que en el caso del fútbol, con qué velocidad converge $(\mathbf{P}^n)$ a $\mathbf{W}$:

$$\mathbf{P} = \begin{pmatrix} 0{,}36 & 0{,}64 \\ 0{,}24 & 0{,}76 \end{pmatrix} \quad (\mathbf{P}^2) = \begin{pmatrix} 0{,}2832 & 0{,}7168 \\ 0{,}2688 & 0{,}7312 \end{pmatrix}$$

$$(\mathbf{P}^5) = \begin{pmatrix} 0{,}27274 & 0{,}72725 \\ 0{,}27272 & 0{,}72727 \end{pmatrix} \quad (\mathbf{P}^{10}) = \begin{pmatrix} 0{,}272727273 & 0{,}727272726 \\ 0{,}272727272 & 0{,}727272727 \end{pmatrix} \approx \mathbf{W}.$$

Podemos comprobar que, al igual que en el caso del fútbol, para $n = 10$ acertamos hasta 8 decimales.

Por último, decir que otra fuente que hemos tomado para la documentación sobre este juego es la Referencia [6].

4.3 Caso real: Google Adwords

4.3.1 Introducción

La herramienta *Google Adwords*, es un servicio de *Google* destinado a aquellos usuarios que quieren realizar posicionamiento web de pago en las páginas de resultados de búsqueda de *Google*, o en la red de *Display*.

Posicionamiento web es, básicamente, conseguir que los anuncios que enlazan a tu página web estén posicionados lo más arriba posible en un motor de búsqueda. En este caso, el motor de búsqueda es *Google* (véase la Figura 4.2).

La red de *Display* son todas las páginas web que cuentan con *slots* (espacios), donde puedas poner tu anuncio.

Para este ejemplo, nos centraremos únicamente en las páginas de resultados de búsqueda de *Google*.

Hay dos tipos de posicionamiento: el posicionamiento de pago, o *SEM* (*Search Engine Marketing*), y el posicionamiento gratuito, o *SEO* (*Search Engine Optimization*).

Para que nos hagamos una idea, aquí tenemos un ejemplo de una búsqueda en el motor de búsqueda *Google*:

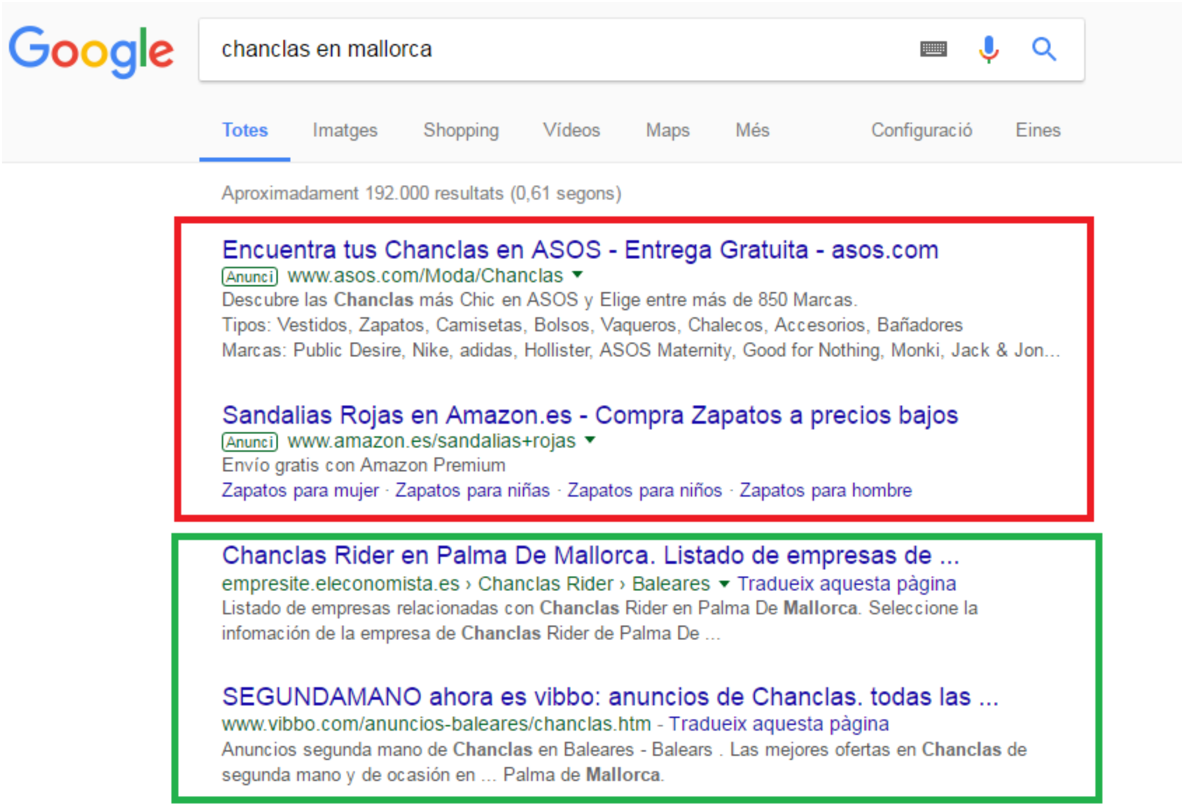

Figura 4.2: Resultado de la búsqueda *chanclas en mallorca* en *Google*.

Los anuncios del primer recuadro que hemos marcado en la Figura 4.2, pertenecen al posicionamiento de pago, es decir, un cliente que utiliza *Google Adwords* ha pagado una cantidad específica para que su anuncio aparezca entre los primeros para esta búsqueda. Los anuncios del segundo recuadro de la Figura 4.2, son de *SEO*, o sea, gracias a la relevancia que tienen estas páginas con la búsqueda realizada, *Google* ha premiado a estos anunciantes colocando sus anuncios en las partes altas del posicionamiento gratuito, sin coste alguno.

Cada tipo de posicionamiento tiene sus ventajas y desventajas, pero ahora no nos adentraremos en estos detalles.

La cuestión es, que si te dedicas al *Marketing Digital* hoy en día, debes saber usar este servicio. En mi caso, he tenido la oportunidad de trabajar optimizando campañas de Marketing en *Google Adwords* para diferentes páginas web.

Actualmente, llevo las campañas de una única página web, que es de donde sacaré los datos que analizaremos a continuación.

Mediante *Google Adwords*, puedes controlar qué anuncios muestras a las personas que te interesa que los vean, y conducirlos a la sección de tu página web que más les pueda interesar, con el objetivo de que realicen una conversión.

Una conversión, es aquello que te proporciona un beneficio. Para cada página web puede ser diferente. Una web sobre venta de zapatos puede tener como conversión, que un usuario compre con su tarjeta unos zapatos. Otra web, puede definir su conversión como una suscripción a un boletín informativo. Y como estos, muchos tipos de conversiones diferentes más.

Por lo tanto, es vital que la relación entre lo que busca la persona en *Google*, y el anuncio y la web a la que lo envías, sea lo más estrecha posible, para conseguir el máximo número de conversiones.

Pero hay que ir con cuidado, porque cada vez que un usuario hace *click* en tu anuncio, pagas una determinada puja a *Google*. Por lo que debes optimizar al máximo qué anuncios muestras, y a quién se los muestras. Eh aquí la optimización necesaria de la que hablábamos antes.

Una de las principales métricas que se deben tener en cuenta cuando manejas esta herramienta, es el *CTR* (*click through rate*). Este valor se calcula como sigue:

$$CTR = \frac{clicks}{impresiones}.$$

Cuando el número de impresiones (veces que tu anuncio ha sido mostrado a un usuario) es suficientemente grande, esta métrica es relevante. Debido a que nos dice cuán relevantes son nuestros anuncios para los usuarios a los que se los mostramos.

4.3.2 Construcción de la matriz de transición

En esta sección explicaremos cómo hemos obtenido la matriz de transición **P**

$$\begin{pmatrix} 0{,}65 & 0{,}35 \\ 0{,}42 & 0{,}58 \end{pmatrix},$$

que hemos usado en la Subsección 4.3 de este caso real, paso por paso:

- De la cuenta de *Google Adwords*[12] descargamos los datos entre las fechas 25/11/2016 y 23/03/2017 en el *Excel* adjunto que hemos llamado *Estadisticas Adwords*.

- En las 4 primeras columnas de este *Excel* adjunto, de la fila 2 a la 120 tenemos: la fecha, clics/día, impresiones/día y *CTR*/día respectivamente.

- La técnica utilizada para obtener la matriz de transición es semejante a la que hemos usado para la del caso real de fútbol:

  La columna *F* de *Estadisticas Adwords*, vale 1 si el valor de la columna *CTR* de la misma fila es mayor que 0,07. Vale 0 en caso contrario. Gracias a esta columna sabremos qué días ha habido un *CTR* alto.

  Por otro lado, la columna *G* vale 1 si el valor de la columna *CTR* de la misma fila es menor o igual que 0,07. Vale 0 en caso contrario. Gracias a esta otra columna sabremos qué días ha habido un *CTR* bajo.

[12]El Id de la cuenta de *Google Adwords* de donde sacaremos los datos es: 621-264-2962. Mi Id de cliente para administrarla es: 349-513-8519.

- Para cada fila n (con $n \in \{3,\ldots,120\}$) de la columna H, escribimos un 1 si en la fila $n-1$ y n de la columna F hay un 1. Esto quiere decir que ha habido un *CTR* alto en el día n y en el $n-1$. En otro caso ponemos un 0.

  Por otro lado, para cada fila n de la columna I, escribimos un 1 si en la fila $n-1$ de la columna F hay un 1, y en la fila n hay un 0. Esto quiere decir que ha habido un *CTR* bajo en el día n y en el $n-1$ ha habido un *CTR* alto. En otro caso escribimos un 0.

- Para cada fila n de la columna J, escribimos un 1 si en la fila $n-1$ de la columna G hay un 1, y en la fila n hay un 0. Esto quiere decir que ha habido un *CTR* alto en el día n y bajo en el $n-1$. En otro caso ponemos un 0.

  Por otro lado, para cada fila n de la columna K, escribimos un 1 si en la fila $n-1$ y n de la columna G hay un 1. Esto quiere decir que ha habido un *CTR* bajo en el día n y en el $n-1$. En otro caso escribimos un 0.

- Sea $\{A,B\} \longrightarrow \{1,2\}$ una biyección entre el conjunto de *CTR*'s altos y bajos (A y B respectivamente) y los dos primeros valores del conjunto de los naturales.

- En la fila 121 sumamos las casillas que valen 1 de las columnas F, G, H, I, J y K, para saber el total en cada caso. En particular, en las casillas de la fila 121 correspondientes a las columnas H e I, tendremos el total de días que ha habido *CTR* alto en $n-1$ y alto o bajo en el día n, respectivamente. Estos valores son 41 y 22 que suman un total de 63 días con *CTR* alto en $n-1$. De esta forma, ya podemos obtener los valores de la primera fila de la matriz de transición:

$$\frac{41}{63} = 0{,}65, \qquad \frac{22}{63} = 0{,}35.$$

- De manera análoga a cómo hemos calculado los valores de las columnas H e I, en las casillas de la fila 121 correspondientes a las columnas J y K, tendremos el total de días que ha habido *CTR* bajo en $n-1$ y alto o bajo en el día n respectivamente.

  Estos valores son 23 y 32 que suman un total de 55 días con *CTR* bajo en $n-1$. Así pues, ya podemos obtener también los valores de la segunda fila de la matriz de transición:

$$\frac{23}{55} = 0{,}42, \qquad \frac{32}{55} = 0{,}58.$$

Por último, añadir que estos *CTR's* que hemos considerado son precisos y fiables porque cada día nuestros anuncios se han mostrado cientos de veces. Esto quiere decir que han obtenido cientos de impresiones, y por tanto el *CTR* es relevante.[13]

[13] Lo podemos comprobar consultando los valores de la columna C del archivo *Excel* adjunto: *Estadisticas Adwords*.

4.3.3 Cálculos

Consideramos Ω, el conjunto formado por cada uno de los días comprendidos entre las fechas 25/11/2016 y 23/03/2017. Sea $\{\mathscr{X}_n\}$ la sucesión de vv.aa., donde cada una de ellas asigna a cada elemento de Ω un estado del espacio de estados S, con $S = \{1,2\}$.

Como ya hemos visto en la Subsección 4.3.2 anterior, la matriz de transición de $\{\mathscr{X}_n\}$ es:

$$\mathbf{P} = \begin{pmatrix} 0{,}65 & 0{,}35 \\ 0{,}42 & 0{,}58 \end{pmatrix}.$$

Supondremos que $\{\mathscr{X}_n\}$ es una C.M.H..

Obviamente, la matriz $\mathbf{P}$ es regular, así que existe un estado de equilibrio. En esta ocasión, vamos a encontrar $\mathbf{w}$ buscando el punto fijo para $\mathbf{P}$:

$$(w_1, w_2)\begin{pmatrix} 0{,}65 & 0{,}35 \\ 0{,}42 & 0{,}58 \end{pmatrix} = (w_1, w_2) \Rightarrow \begin{cases} 0{,}65 w_1 + 0{,}42 w_2 = w_1 \\ 0{,}35 w_1 + 0{,}58 w_2 = w_2 \end{cases} \Rightarrow \begin{cases} -0{,}35 w_1 + 0{,}42 w_2 = 0 \\ \ 0{,}35 w_1 - 0{,}42 w_2 = 0 \end{cases} \Rightarrow$$

$$-0{,}35 w_1 + 0{,}42 w_2 = 0.$$

Ahora, sea $w_2 = \gamma$. Entonces:

$$-0{,}35 w_1 + 0{,}42 w_2 = 0 \Rightarrow w_1 = \frac{0{,}42}{0{,}35}\gamma = \frac{6}{5}\gamma.$$

Y, como $\mathbf{w}$ es un vector estocástico:

$$w_1 + w_2 = 1 \Rightarrow \frac{6}{5}\gamma + \gamma = 1 \Rightarrow \gamma = \frac{5}{11}.$$

Por tanto, el estado de equilibrio viene dado por:

$$\mathbf{w} = (w_1, w_2) = (\frac{6}{11}, \frac{5}{11}) = (0{,}545454, 0, 454545).$$

Esto quiere decir que en el futuro, si todo sigue igual, tendré más días de *CTR* alto que de *CTR* bajo.

CAPÍTULO 5

Conclusión

Hemos tratado de ilustrar con la aplicación de la teoría expuesta a los casos reales (Fútbol, Póquer y *Google Adwords*), la importancia y utilidad de las C.M.H. (absorbentes y regulares). Independientemente de cómo sea la matriz de transición $\mathbf{P}$ (regular o no), si hemos planteado un espacio de probabilidad y una sucesión de vv.aa. adecuadamente, siempre podemos obtener información relevante para sacar nuestras conclusiones.

En el caso de que $\mathbf{P}$ no sea regular, podemos saber qué estados son transitorios y qué estados son recurrentes identificando las clases cerradas minimales. Con ellas, obtenemos la representación canónica de $\mathbf{P}$ y podemos calcular la matriz fundamental $\mathbf{N}$ y la matriz $\mathbf{A} = \mathbf{NR}$, que nos darán información interesante sobre el proceso estocástico definido.

Si $\mathbf{P}$ es regular, entonces podemos encontrar el estado de equilibrio $\mathbf{w}$. Además, tenemos varias formas de obtenerlo: diagonalizando o encontrando la descomposición de Jordan, calculando el punto fijo, o calculando las potencias de la matriz $\mathbf{P}$ para un n suficientemente grande. El estado de equilibrio es vital si queremos saber a qué convergen las probabilidades de que ocurran cada uno de los estados de la C.M. en un futuro.

Por otro lado, notemos que si $\mathbf{P}$ es regular, entonces es irreducible. Y, al estar todos los estados comunicados entre sí, tenemos que la única clase cerrada minimal es el propio espacio de estados S. No se hace una distinción de estados recurrentes y transitorios. Este es el caso en que podemos encontrar un estado de equilibrio. En cambio, si $\mathbf{P}$ no es regular, habrá una o más clases cerradas minimales diferentes de S. En este caso no podemos encontrar un estado de equilibrio, pero sí una representación canónica de la matriz $\mathbf{P}$ y hacer una distinción de qué estados son transitorios y qué estados son recurrentes.

Los diagramas de estados nos han sido muy útiles para hacer una mejor interpretación de los valores de $\mathbf{P}$. Gracias a ellos incluso hemos podido identificar los estados recurrentes y transitorios de forma intuitiva.

Este trabajo también me ha servido para autoevaluar los conocimientos que he aprendido durante los estudios de Grado. Muchos de ellos los tenía claros y los recordaba, pero otros he tenido que refrescarlos consultando apuntes de asignaturas. De echo, como podemos comprobar en la bibliografía, hay varios apuntes de asignaturas que he necesitado repasar para aclarar conceptos que he usado aquí ([7], [8], [9], [10], [11], [12], [13], [14]).

Además, he consultado en repetidas ocasiones la enciclopedia online [15] que, aunque no debe ser la única fuente de información, me ha servido para aclarar algunos detalles sobre definiciones que no me habían quedado del todo claras.

He pecado al haber hecho este trabajo demasiado extenso. Pero ha sido debido a que me gusta aclarar todos los términos que utilizo y creo que no tienen porque ser del conocimiento del lector. Aunque este trabajo esté orientado a cualquier estudiante del Grado de Matemáticas, dependiendo del año en el que se encuentre, no tienen porqué saber según que cosas.

A parte de esto, he creado un extenso apéndice debido a que, personalmente, me gusta dejar claro al lector que he entendido todo lo que he explicado en este trabajo. Por ello, me preocupo por definir conceptos, dar detalles de cada paso que se sigue en las demostraciones de las proposiciones y/o teoremas (aunque algunas ya estén hechas en los libros que he consultado, procuro revisar y entender cómo se han hecho) y mostrar ejemplos claros y concisos.

Estas dos razones tienen como objetivo común la mejor comprensión posible, por parte del lector, de este trabajo.

Finalmente, las principales aportaciones por mi parte a este trabajo se pueden resumir en los siguientes puntos:

- Extracción, organización y estudio de la información proporcionada por mi tutor.
- Realización de los casos reales que no se encontraban en ninguna de las referencias empleadas y extracción de conclusiones basadas en dichos casos reales.
- Mostrar ejemplos cuyos cálculos no se encontraban en las referencias empleadas.
- Demostración de resultados que tampoco se encontraban en las referencias utilizadas.

APÉNDICE A

CONOCIMIENTOS ADICIONALES

Expondremos una serie de conocimientos (con sus respectivas demostraciones si es necesario) de diferentes ramas de las matemáticas que utilizamos en varias secciones de nuestro trabajo, pero que son de menor relevancia (con respecto al tema que tratamos) como para explicarlas en él.

A.1 Conceptos relacionados con Álgebra lineal

Si no decimos lo contrario en las definiciones, consideraremos el $\mathbb{R}$-espacio vectorial $\mathbb{R}^N$ por defecto.

Definición A.1.1. *Sea* $\|\cdot\| : \mathbb{R}^N \longrightarrow \mathbb{R}$ *una aplicación que satisfaga:*

- $\|\mathbf{x}\| \geq 0,\ \forall \mathbf{x} \in \mathbb{R}^N$ *y* $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ *(donde este* $\mathbf{0}$ *es el vector de* $\mathbb{R}^N$ *con todas sus componentes igual a* 0*).*

- $\|\alpha\mathbf{x}\| = \alpha\|\mathbf{x}\|,\ \forall \mathbf{x} \in \mathbb{R}^N$ *y* $\forall \alpha \in \mathbb{R}$.

- $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|,\ \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^N$.

Entonces $\|\cdot\|$ *es una **norma en*** $\mathbb{R}^N$ *y* $(\mathbb{R}^N, \|\cdot\|)$ *es un **espacio normado**.*

Definición A.1.2. *Sea S un conjunto numerable y $1 \leq p < +\infty$ un valor fijado. Denotaremos por* $\ell_{\mathbf{p}}$ *al **espacio*** $\ell_{\mathbf{p}}$*, que es el conjunto de los vectores* $\mathbf{x} = (x_j)$ *indexados por S tt.q.:*

$$\sum_{j\in S} |x_j|^p < +\infty.$$

Además, ℓ_p *es un subespacio vectorial de* $\mathbb{R}^N$ *y*

$$\|\mathbf{x}\|_p = \left(\sum_{j\in S} |x_j|^p\right)^{\frac{1}{p}},$$

con $\mathbf{x} \in \ell_p$, *es una* ***norma*** *en* ℓ_p.

Ejemplo A.1.1. *En particular, para* $p = 1$, *tenemos que si* $\mathbf{x} \in \ell_1$, *entonces:*

$$\|\mathbf{x}\|_1 = \sum_{j\in S} |x_j|.$$

Trivialmente, $\mathcal{T} \subset \ell_1$.

Ejemplo A.1.2. *Para* $p = \infty$, *tenemos que si* $\mathbf{x} \in \ell_\infty$, *entonces:*

$$\|\mathbf{x}\|_\infty = \max_{j\in S}\{|x_j|\}.$$

Definición A.1.3. *Sea* $\|\cdot\| : M_{N\times N}(\mathbb{R}) \longrightarrow \mathbb{R}$, *donde* $M_{N\times N}(\mathbb{R})$ *es el conjunto de las matrices* $N \times N$ *con valores en* $\mathbb{R}$, *una aplicación que satisfaga:*

- $\|\mathbf{A}\| \geq 0$, $\forall \mathbf{A} \in M_{N\times N}(\mathbb{R})$ *y* $\|\mathbf{A}\| = 0 \Leftrightarrow \mathbf{A} = \mathbf{0}$ *(donde este* $\mathbf{0}$ *denota la matriz* $N \times N$ *con todos los valores igual a* 0*).*
- $\|\alpha\mathbf{A}\| = \alpha\|\mathbf{A}\|$, $\forall \mathbf{A} \in M_{N\times N}(\mathbb{R})$ *y* $\forall \alpha \in \mathbb{R}$.
- $\|\mathbf{A}+\mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$, $\forall \mathbf{A}, \mathbf{B} \in M_{N\times N}(\mathbb{R})$.
- $\|\mathbf{A}\cdot\mathbf{B}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$, $\forall \mathbf{A}, \mathbf{B} \in M_{N\times N}(\mathbb{R})$.

Entonces $\|\cdot\|$ *es una* ***norma matricial en*** $M_{N\times N}(\mathbb{R})$.

Ejemplo A.1.3. *Sea* S *un conjunto finito t.q.* $|S| = N$ *y* $\mathbf{A} \in M_{N\times N}(\mathbb{R})$ *una matriz indexada por* S. *Entonces:*

$$\|\mathbf{A}\|_\infty = \sup_{i\in S}\{\sum_{j\in S} |\mathbf{A}^i_j|\},$$

es una norma matricial y se conoce como la ***norma matricial inducida por la norma*** ℓ_∞.

Definición A.1.4. *Sea* $X \neq \emptyset$ *un conjunto y* $d : X \times X \longrightarrow \mathbb{R}$ *una aplicación que satisfaga:*

- $d(x, y) \geq 0$, $\forall x, y \in X$ *y* $d(x, y) = 0 \Leftrightarrow x = y$.
- $d(x, y) = d(y, x)$, $\forall x, y \in X$ *(simetría).*
- $d(x, z) \leq d(x, y) + d(y, z)$, $\forall x, y, z \in X$ *(desigualdad triangular).*

Entonces d *es una* ***distancia en* X** y (X, d) *un* ***espacio métrico***.

Ejemplo A.1.4. *Sea* d_1 *la distancia en* $\mathbb{R}^N$ *definida como sigue:*

$$d_1(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_1 = |x_1 - y_1| + \cdots + |x_N - y_N|, \ \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^N.$$

Sean $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^N$ *arbitrarios pero fijados, veamos que* d_1 *es distancia:*

- *Obviamente* $d_1(\mathbf{x},\mathbf{y}) \geq 0$. *Por otro lado, si* $\mathbf{x} = \mathbf{y}$ *entonces* $x_1 = y_1, \dots, x_N = y_N$, *y esto quiere decir que* $d_1(\mathbf{x},\mathbf{y}) = 0$. *Y, si* $d_1(\mathbf{x},\mathbf{y}) = 0$, *entonces* $|x_i - y_i| = 0$, $\forall i = 1, \dots, N$. *Ya que, como la función valor absoluto es mayor o igual a* 0 *en todo su dominio, para que una suma de valores absolutos sea* 0, *cada uno de los sumandos debe valer* 0.

  Por lo tanto, si $|x_i - y_i| = 0$, $\forall i = 1, \dots, N$, *esto quiere decir que* $x_i = y_i \, \forall i = 1, \dots, N$, *y, en conclusión,* $\mathbf{x} = \mathbf{y}$.

- *Por propiedades de la función valor absoluto sabemos que* $|x_i - y_i| = |y_i - x_i|$, $\forall i = 1, \dots, N$. *Por tanto* $d_1(\mathbf{x},\mathbf{y}) = d_1(\mathbf{y},\mathbf{x})$.

- *Por propiedades de la función valor absoluto de nuevo, tenemos que* $|x_i - z_i| \leq |x_i - y_i| + |y_i - z_i|$, $\forall i = 1, \dots, N$. *Por tanto* $d_1(\mathbf{x},\mathbf{z}) \leq d_1(\mathbf{x},\mathbf{y}) + d_1(\mathbf{y},\mathbf{z})$.

Definición A.1.5. *Un conjunto* X *es* ***convexo*** *si el segmento de extremos* x *e* y *está totalmente contenido dentro de* X, $\forall x, y \in X$.

Definición A.1.6. *La* ***envoltura convexa*** *de un conjunto* X *es el conjunto convexo más pequeño que contiene a* X. *Se puede obtener haciendo la intersección de todos los conjuntos convexos que contienen a* X.

Definición A.1.7. *El* ***método de Gauss*** *consiste en transformar un sistema de ecuaciones en otro equivalente, mediante una serie de operaciones, para conseguir un sistema escalonado y así facilitar su resolución.*

La idea es interpretar los coeficientes de las variables del sistema como una matriz $m \times n$. *Una vez tenemos esto, solo debemos ir multiplicando números por filas (una fila de la matriz equivale a una igualdad entera del sistema) y sumarlas o restarlas a otras filas para procurar obtener una matriz triangular superior, como vemos aquí:*

$$\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1(n-1)}x_n = a_{1n} \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2(n-1)}x_n = a_{2n} \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{m(n-1)}x_n = a_{mn} \end{array}\right. \Rightarrow \left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1(n-1)}x_n = a_{1n} \\ \;0 \quad + b_{22}x_2 + \cdots + b_{2(n-1)}x_n = b_{2n} \\ \qquad\qquad \vdots \\ \;0 \quad + \cdots\cdots + 0 + b_{m(n-1)}x_n = b_{mn} \end{array}\right.$$

Los coeficientes b_{ij} *son los obtenidos después de realizar las operaciones entre filas que hemos explicado antes.*

Dependiendo de cuántas ecuaciones y variables haya, la solución del sistema será incompatible (no hay solución), compatible determinada (soluciones únicas) o compatible indeterminada (soluciones infinitas).

Definición A.1.8. *Sea* $\mathbb{R}^N$ *el espacio vectorial y* $\mathbb{C}$[1] *el cuerpo donde habitan los escalares. Sea* $\mathbf{P}$ *una matriz* $N \times N$.

Entonces decimos que $\mathbf{v} \in \mathbb{R}^N$, *t.q.* $\mathbf{v} \neq \mathbf{0}$, *es un* ***vector propio*** *para* $\mathbf{P}$ *si se cumple lo siguiente:*

$$\mathbf{P}\mathbf{v}^t = \lambda \mathbf{v}^t,$$

[1] Véase la Definición A.6.1 del Apéndice A para más información sobre $\mathbb{C}$.

con $\lambda \in \mathbb{C}$ y $\mathbf{v}^t$ el vector traspuesto de **v** *(vector columna). O sea, la dirección del vector no varía cuando le aplicamos* **P**.

Al escalar λ se le conoce por ***valor propio*** *asociado al vector propio* **v**.

Definición A.1.9. *Sea λ_v un valor propio asociado a un vector propio* **v** *de la matriz* **P**. *Definimos la* ***multiplicidad geométrica*** *del valor propio λ_v, como la dimensión del espacio generado por la base necesaria para obtener el vector asociado* **v**, *y la denotamos por $m_g(\lambda_v)$.*

Por otro lado, la ***multiplicidad algebraica*** *de λ_v, es el orden de λ_v como raíz del polinomio característico de* **P**, *y la denotamos por $m_a(\lambda_v)$.*

Estas multiplicidades cumplen la siguiente desigualdad:

$$m_g(\lambda_v) \leq m_a(\lambda_v).$$

Definición A.1.10. *Sea* **P** *la matriz cuadrada $N \times N$ sobre $\mathbb{R}$. El* ***polinomio característico*** *de* **P** *es el polinomio calculado como sigue:*

$$det(\mathbf{P} - x_i \mathbf{Id}),$$

donde $\mathbf{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$, $x_i\mathbf{Id}$ denota la matriz diagonal:

$$x_i\mathbf{Id} = \begin{pmatrix} x_1 & 0 & \cdots & \cdots & 0 \\ 0 & x_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_{N-1} & 0 \\ 0 & \cdots & \cdots & 0 & x_N \end{pmatrix}$$

y $det(\mathbf{P} - x_i\mathbf{Id})$ es el determinante de la matriz $\mathbf{P} - x_i\mathbf{Id}$.

Este polinomio es de grado N y sus raíces son los valores propios de la matriz **P** *y, con ellos, podemos averiguar también los vectores propios asociados.*

Definición A.1.11. *Supongamos que el polinomio característico de la matriz* **P** *factoriza completamente sobre $\mathbb{R}$ (o sea, todas las raíces del polinomio pertenecen a $\mathbb{R}$). Sean $\lambda_1, \ldots, \lambda_k$ los valores propios, diferentes entre sí, asociados a los vectores propios $v_1, \ldots, v_k$ de* **P**. *Entonces existe una base para la cuál* **P** *se transforma en una matriz diagonal por bloques. Esta descomposición se conoce como* ***descomposición de Jordan***, *y se obtiene de la siguiente forma:*

$$\mathbf{P} = \mathbf{SJS}^{-1},$$

donde **J** *es la matriz diagonal por bloques:*

$$\mathbf{J} = \begin{pmatrix} \mathbf{A}_1 & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{A}_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{A}_{k-1} & 0 \\ 0 & \cdots & \cdots & 0 & \mathbf{A}_k \end{pmatrix},$$

con $\mathbf{A}_i$*, e* $i \in \{1, \ldots, k\}$*, un* ***bloque de Jordan*** *t.q.*

$$\mathbf{A}_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda_i & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{pmatrix}_{n_i \times n_i},$$

donde n_i *es la multiplicidad algebraica del valor propio* λ_i*. Y* $\mathbf{S}$ *y* $\mathbf{S}^{-1}$ *son las matrices del cambio de base y su inversa, respectivamente. Notemos que* $\sum_{i=1}^{k} n_i = N$*.*

Por otro lado, si se cumpliera:

$$m_g(\lambda_i) = m_a(\lambda_i),\ \forall\, i \in \{1, \ldots, k\},$$

entonces $\mathbf{P}$ *es* ***diagonalizable****. Esto significaría que existe una base para la cuál* $\mathbf{P}$ *se transforma en una matriz diagonal:*

$$\mathbf{P} = \mathbf{U}\mathbf{D}\mathbf{U}^{-1},$$

con $\mathbf{U}$ *y* $\mathbf{U}^{-1}$ *las matrices del cambio de base y su inversa, respectivamente y* $\mathbf{D}$ *la matriz diagonal:*

$$\mathbf{D} = \begin{pmatrix} \mathbf{A}_1 & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{A}_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{A}_{k-1} & 0 \\ 0 & \cdots & \cdots & 0 & \mathbf{A}_k \end{pmatrix},$$

con $\mathbf{A}_i$ *una matriz diagonal t.q.*

$$\mathbf{A}_i = \begin{pmatrix} \lambda_i & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_i & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \lambda_i & 0 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{pmatrix}_{n_i \times n_i}$$

$\forall\, i \in \{1, \ldots, k\}$.

A.2 Introducción a los Grafos Dirigidos

Vamos a realizar una pequeña introducción sobre grafos dirigidos definiendo una serie de conceptos. La notación que expondremos a continuación será la usada a lo largo del trabajo.

Definición A.2.1. *Un* ***Grafo Dirigido*** *está definido por un par de conjuntos* (S,E)*, donde* S *es un conjunto finito o numerable de puntos llamados* ***nodos****, y* E *es un conjunto ordenado de pares de nodos tt.q.*

$$E = \{(i,j) \,/\, i,j \in S\}$$

llamados ***arcos****. Cuando escribimos* (i,j) *decimos que el arco* (i,j) *va de* i *a* j*.*

Definición A.2.2. *Un* ***camino dirigido*** *de* i *a* j *es una unión finita de arcos de la forma:*

$$(i,i_1)(i_1,i_2)\cdots(i_k,j).$$

Definición A.2.3. *Un grafo dirigido puede ser representado por una matriz* $\mathbf{A}$ *indexada por* S*, llamada* ***matriz de incidencia*** *del grafo, y definida como sigue:*

$$\mathbf{A}^i_j = \begin{cases} 1 & si\,(i,j) \in E \\ 0 & sinó \end{cases}$$

Nota A.2.1. *A raíz de esta definición, para cada* $i,j \in S$ *tenemos que:*

$$(\mathbf{A}^2)^i_j = \sum_{k\in S} \mathbf{A}^i_k \mathbf{A}^k_j > 0 \Leftrightarrow \exists k \in S \; t.q. \; \mathbf{A}^i_k \mathbf{A}^k_j = 1.$$

Esto quiere decir que si la posición (i,j) *de la matriz* $\mathbf{A}^2$ *es mayor que* 0*, entonces existe al menos un camino dirigido que va de* i *a* j *en máximo* 2 *pasos (o* 2 *aristas). Podría ser que* $\mathbf{A}^i_j > 0$*, o que exista un camino dirigido de la forma*

$$(i,k)(k,j),$$

con $k \in S$*.*

Mediante el método de inducción podríamos demostrar esto para n*. Este caso era para* $n = 2$*.*

Definición A.2.4. *Un* ***Grafo Dirigido con pesos*** *es un grafo dirigido t.q. existe un peso* p_{ij} *adjunto a cada arco* (i,j) *del grafo. La matriz* $\mathbf{A}$*, de la Definición A.2.3 anterior, nos quedaría de la siguiente forma:*

$$\mathbf{A}^i_j = \begin{cases} p_{ij} & si\,(i,j) \in E \\ 0 & sinó \end{cases}$$

Así tendríamos un grafo dirigido con pesos representado por una matriz.

A.3 Probabilidad y Estadística

Todo lo que vamos a definir en esta sección es importante que se entienda, debido a que nuestro trabajo trata fundamentalmente sobre la probabilidad.

Definición A.3.1. *Sea* Ω *un espacio muestral y* $\mathscr{E}$ *una* σ*-álgebra*[2] *sobre* Ω*. Una* ***medida de probabilidad****, que le llamaremos* $\mathbb{P}$*, es una medida que le asigna a cada subconjunto de* $\mathscr{E}$ *un valor perteneciente al intervalo* $[0,1]$ *y cumple con las siguientes condiciones:*

[2] Consúltese la Definición A.4.2 del Apéndice A para más información.

1) *A cada suceso* $E \in \mathscr{E}$ *le corresponde un número real* $\mathbb{P}(E)$ *t.q.* $0 \leq \mathbb{P}(E) \leq 1$.

2) $\mathbb{P}(\Omega) = 1$.

3) *Sean* $E_1, E_2, \dots \in \mathscr{E}$ *tt.q. los subíndices pertenecen a un conjunto* I *finito o numerable y que los sucesos son disjuntos dos a dos* ($E_i \cap E_j = \emptyset$, $\forall i, j \in I$). *Entonces*

$$\mathbb{P}\left(\bigcup_{i \in I} E_i\right) = \sum_{i \in I} \mathbb{P}(E_i).$$

Definición A.3.2. *Un **espacio de probabilidad*** $(\Omega, \mathscr{E}, \mathbb{P})$ *está compuesto por* 3 *componentes:* Ω, *el espacio muestral, que es el conjunto de los posibles resultados del experimento que llevamos a cabo,* $\mathscr{E}$, *la* σ*-álgebra sobre* Ω, *y* $\mathbb{P}$, *que es la medida de probabilidad.*

Al par $(\Omega, \mathscr{E})$ *se le conoce como **espacio medible**.*

A partir de ahora, a no ser que se defina de manera diferente, todos los conceptos estarán definidos sobre el espacio de probabilidad $(\Omega, \mathscr{E}, \mathbb{P})$ y esta será la notación que utilizaremos.

Definición A.3.3. *Sea* S *un conjunto. Una **variable aleatoria** (v.a.) es una aplicación*

$$\begin{array}{rccc} \mathscr{X}: & \Omega & \longrightarrow & S \\ & x & \longrightarrow & \mathscr{X}(x) = i \end{array}$$

t.q. a cada resultado, $x \in \Omega$, *posible del experimento, se le asigna un valor* $i \in S$.

Se dice que $\mathscr{X}$ *es **discreta** si* S *es finito o numerable, sino es **continua**.*

Definición A.3.4. *Sea* $\mathscr{X}$ *una v.a. discreta y* S *un conjunto donde toma los valores* $\mathscr{X}$.

*Una **función de probabilidad*** ρ, *asocia a cada elemento de* Ω *la probabilidad de que ocurra. Lo denotaremos por* $\rho(i) = \mathbb{P}(\mathscr{X}(x) = i)$, *con* $x \in \Omega$ *y* $i \in S$.

*La **función de distribución*** F, *describe la probabilidad de que* $\mathscr{X}(x)$ *sea menor o igual que* $i \in S$ *para un* $x \in \Omega$ *concreto. Esta función la podemos definir gracias a la función de probabilidad:*

$$F(i) = \sum_{\substack{j \in S, \\ j \leq i}} \mathbb{P}(\mathscr{X}(x) = j).$$

Definición A.3.5. *Sea* $\mathscr{X}$ *una v.a. continua y* S *un conjunto donde toma los valores* $\mathscr{X}$.

*Una **densidad*** f, *describe las probabilidades de que la v.a. tome ciertos valores en diferentes intervalos de* S.

En este caso, la función de distribución F, *es una función de los valores* i:

$$F(i) = \mathbb{P}(\mathscr{X} \leq i).$$

Una densidad y la función de distribución están estrechamente relacionadas:

$$F(i) = \int_{-\infty}^{i} f(t)dt \quad \text{y} \quad f(i) = \frac{\partial}{\partial i} F(i).$$

Por ello, si conocemos una de las dos funciones, conoceremos la otra.

Definición A.3.6. *Sea $\Omega = \{x_0, x_1\}$ el espacio muestral y $\mathscr{X}$ una v.a. que toma valores en un espacio finito $S = \{0,1\}$ de tal forma que, por abuso de notación, escribiremos $\mathscr{X} = a$ en lugar de $\mathscr{X}(x_a) = a$, con $a \in S$.*

Con esta notación, sea $p = \mathbb{P}(\mathscr{X} = 0)$ y $q = \mathbb{P}(\mathscr{X} = 1) = 1 - p$. Entonces, la función de probabilidad siguiente:

$$\begin{array}{rccc} \rho: & S & \longrightarrow & [0,1] \\ & 0 & \longrightarrow & p \\ & 1 & \longrightarrow & q \end{array}$$

se conoce como la ***función de probabilidad de Bernoulli*** *o* ***ley de Bernoulli*** *con parámetro p, y se denota por $B(1,p)$. O sea, la v.a. $\mathscr{X}$ sigue una $B(1,p)$.*

El experimento es conocido como ***prueba de Bernoulli****.*

Definición A.3.7. *Sea $\mathscr{X}$ una v.a. que toma valores en $\mathbb{R}$ con densidad dada por:*

$$f(x) = \begin{cases} \frac{1}{b-a} & si\ x \in [a,b] \\ 0 & si\ x \notin [a,b] \end{cases}$$

y función de distribución

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0 & si\ x \leq a \\ \frac{x-a}{b-a} & si\ a \leq x \leq b \\ 1 & si\ x \geq b \end{cases}$$

Entonces $\mathscr{X}$ tiene una ***distribución uniforme continua en $[\mathbf{a},\mathbf{b}]$*** *y se denota por $U([a,b])$. O sea, la v.a. $\mathscr{X}$ sigue una $U([a,b])$.*

Definición A.3.8. *Sean $\mathscr{X}_1$ y $\mathscr{X}_2$ dos vv.aa. y S un e.e.f.n.. Diremos que son* ***independientes*** *si:*

$$\mathbb{P}(\{\mathscr{X}_1(x) = i\} \cap \{\mathscr{X}_2(x) = j\}) = \mathbb{P}(\{\mathscr{X}_1(x) = i\})\mathbb{P}(\{\mathscr{X}_2(x) = j\}),$$

para cualquier $i, j \in S$. Donde $x \in \Omega$ y $\{\mathscr{X}_2(x) = j\}, \{\mathscr{X}_1(x) = i\} \in \mathscr{E}$.

Teorema A.3.1 (Teorema de la Probabilidad Total)**.** *Sea $A_1, A_2, \ldots, A_n$ una partición sobre el espacio muestral Ω. Sea $B \in \mathscr{E}$ un suceso cualquiera t.q. las probabilidades condicionales $\mathbb{P}(B/A_i)$ son conocidas. Entonces:*

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B/A_i)\mathbb{P}(A_i).$$

Definición A.3.9. *Sean A, B dos sucesos de $\mathscr{E}$ cualesquiera con $\mathbb{P}(B) > 0$. La* ***probabilidad condicional*** *de A dado B, es la probabilidad de que ocurra A sabiendo que también ha sucedido B. Se denota por $\mathbb{P}(A/B)$ y se calcula con la siguiente fórmula:*

$$\mathbb{P}(A/B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{A.1}$$

Si A y B son independientes, entonces:

$$\mathbb{P}(A/B) = \mathbb{P}(A) \quad y \quad \mathbb{P}(B/A) = \mathbb{P}(B). \tag{A.2}$$

Definición A.3.10. *La **esperanza** de una v.a. $\mathscr{X}$ se denota por $\mathbb{E}[\mathscr{X}]$ y representaría su valor medio. Dependiendo de si la v.a. es discreta o continua se define de una manera u otra:*

- *Sea $\mathscr{X} : \Omega \longrightarrow S$ la v.a. discreta, entonces si se cumple que:*

$$\sum_{i=1}^{+\infty} |i| \cdot \mathbb{P}(\mathscr{X} = i) < +\infty,$$

  con $i \in S$, o sea, que tiene esperanza finita, tenemos que:

$$\mathbb{E}[\mathscr{X}] = \sum_{i=1}^{+\infty} i \cdot \mathbb{P}(\mathscr{X} = i) < +\infty.$$

- *Sea $\mathscr{X}$ la v.a. continua y $f(i)$ su densidad, con $i \in S$ (S en este caso es infinito y no numerable). Si se cumple que:*

$$\int_{-\infty}^{+\infty} |i| \cdot f(i) di < +\infty,$$

  tenemos que:

$$\mathbb{E}[\mathscr{X}] = \int_{-\infty}^{+\infty} i \cdot f(i) di.$$

Propiedad A.3.1. *Sean $\mathscr{X}, \mathscr{Y}$ vv.aa. con esperanza finita y $a, b \in \mathbb{R}$. Aquí tenemos algunas de las propiedades de la esperanza matemática:*

1) $a\mathscr{X} + b\mathscr{Y}$ *tiene esperanza finitay* $\mathbb{E}[a\mathscr{X} + b\mathscr{Y}] = a\mathbb{E}[\mathscr{X}] + b\mathbb{E}[\mathscr{Y}]$.

2) $\mathbb{E}[a\mathscr{X} + b] = a\mathbb{E}[\mathscr{X}] + b$.

3) $|\mathbb{E}[\mathscr{X}]| \leq \mathbb{E}[|\mathscr{X}|]$.

4) $\mathbb{P}(\mathscr{X} = b) = 1 \Rightarrow \mathbb{E}[\mathscr{X}] = b$.

A.4 Conceptos de Topología

Definición A.4.1. *Sea $X \neq \emptyset$ un conjunto y $\mathfrak{T}$ una familia de subconjuntos de X que satisface:*

1) $\emptyset, X \in \mathfrak{T}$.

2) $A \cap B \in \mathfrak{T}, \forall A, B \in \mathfrak{T}$.

3) $\bigcup_{i \in I} A_i \in \mathscr{T}, \forall A_i \in \mathfrak{T}$ *con $i \in I$, donde I es un conjunto de índices.*

*Entonces $\mathfrak{T}$ es la **topología** en X y el par $(X, \mathfrak{T})$ es conocido como **espacio topológico**. Los elementos de $\mathfrak{T}$ se conocen como **conjuntos abiertos** del espacio, o sea:*

$$U \text{ abierto en } (X, \mathfrak{T}) \Leftrightarrow U \in \mathfrak{T}.$$

Definición A.4.2. *Una familia de subconjuntos de Ω, que llamaremos $\mathscr{E}$, es una **σ-álgebra** sobre Ω si cumple con las condiciones siguientes:*

1) $\emptyset \in \mathscr{E}$.

2) Si $E \in \mathscr{E}$, entonces $\Omega \setminus E \in \mathscr{E}$.

3) Sean $E_1, E_2, \cdots \in \mathscr{E}$ tt.q. los subíndices pertenecen a un conjunto I finito o numerable, entonces

$$\bigcup_{i \in I} E_i \in \mathscr{E}.$$

Definición A.4.3. *La **σ-álgebra de Borel** sobre un espacio topológico $(X, \mathfrak{T})$, es una σ-álgebra de subconjuntos de X, digámosle $\mathscr{B}$, generada por los elementos de $\mathfrak{T}$.*

*Al par $(X, \mathscr{B})$ se le conoce como **espacio de Borel**.*

Ejemplo A.4.1. *Sea $X = \mathbb{R}$ y $(\mathbb{R}, \mathfrak{T})$ un espacio topológico. Una σ-álgebra de Borel, $\mathscr{B}$, sobre $\mathbb{R}$ sería la σ-álgebra generada por los elementos de $\mathfrak{T}$, o sea, los subconjuntos abiertos de $\mathbb{R}$.*

Esta $\mathscr{B}$ también se puede generar mediante: los intervalos abiertos de $\mathbb{R}$, los intervalos cerrados, los intervalos semiabiertos de la forma $(a, b]$, con $a, b \in \mathbb{R}$, tt.q. $a < b$, etc.

Definición A.4.4. *Sea $(\Omega, \mathscr{E})$ un espacio medible. Decimos que $f : (\Omega, \mathscr{E}) \longrightarrow (\Omega, \mathscr{E})$ es una **función $\mathscr{E}$-medible** si preserva la estructura entre dos espacios medibles. O sea, si la anti-imagen de cualquier conjunto medible es medible:*

$$f^{-1}(A) \subseteq B, \quad \forall A, B \in \mathscr{E}.$$

Definición A.4.5. *Sean $(X, \mathscr{B}_1)$ y $(Y, \mathscr{B}_2)$ espacios de Borel. Si $f : (X, \mathscr{B}_1) \longrightarrow (Y, \mathscr{B}_2)$ es una función medible, entonces f es una **función de Borel**.*

Teorema A.4.1 (Teorema de Beppo Levi)**.** *Sea $(\Omega, \mathscr{E})$ un espacio medible y $\{f_k\}$ una sucesión de funciones no-decrecientes $\mathscr{E}$-medibles no-negativas donde:*

$$f_k : \Omega \longrightarrow \overline{\mathbb{R}}_+ \quad y \quad f(x) = \lim_{k \to +\infty} f_k(x).$$

Entonces f es integrable y se cumple la siguiente igualdad:

$$\int_\Omega f(x)\mathbb{P}(dx) = \lim_{k \to +\infty} \int_\Omega f_k(x)\mathbb{P}(dx).$$

Teorema A.4.2 (Fórmula de Cavalieri)**.** *Sea $f : \Omega \longrightarrow \overline{\mathbb{R}}_+$ una función no-negativa $\mathscr{E}$-medible. Entonces tenemos que:*

$$\int_\Omega f(x)\mathbb{P}(x) = \int_0^{+\infty} \mathbb{P}(f(x) > t)dt.$$

Definición A.4.6. *Sea $\{x_n\}$ una sucesión. El* ***límite superior*** *y el* ***límite inferior*** *de $\{x_n\}$ son el mayor y menor límite convergente de las sub-sucesiones*[3] *de $\{x_n\}$ respectivamente.*

El límite inferior de $\{x_n\}$ se calcula de la siguiente forma:

$$\liminf_{n\to+\infty} x_n = \lim_{n\to+\infty}\left(\inf_{k\geq n} x_k\right).$$

Por otro lado, el límite superior de $\{x_n\}$ se calcula como sigue:

$$\limsup_{n\to+\infty} x_n = \lim_{n\to+\infty}\left(\sup_{k\geq n} x_k\right).$$

Si los límites superior e inferior existen, entonces se cumple que:

$$\liminf_{n\to+\infty} x_n \leq \limsup_{n\to+\infty} x_n.$$

Además, si el límite de la sucesión existe, entonces:

$$\lim_{n\to+\infty} x_n = \liminf_{n\to+\infty} x_n = \limsup_{n\to+\infty} x_n.$$

Lema A.4.1 (Lema de Fatou)**.** *Sea $(\mathscr{E},\mathbb{P})$ una medida en Ω y sea $\{f_k\}$ una sucesión de funciones no-negativas $\mathscr{E}$-medibles. Entonces:*

$$\int_\Omega \liminf_{k\to+\infty} f_k(x)\mathbb{P}(dx) \leq \liminf_{k\to+\infty}\int_\Omega f_k(x)\mathbb{P}(dx).$$

Demostración A.4.1. *Sea $g_n(x) = \inf_{k\geq n} f_k(x)$. Entonces $\{g_n(x)\}$ es una sucesión creciente de funciones no-negativas $\mathscr{E}$-medibles. Es una sucesión creciente debido a que si $n_1 > n_2$ tt.q. $n_1, n_2 \in \mathbb{N}$, entonces el ínfimo de los $f_k(x)$ con $k \geq n_2$, será menor o igual que el ínfimo de los $f_k(x)$ con $k \geq n_1$, ya que hay más posibilidades de encontrar un valor más pequeño.*

Además, al ser $g_n(x) = \inf_{k\geq n} f_k(x)$, tenemos que:

a) $0 \leq g_n(x) \leq f_k(x)$, $\forall k \geq n$.

b) $\liminf_{n\to+\infty} f_n(x) = \lim_{n\to+\infty}\left(\inf_{k\geq n} f_n(x)\right) = \lim_{n\to+\infty} g_n(x)$.

c) Por a) y b):

$$\int_\Omega g_n(x)\mathbb{P}(dx) \leq \inf_{k\geq n}\int_\Omega f_k(x)\mathbb{P}(dx).$$

Ahora sí estamos en disposición de demostrar la desigualdad del lema:

$$\int_\Omega \liminf_{k\to+\infty} f_k(x)\mathbb{P}(dx) = \int_\Omega \lim_{n\to+\infty} g_n(x)\mathbb{P}(dx),$$

aplicamos el Teorema de Beppo Levi[4] *y nos queda:*

[3] Véase la Definición A.4.11 del Apéndice A .

[4] Véase el Teorema A.4.1 del Apéndice A para consultar el enunciado.

$$\int_\Omega \lim_{n\to+\infty} g_n(x)\mathbb{P}(dx) = \lim_{n\to+\infty}\int_\Omega g_n(x)\mathbb{P}(dx),$$

aplicamos c):

$$\lim_{n\to+\infty}\int_\Omega g_n(x)\mathbb{P}(dx) \le \lim_{n\to+\infty}\left(\inf_{k\ge n}\int_\Omega f_k(x)\mathbb{P}(dx)\right) = \liminf_{k\to+\infty}\int_\Omega f_k(x)\mathbb{P}(dx).$$

Por tanto:

$$\int_\Omega \liminf_{k\to+\infty} f_k(x)\mathbb{P}(dx) \le \liminf_{k\to+\infty}\int_\Omega f_k(x)\mathbb{P}(dx),$$

tal y como queríamos ver.

□

A.4.1 Conjuntos compactos

Definición A.4.7. *Sea X un conjunto y $A \subseteq X$ un subconjunto. Una familia $\{U_i\}_{i\in I}$ de subconjuntos de X (con I un conjunto numerable de índices) se conoce como* ***cubrimiento de* A** *si $A \subseteq \bigcup_{i\in I} U_i$.*

Definición A.4.8. *Con la misma notación de la definición anterior, un* ***subcubrimiento*** *de un cubrimiento $\{U_i\}_{i\in I}$ de A es una subfamilia $\{U_j\}_{j\in J}$, para algún subconjunto $J \subseteq I$, t.q. $\{U_j\}_{j\in J}$ sea aún un cubrimiento de A. Si J es finito, le llamaremos* ***subcubrimiento finito*** *a la subfamilia $\{U_j\}_{j\in J}$.*

Definición A.4.9. *Consideramos un espacio topológico $(X,\mathfrak{T})$ y un subconjunto $A \subseteq X$. Si $\mathscr{U} = \{U_i\}_{i\in I}$ es un cubrimiento de A y $U_i \in \mathfrak{T}$, $\forall i \in I$, entonces $\mathscr{U}$ es un* ***cubrimiento abierto*** *de A.*

Definición A.4.10. *Un subconjunto A de un espacio topológico $(X,\mathfrak{T})$ es* ***compacto*** *si todo cubrimiento abierto de A tiene un subcubrimiento finito.*

Propiedad A.4.1. *Todo subconjunto compacto de $\mathbb{R}^N$ está acotado y es cerrado en $\mathbb{R}^N$.*

Definición A.4.11. *Sea (x_n) una sucesión en un conjunto X y $n_1 < n_< \cdots < n_r$ una sucesión creciente de números enteros positivos. La sucesión (y_r) dada por:*

$$y_r = x_{n_r}$$

se conoce como ***sub-sucesión*** *de la sucesión (x_n).*

Definición A.4.12. *Un subconjunto $S \subseteq \mathbb{R}$ es* ***secuencialmente compacto*** *si toda sucesión en S tiene al menos una sub-sucesión que converge a un punto de S.*

Teorema A.4.3. *Todo espacio métrico X compacto es secuencialmente compacto.*

A.4.2 Espacios métricos completos

Definición A.4.13. *Un espacio métrico* (X, d) *es* ***completo*** *si toda sucesión de Cauchy en* X *converge a un punto de* X.

Ejemplo A.4.2. *El espacio métrico* $(\mathbb{R}^n, d)$*, con* d *la distancia de la Definición A.1.4, es completo* $\forall n \in \mathbb{N}$.

Definición A.4.14. *Una* ***aplicación contractiva*** *o* ***contracción*** *en el espacio métrico* (X, d)*, con* d *la distancia A.1.4, es una aplicación* $f : X \longrightarrow X$ *que reduce distancias uniformemente, o sea, analíticamente sería:*

Una aplicación para la cual existe una constante L*, con* $0 < L < 1$*, t.q.*

$$d(f(x) - f(y)) \leq L \cdot d(x - y),\ \forall x, y \in X.$$

Decimos que L *es el* ***factor de contracción*** *de* f.

Propiedad A.4.2. *Si* f *es una aplicación contractiva en* X*, entonces es continua en* X.

Demostración A.4.2. *Sea* $\epsilon > 0$ *y* $x_0 \in X$ *arbitrarios pero fijados y* $\delta = \epsilon$*. Queremos ver que si* $d(x - x_0) < \delta$ *entonces* $d(f(x) - f(x_0)) < \epsilon$ *para ver así que* f *es continua:*

Por ser f *contractiva y* $d(x - x_0) < \delta$*:*

$$d(f(x) - f(x_0)) \leq L \cdot d(x - x_0) < d(x - x_0) < \delta = \epsilon.$$

Por tanto f *es continua.*

$\square$

A.5 Teoría del Orden

Definición A.5.1. *Una* ***relación matemática****, denotada por* $\mathscr{R}$*, de los conjuntos* $A_1, \ldots, A_n$ *es un subconjunto del producto cartesiano:*

$$\mathscr{R} \subseteq A_1 \times \cdots \times A_n.$$

Definición A.5.2. *Una* ***relación binaria****, denotada por* $\mathscr{R}$*, es una relación matemática definida entre los elementos de dos conjuntos* A *y* B.

Esta relación se puede representar con pares ordenados (a, b) *tt.q.* $(a, b) \in A \times B$*, para los cuales se cumple una propiedad* $\mathscr{P}$*:*

$$\mathscr{R} = \{(a, b) \in A \times B \,/\, \mathscr{P}(a, b)\}.$$

Definición A.5.3. *Un* ***orden parcial*** *sobre un conjunto* X *es una relación binaria* $\mathscr{R}$ *que es reflexiva, anti-simétrica y transitiva. O sea, para cualesquiera* $a, b, c \in X$*:*

1) *Si* $a\mathscr{R}a$*, entonces* $\mathscr{R}$ *es reflexiva.*

2) *Si* $a\mathscr{R}b$ *y* $b\mathscr{R}a$ *implica que* $a = b$*, entonces* $\mathscr{R}$ *es anti-simétrica.*

3) *Si* $a\mathscr{R}b$ *y* $b\mathscr{R}c$ *implica que* $a\mathscr{R}c$*, entonces* $\mathscr{R}$ *es transitiva.*

Definición A.5.4. *Un* ***conjunto parcialmente ordenado*** *o* ***poset*** *es un conjunto que cuenta con un orden parcial para relacionar sus elementos.*

A.6 Pequeña introducción a los números complejos

Definición A.6.1. *Los **números complejos** son una extensión de los números reales. Sea z un número complejo arbitrario pero fijado, se puede representar mediante la suma de dos valores: un valor real, llamado $Re(z)$ y un valor imaginario, llamado $Im(z)$.*

*La parte imaginaria es un múltiplo real de **la unidad imaginaria** $\mathbf{i} = \sqrt{-1}$. De esta forma, el complejo z anterior quedaría representado como sigue:*

$$z = a + bi, \text{ t.q. } a, b \in \mathbb{R},$$

donde $a = Re(z)$ y $b = Im(z)$. Al conjunto de los números complejos lo denotamos por $\mathbb{C}$. Obviamente $\mathbb{R} \subset \mathbb{C}$.

Definición A.6.2. *Sea $z \in \mathbb{C}$ t.q. $z = a + bi$. Se define el **módulo** de z o **valor absoluto** de z como sigue:*

$$|z| = \sqrt{a^2 + b^2}.$$

Este módulo, además, cumple las siguientes propiedades:

a) $|z| = 0 \Leftrightarrow z = 0$.

b) $|z + w| \leq |z| + |w|$.

c) $|zw| = |z||w|$.

d) $|z - w| \geq ||z| - |w||$.

Donde $z, w \in \mathbb{C}$. Notemos que estas propiedades son semejantes a las propiedades que cumple el valor absoluto para los números reales.

A.7 Aplicaciones iteradas y puntos fijos

En esta sección se cumplirá siempre lo siguiente:

Sea (X, d) un espacio métrico[5] y $f : X \longrightarrow X$ una aplicación lineal continua de X en sí misma.

Cada vez que expongamos una definición, proposición o teorema nuevo en esta sección, tendremos en cuenta las condiciones y notaciones que habremos usado hasta la definición, proposición o teorema justo anterior a éste, junto con las nuevas condiciones o notaciones que expliquemos en el mismo.

Definición A.7.1. *La función $f^k : X \longrightarrow X$ es la **iteración k-ésima de f**, o sea:*

$$f^k = \underbrace{f \circ f \circ \cdots \circ f}_{k \text{ veces}}.$$

*Por conveniencia, denotaremos por $f^0(x) = x$ a **la iteración 0-ésima**, que se correspondería con la función identidad.*

[5] Véase la Definición A.1.4 del Apéndice A.

Definición A.7.2. *Para cada $x \in X$, la sucesión $\{f^k(x)\}_{k\in\mathbb{N}}$ se conoce como* ***el camino de x*** *definido por f.*

Definición A.7.3. *Decimos que $p \in X$ es un* ***punto fijo para f*** *si cumple:*

$$f(p) = p.$$

Definición A.7.4. *Si para cada $x \in X$ se cumple:*

$$f^k(x) \xrightarrow[k\to\infty]{} p.$$

Entonces p es un ***punto fijo atractivo de f****.*

Proposición A.7.1. *Las siguientes afirmaciones se mantienen ciertas:*

i) Si p es un punto fijo para f, entonces para cada $k \in \mathbb{N} \cup \{0\}$ se tiene que $f^k(p) = p$. O sea, p es punto fijo de f^k, para cada k.

ii) Sea $n \in \mathbb{N}$ y p un punto fijo de f^n, entonces $p, f(p), \ldots, f^{n-1}(p)$ son puntos fijos de f^n también.

En particular, si f^n tiene un único punto fijo, para algún n, entonces p es el único punto fijo de f.

iii) Si para algún $x \in X$ se tiene que

$$f^k(x) \xrightarrow[k\to\infty]{} p,$$

entonces p es un punto fijo de f.

iv) Si p es el punto fijo atractivo de f, entonces p es el único punto fijo de f.

v) Existen aplicaciones que tienen puntos fijos pero no puntos fijos atractivos.

vi) Sea $n \in \mathbb{N}$, entonces p es el punto fijo atractivo de f sii p es el punto fijo atractivo de f^n.

A continuación realizaremos la demostración de cada una de las afirmaciones que hemos expuesto en esta proposición:

Demostración A.7.1. *Realizaremos las demostraciones por orden:*

i) Realizaremos la demostración por inducción sobre k:

- *Como caso base, sea $k = 1$. Por hipótesis sabemos que p es punto fijo de f, por tanto*

$$f^1(p) = f(p) = p.$$

- *Paso de inducción: supongamos que la implicación es cierta para* k *y veamos si se cumple para* $k+1$*:*

$$f^{k+1}(p) = f^k \circ (f(p)),$$

por hipótesis $f(p) = p$*, entonces:*

$$f^k \circ (f(p)) = f^k(p).$$

Ahora, por hipótesis de inducción tenemos que $f^k(p) = p$*, por tanto:*

$$f^{k+1}(p) = p,$$

tal y como queríamos ver.

En definitiva, por inducción hemos demostrado que para cada $k \in \mathbb{N} \cup \{0\}$ *se tiene que* $f^k(p) = p$*.*

ii) Sea p *un punto fijo de* f^n *con* $n > 1$ *un entero positivo. Realizaremos la demostración por inducción como en el caso anterior. Lo que demostraremos por inducción, concretamente, será:*

$$f^n(f^k(p)) = f^k(p), \forall 0 \leq k \leq n-1.$$

Veámoslo:

- *Caso base: sea* $k = 0$*. Por la Definición A.7.1, sabemos que* $f^0(p) = p$ *y, por hipótesis,* p *es punto fijo de* f^n*, por tanto:*

$$f^n(f^0(p)) = f^n(p) = p = f^0(p).$$

- *Paso de inducción: supongamos que la implicación es cierta para* $k < n-1$ *y veamos si se cumple para* $k+1 \leq n-1$*:*

$$f^n(f^{k+1}(p)) = f^{k+1}(f^n(p)),$$

por hipótesis $f^n(p) = p$*, entonces:*

$$f^{k+1}(f^n(p)) = f^{k+1}(p),$$

tal y como queríamos ver.

En definitiva, por inducción hemos demostrado que para cada $0 \leq k \leq n-1$ *se tiene que* $f^n(f^k(p)) = f^k(p)$*.*

Ahora, si f^n *tiene un único punto fijo, digámosle* p*, entonces* $p = f(p) = \cdots = f^{n-1}(p)$ *por lo que acabamos de ver, y* p *sería, en particular, el único punto fijo de* f*. Esto es debido a que si existiera otro punto fijo, digámosle* q*, entonces por el apartado i) anterior tendríamos que* $f^n(q) = q$*, y esto implicaría que* f^n *no tendría un único punto fijo.*

iii) *Por hipótesis tenemos que* $\lim_{k\to+\infty} f^k(x) = p$*, por tanto:*

$$\lim_{k\to+\infty} f^{k+1}(x) = p.$$

Por otro lado, $f^{k+1}(x) = f(f^k(x))$ *y como* $f^k(x)$ *tiende a* p *y* f *es continua, tenemos que* $f(f^k(x))$ *tiende a* $f(p)$*. En definitiva:*

$$\lim_{k\to+\infty} f^{k+1}(x) = p \;\; y \;\; \lim_{k\to+\infty} f^{k+1}(x) = f(p) \Rightarrow f(p) = p,$$

ya que el límite de una función, si existe, es único. Esto quiere decir que p *es punto fijo para* f*, tal y como queríamos ver.*

iv) *Realizaremos la demostración por reducción al absurdo. Supongamos que* p *no es el único punto fijo para* f*. O sea,* $\exists q \neq p$ *t.q.* $q \in X$ *y* $f(q) = q$*.*

Por el apartado i*), tenemos que* $f^k(q) = q$*,* $\forall k \in \mathbb{N} \cup \{0\}$*, por lo que* $\lim_{k\to+\infty} f^k(q) = q$*.*

Ahora bien, por hipótesis sabemos que $\lim_{k\to+\infty} f^k(x) = p$*,* $\forall x \in X$*, por lo que, en particular* $\lim_{k\to+\infty} f^k(q) = p$*. Esto significa que* $p = q$ *y tendríamos una contradicción, ya que* $p \neq q$*.*

Por tanto p *es el único punto fijo de* f*, tal y como queríamos ver.*

v) *Consideremos el conjunto* $X = [-1, 1]$ *y la aplicación lineal continua* $f : X \longrightarrow X$ *t.q.* $f(x) = -x$*,* $\forall x \in X$*. Notemos que el punto* $0 \in X$ *es un punto fijo para* f*, ya que* $f(0) = 0$*.*

Pero, para $x \in X$ *t.q.* $x \neq 0$*, tenemos que:*

$$f(x) = -x,\; f^2(x) = f(f(x)) = f(-x) = x, \dots,\; f^k(x) = (-1)^k x.$$

Esto quiere decir que $f^k(x)$ *no converge a un valor concreto (que exista), por ello no tiene límite y por tanto* f *no tiene punto fijo atractivo.*

vi) *Empecemos con la implicación hacia la derecha (*$\Rightarrow$*):*

Si $p \in X$ *es el punto fijo atractivo de* f*, entonces* $f^k(x) \xrightarrow[k\to\infty]{} p$*,* $\forall x \in X$ *y, en particular, para un* $n \in \mathbb{N}$ *entero fijado, se cumple que* $f^{nk}(x) \xrightarrow[k\to\infty]{} p$*,* $\forall x \in X$*. Por tanto:*

$$(f^n)^k(x) = f^{nk}(x) \xrightarrow[k\to\infty]{} p,\; \forall x \in X.$$

Así pues, p *es un punto fijo atractivo para* f^n*.*

Veamos ahora la implicación hacia la izquierda ($\Leftarrow$*):*

Por ser p *un punto fijo atractivo de* f^n *tenemos que, por el apartado* iv*),* p *es el único punto fijo de* f^n *y, por el apartado* ii*),* p *es el único punto fijo de* f*.*

Ahora, por ser p *el punto fijo atractivo de* f^n *tenemos que* $f^{nh}(x) \xrightarrow[h\to\infty]{} p$*,* $\forall x \in X$ *y para* $n \in \mathbb{N}$ *entero fijado.*

Para cada $j \in \mathbb{N} \cup \{0\}$, por lo que acabamos de explicar y por la continuidad de f:

$$f^{j+hn}(x) = f^j(f^{hn}(x)) \xrightarrow[h\to\infty]{} f^j(p),\ \forall x \in X,$$

y por el apartado i), al ser p punto fijo de f, tenemos que p es punto fijo de f^j y por ello:

$$f^{j+hn}(x) \xrightarrow[h\to\infty]{} p,\ \forall x \in X.$$

Podemos escribir cualquier k entero positivo como $k = hn + j$, con $j \in \mathbb{N} \cup \{0\}$ (resto de la división de k por n). Por lo que acabamos de ver tendríamos, sustituyendo, que $f^k(x) \xrightarrow[k\to\infty]{} p$, $\forall x \in X$, y así p sería el punto fijo atractivo de f.

$\square$

A continuación expondremos dos famosos teoremas sobre la existencia de puntos fijos. Notaremos que la existencia de dichos puntos, puede ser garantizada bajo ciertas condiciones generales:

Teorema A.7.1 (Teorema de Brouwer)**.** *Sea X un conjunto compacto*[6] *y convexo*[7] *en $\mathbb{R}^N$. Entonces f tiene un punto fijo.*

Demostración A.7.2. *Como hemos expuesto en el Ejemplo A.4.2, tenemos que $(\mathbb{R}^N, d_1)$ es un espacio métrico completo.*

Sea $x \in X \subset \mathbb{R}^N$, consideramos el camino $\{f^n(x)\}$ de x. Para cada n, consideramos:

$$w_n = \frac{1}{n+1}\sum_{k=0}^{n} f^k(x),$$

que es una combinación convexa de $f^0(x), f(x), f^2(x), \ldots, f^n(x)$. Como X es convexo, tenemos que $w_n \in X$, $\forall n \in \mathbb{N}$. Esto quiere decir que la sucesión $\{w_n\}$ está totalmente contenida en X. Por ser X compacto, tenemos que X es secuencialmente compacto, por el Teorema A.4.3, esto quiere decir que existe una sub-sucesión de $\{w_n\}$, digámosle $\{w_{n_k}\}$, t.q. $w_{n_k} \xrightarrow[n_k\to\infty]{} w \in X$. Veamos que $f(w) = w$:

Por cómo hemos definido w_n, por ser f lineal y por definición de la distancia d tenemos que:

$$d_1(f(w_{n_k}), w_{n_k}) = d_1\Big(\frac{1}{n+1}\sum_{k=0}^{n} f^{k+1}(x), \frac{1}{n+1}\sum_{k=0}^{n} f^k(x)\Big) = \frac{1}{n+1} d_1\Big(\sum_{k=0}^{n} f^{k+1}(x), \sum_{k=0}^{n} f^k(x)\Big) =$$

$$\frac{1}{n+1}|f^1(x)_1 + \cdots + f^{n+1}(x)_1 - (f^0(x)_1 + \cdots f^n(x)_1)| + \cdots +$$

$$|f^1(x)_N + \cdots + f^{n+1}(x)_N - (f^0(x)_N + \cdots f^n(x)_N)| =$$

[6] Consúltese la Subsección A.4.1 del Apéndice A para más información.

[7] Véase la Definición A.1.5 del Apéndice A.

$$\frac{1}{n+1}|f^{n+1}(x)_1 - f^0(x)_1| + \cdots + |f^{n+1}(x)_N - f^0(x)_N| =$$

$$\frac{1}{n+1}d_1(f^{n+1}(x), f^0(x)) = \frac{1}{n+1}d_1(f^{n+1}(x), x) \leq \frac{diam(X)}{n+1},$$

donde $diam(X) = \max_{x,y \in X} d_1(x,y)$. *Aplicando límites a cada lado de la desigualdad obtenemos lo siguiente:*

$$\lim_{n_k \to +\infty} d_1(f(w_{n_k}), w_{n_k}) \leq \lim_{n_k \to +\infty} \frac{diam(X)}{n+1} \Rightarrow \lim_{n_k \to +\infty} d_1(f(w_{n_k}), w_{n_k}) \leq 0.$$

Ahora, como $w_{n_k} \xrightarrow[n_k \to \infty]{} w$ *y* f *es continua, tenemos que* $f(w_{n_k}) \xrightarrow[n_k \to \infty]{} f(w)$. *Sustituyendo en la desigualdad anterior y por ser* d *distancia:*

$$\lim_{n_k \to +\infty} d_1(f(w_{n_k}), w_{n_k}) \leq 0 \Rightarrow d_1(f(w), w) \leq 0 \Rightarrow d_1(f(w), w) = 0 \Rightarrow f(w) = w.$$

$\square$

Teorema A.7.2 (Teorema del punto fijo de Banach)**.** *Sea* (X, d) *un espacio métrico completo, con* d *la distancia de la Definición A.1.4, y* $f : X \longrightarrow X$ *una contracción con factor de contracción* $0 < L < 1$. *Entonces* f *tiene un único punto fijo atractivo* $p \in X$.

Además, para cada $x \in X$, $\{f^n(x)\}$ *converge a* p *al menos exponencialmente:*

$$d(f^n(x), p) \leq \frac{L^n}{1-L} d(f(x), x), \; \forall n \in \mathbb{N}. \tag{A.3}$$

Demostración A.7.3. *Esta demostración constará de dos partes: la unicidad y la existencia del punto fijo atractivo:*

- *Empezaremos demostrando la unicidad:*

  Supongamos que existen dos puntos fijos atractivos para f, p *y* q, *tt.q.* $p \neq q$.

  Recordemos por el apartado $iv)$ *de la Proposición A.7.1 que si* p *y* q *son puntos fijos atractivos de* f, *entonces son puntos fijos de* f. *por tanto* p *y* q *son puntos fijos de* f.

  Ahora, como f *es contractiva tenemos que:*

  $$d(p,q) = d(f(p), f(q)) \leq Ld(p,q) \Rightarrow d(p,q) \leq Ld(p,q),$$

  que es una contradicción, debido a que $0 < L < 1$. *Por tanto, si existe, solo puede haber un único punto fijo atractivo.*

- *Vayamos ahora a demostrar la existencia:*

  Sea $x \in X$ *arbitrario pero fijado. Definimos el camino de* x *como* $x_k = f^k(x), \forall k \in \mathbb{N}$, *donde:*

$$\begin{gathered}x_0 = f^0(x) = x\\ x_1 = f(x)\\ x_2 = f^2(x) = f(f(x)) = f(x_1)\\ \vdots\\ x_k = f^k(x) = f(f^{k-1}(x)) = f(x_{k-1})\\ \vdots\end{gathered}$$

Sea ahora $k \in \mathbb{N}$ arbitrario pero fijado. Por ser f contractiva tenemos lo siguiente:

$$d(x_{k+1}, x_k) = d(f(x_k), f(x_{k-1})) \leq Ld(x_k, x_{k-1}) = Ld(f(x_{k-1}), f(x_{k-2})) \leq L(Ld(x_{k-1}, x_{k-2})) =$$

$$L^2 d(x_{k-1}, x_{k-2}) \leq \cdots \leq L^k d(x_1, x_0) = L^k d(f(x), x),$$

o sea que:

$$d(x_{k+1}, x_k) \leq L^k d(f(x), x), \ \forall k \in \mathbb{N}. \tag{A.4}$$

Sean $r, s \in \mathbb{N}$ tt.q. $s \geq r \geq 1$. Por ser d distancia, se cumple la desigualdad triangular, y por ello:

$$d(x_r, x_s) \leq \sum_{k=r}^{s-1} d(x_k, x_{k+1}),$$

por la desigualdad A.4:

$$\sum_{k=r}^{s-1} d(x_k, x_{k+1}) \leq \sum_{k=r}^{s-1} L^k d(x, f(x)) = d(x, f(x)) \sum_{k=r}^{s-1} L^k.$$

Ahora, $\sum_{k=r}^{s-1} L^k$ es una serie geométrica de razón L, por lo que:

$$d(x, f(x)) \sum_{k=r}^{s-1} L^k = d(x, f(x)) \frac{L^r - L^s}{1 - L} < d(x, f(x)) \frac{L^r}{1 - L}.$$

Por tanto,

$$d(x_r, x_s) < d(x, f(x)) \frac{L^r}{1 - L}, \ \forall r, s \in \mathbb{N} \text{ t.q. } s \geq r \geq 1. \tag{A.5}$$

Notemos que a medida que $r \to +\infty$ la ecuación de la derecha de la desigualdad tiende a 0, ya que $0 < L < 1$. Esto quiere decir que la sucesión $\{x_n\}_{n \in \mathbb{N}}$ es de Cauchy, veámoslo:

Lo que debemos ver es que $\forall \epsilon > 0, \exists N \in \mathbb{N}$ t.q. si $n, m > N$, entonces $d(x_m, x_n) < \epsilon$.

Sean $m, n \in \mathbb{N}$ t.q. $m \geq n \geq 1$, por la Desigualdad (A.5) tenemos que

$$d(x_m, x_n) < d(x, f(x))\frac{L^n}{1-L}.$$

Por otro lado, sabemos que $\lim_{n\to+\infty} \frac{L^n}{1-L} d(f(x), x) = 0$. *Esto quiere decir que* $\forall \epsilon > 0$, $\exists n_0 \in \mathbb{N}$ *t.q.* $\frac{L^n}{1-L} d(f(x), x) < \epsilon$, $\forall n > n_0$.

Sea $\epsilon > 0$ *arbitrario pero fijado y* $N = n_0$ *(donde este* n_0 *es el de la definición de límite que acabamos de ver). Tenemos que:*

$$d(x_m, x_n) < d(x, f(x))\frac{L^n}{1-L} < \epsilon,\ \forall n > N.$$

Y, como $m \geq n$, *esto se cumple* $\forall m, n > N$, *y por tanto:*

$$d(x_m, x_n) < \epsilon,\ \forall m, n > N,$$

tal y como queríamos ver.

Por ser X *un espacio métrico completo, tenemos que toda sucesión de Cauchy es convergente en* X, *por tanto* $\{x_n\}$ *converge a un valor* $p \in X$. *Recordemos que* $x_k = f^k(x)$, $\forall k \in \mathbb{N}$, *así que por el apartado* $iii)$ *de la Proposición A.7.1 tenemos que* p *es un punto fijo de* f. *Ahora, el* x *elegido inicialmente era arbitrario pero fijado, así que si elegimos otro* $y \in X$ *t.q.* $x \neq y$ *también llegaríamos a tener que existe un punto fijo para* f, *pero hemos visto antes en la demostración de la unicidad que solo existe un único punto fijo. Esto significa que:*

$$f^k(x) \xrightarrow[k\to\infty]{} p,\ \forall x \in X,$$

y por tanto p *es un punto fijo atractivo de* f.

Sea $n \in \mathbb{N}$ *arbitrario pero fijado, la Desigualdad (A.3) se deduce igual que como hemos deducido la Desigualdad (A.5), veámoslo:*

$$d(f^n(x), p) = d(x_n, p) \leq \sum_{k=n}^{+\infty} L^k d(x, f(x)) = d(x, f(x))\frac{L^n - L^{+\infty}}{1-L} = d(x, f(x))\frac{L^n}{1-L},$$

tal y como queríamos ver.

□

Combinando este Teorema A.7.2 con el apartado $vi)$ de la Proposición A.7.1, tenemos el siguiente corolario:

Corolario A.7.1. *Sea* (X, d) *un espacio métrico completo, con* d *la distancia de la Definición A.1.4. Si para algún* $n \in \mathbb{N}$, f^n *es una contracción con factor de contracción* $0 < L < 1$, *entonces* f *tiene un único punto fijo atractivo* $p \in X$ *y para cada* $x \in X$*:*

$$d(f^k(x), p) \leq \frac{L^{\lfloor \frac{k}{n} \rfloor}}{1-L} \max_{0\leq j\leq n} d(f^n(x_j), x_j),\ \forall n \in \mathbb{N}, \tag{A.6}$$

donde $x_j = f^j(x)$, $\forall j = 0, \ldots, n-1$. *Además, si* X *está acotado, o sea:* $diam(X) < +\infty$, *tenemos que:*

$$d(f^k(x), p) \leq diam(X) \cdot L^{\lfloor k/n \rfloor}, \; \forall k \geq n.$$

Demostración A.7.4. *Por el Teorema A.7.2,* f^n *tiene un punto fijo atractivo* $p \in X$ *y se cumple, por la Desigualdad (A.3), para cualquier* $x \in X$ *que:*

$$d(f^{kn}(x), p) \leq \frac{L^{kn}}{1-L} d(f^n(x), x), \; \forall kn \in \mathbb{N}.$$

Sea $m = kn$, *entonces la desigualdad anterior nos quedaría:*

$$d(f^m(x), p) \leq \frac{L^m}{1-L} d(f^n(x), x), \; \forall m \in \mathbb{N}.$$

Y, por el apartado $vi)$ *de la Proposición A.7.1, tenemos que* p *es un punto fijo atractivo de* f.

Vamos ahora a probar la Desigualdad (A.6):

Sean $k, n \in \mathbb{N}$ *t.q.* $k \geq n$. *Dividimos* k *por* n *y, por la división euclídea, tenemos que* $k = q \cdot n + j$, *con* $0 \leq j < n$ *y* $q = \lfloor \frac{k}{n} \rfloor$. *Sea ahora* $x \in X$ *arbitrario pero fijado y definimos el camino de* x *en* f *como* $x_k = f^k(x)$, $\forall k \in \mathbb{N}$ *(tal y como lo hemos hecho en la demostración del teorema anterior). Así pues:*

$$d(f^k(x), p) = d(f^{qn}(f^j(x)), p) \leq \frac{L^{qn}}{1-L} d(f^n(x_j), x_j)) \leq$$

$$\frac{L^q}{1-L} d(f^n(x_j), x_j)) \leq \frac{L^q}{1-L} \max_{0 \leq j \leq n} d(f^n(x_j), x_j),$$

tal y como queríamos ver.

Ahora demostraremos la segunda parte del corolario suponiendo que $diam(X) < +\infty$:

Mantenemos la igualdad $k = q \cdot n + j$ *para cualquier par de enteros* k *y* n *tt.q.* $k \geq n$. *Así pues:*

$$d(f^k(x), p) = d(f^{qn}(f^j(x)), p).$$

Por el apartado $iv)$ *de la Proposición A.7.1 como* p *es un punto fijo atractivo de* f, *tenemos que* p *es el único punto fijo de* f *y por ello* $f(p) = p$:

$$d(f^{qn}(f^j(x)), p) = d(f^{qn}(f^j(x)), f(p)),$$

y, por el apartado $i)$ *de la misma Proposición A.7.1, se tiene que* p *es también punto fijo de* f^n, *con* $n \in \mathbb{N}$, *por lo que:*

$$d(f^{qn}(f^j(x)), f(p)) = d(f^{qn}(f^j(x)), f^n(p)).$$

Ahora, x_j *es el paso* j*-ésimo del camino marcado por* f *partiendo desde* x, *ya que* $x_j = f^j(x)$. *Pero, por otro lado, podemos empezar un nuevo camino partiendo desde* x_j, *mediante la aplicación* f^n, *de tal forma que*

$$
\begin{gathered}
x_j = (x_j)_0 = f^0(x_j), \\
(x_j)_1 = f^n(x_j), \\
\vdots \\
(x_j)_{q-1} = f^{(q-1)n}(x_j) = f^n(x_j)_{q-2}, \\
(x_j)_q = f^{qn}(x_j) = f^n(x_j)_{q-1}, \\
\vdots
\end{gathered}
$$

Aplicando esto a la igualdad anterior nos queda:

$$d(f^{qn}(f^j(x)), f^n(p)) = d(f^n(f^j(x)_{q-1}), f^n(p)),$$

por ser f^n contractiva y aplicando lo que hemos visto antes reiteradas veces, obtenemos:

$$d(f^n(f^j(x)_{q-1}), f^n(p)) \leq L \cdot d(f^j(x)_{q-1}, p) = L \cdot d(f^n(f^j(x)_{q-2}), f^n(p)) \leq$$

$$L \cdot (L \cdot d(f^j(x)_{q-2}, p)) = L^2 \cdot d(f^j(x)_{q-2}, p) \leq \cdots \leq L^q \cdot d(f^j(x)_0, p) = L^q \cdot d(x_j, p) \leq L^q \cdot diam(X),$$

ya que $diam(X) < +\infty$. En definitiva:

$$d(f^k(x), p) \leq L^q \cdot diam(X),$$

con $q = \lfloor \frac{k}{n} \rfloor$ y $k \geq n$.

$\square$

A.8 Curiosidades

Definición A.8.1. *El **triángulo de Pascal** es una representación ordenada mediante un triángulo, de los coeficientes que aparecen en las potencias del binomio $(a+b)^n$.*

Figura A.1: Imagen de las 11 primeras filas del triángulo de Pascal.

Fijada una fila $i \in \{0,\dots,10\}$ del triángulo de la Figura A.1 (van del 0 al 10 en lugar del 1 al 11 para que se correspondan los coeficientes de las potencias del binomio con los valores de las filas del triángulo de la imagen), los valores de i se corresponden con los coeficientes de los términos del binomio $(a+b)^i$, con $a, b \in \mathbb{N}$.

Los valores de la fila $i+1$ se obtienen sumando los dos valores de justo encima que se encuentran en la fila i, $\forall i \in \mathbb{N}$.

Por ejemplo, si nos fijamos en la Figura A.1, el tercer valor, empezando por la izquierda, de la fila 4 (donde la primera fila es la 0) es un 6, que se obtiene sumando el segundo y tercer valor de la fila 3.

A.9 Más sobre las CC.M.

En esta sección explicaremos un concepto y un teorema sobre las CC.M. que, por espacio, no hemos podido introducir en el nudo de este trabajo. Este teorema es de vital importancia para realizar las demostraciones (de las proposiciones y teoremas del nudo del trabajo) que se encuentran en el Apéndice A y B:

Primero necesitamos definir un concepto que nos hará falta para entender mejor el teorema:

Definición A.9.1. *Sea* $(\Omega, \mathscr{E}, \mathbb{P})$ *un espacio de probabilidad,* $\{\mathscr{X}_n\}$ *un proceso estocástico con valores en* S *y* $A \in \mathscr{E}$ *un suceso. Decimos que* A *está* ***identificado*** *por las vv.aa.* $\mathscr{X}_0, \dots, \mathscr{X}_{n-1}$ *si es la unión disjunta de sucesos del siguiente tipo:*

$$\{x \in \Omega \,/\, \mathscr{X}_{n-1}(x) = i_{n-1}, \dots, \mathscr{X}_1(x) = i_1, \mathscr{X}_0(x) = i_0\},$$

con $i_0, \dots, i_{n-1} \in S$.

Teorema A.9.1. *Sea* $\{\mathscr{X}_n\}$ *un proceso estocástico.*

La sucesión $\{\mathscr{X}_n\}$ *es una C.M. sii para cualesquiera 3 enteros* r, n *y* k *con* $0 \le r < n < n+k$ *y para cualesquiera 3 sucesos:* G *identificado por* $\mathscr{X}_0, \dots, \mathscr{X}_r$*,* F *identificado por* $\mathscr{X}_{r+1}, \dots, \mathscr{X}_n$ *y* E *identificado por* $\mathscr{X}_{n+1}, \dots, \mathscr{X}_{n+k}$*, tenemos que*

$$\mathbb{P}(E/\{F \cap G\}) = \mathbb{P}(E/F). \tag{A.7}$$

Además, si $\mathbb{P}(F \cap G) > 0$ *entonces:*

$$\mathbb{P}(\{E \cap F\}/G) = \mathbb{P}(E/F)\mathbb{P}(F/G). \tag{A.8}$$

Demostración A.9.1. *Primero veremos la implicación de izquierda a derecha* ($\Rightarrow$)*, y luego la implicación de derecha a izquierda* ($\Leftarrow$)*:*

- *Realizaremos la demostración por inducción sobre el entero* k*:*

  - *Caso base: para* $k = 1$ *hay que ver que para* $0 \le r < n < n+1$ *con los sucesos definidos como:* G *identificado por* $\mathscr{X}_0, \dots, \mathscr{X}_r$*,* F *identificado por* $\mathscr{X}_{r+1}, \dots, \mathscr{X}_n$ *y* E *identificado por* $\mathscr{X}_{n+1}$*, se tiene que* $\mathbb{P}(E/\{F \cap G\}) = \mathbb{P}(E/F)$*.*
    Sea $\{i_n\}$ *una sucesión de valores de* S *y*

$$A_n = \{x \in \Omega \,/\, \mathscr{X}_n(x) = i_n\} \in \mathscr{E},$$

para cada $n \in \mathbb{N} \cup \{0\}$. *Fijemos un* $n \in \mathbb{N}$ *arbitrario (mayor o igual que* 1 *porque* $0 \leq r < n$ *y* r *y* n *son enteros). Para este* n*, sea* $A_n \cap B = F \cap G$*, donde* $B \in \mathscr{E}$ *está identificado por las vv.aa.* $\mathscr{X}_0, \ldots, \mathscr{X}_{n-1}$. *Definimos pues* B *como:*

$$B = \bigcup_{i \in I} C_i, \quad C_i \cap C_j = \emptyset, \quad \forall i, j \in I \text{ t.q. } i \neq j,$$

con I *un conjunto de índices numerable y los* C_i *sucesos de la forma*

$$C_i = \{x \in \Omega \,/\, \mathscr{X}_{n-1}(x) = i_{n-1}, \ldots, \mathscr{X}_0(x) = i_0\}.$$

Sea $F = A_n$*,* $G = B$ *y* $E = A_{n+1}$. *Consideramos* $r = n-1$. *Tenemos que:*

$$\mathbb{P}(E/\{F \cap G\}) = \mathbb{P}(A_{n+1}/\{A_n \cap B\}) =$$

$$\mathbb{P}\left(A_{n+1}/\left\{A_n \cap \bigcup_{i \in I} C_i\right\}\right) = \mathbb{P}\left(A_{n+1}/\left\{\bigcup_{i \in I}(A_n \cap C_i)\right\}\right),$$

suponiendo que $\mathbb{P}(A_n \cap B) > 0$*:*

$$\mathbb{P}\left(A_{n+1}/\left\{\bigcup_{i \in I}(A_n \cap C_i)\right\}\right) = \frac{\mathbb{P}\left(A_{n+1} \cap \left\{\bigcup_{i \in I}(A_n \cap C_i)\right\}\right)}{\mathbb{P}\left(\bigcup_{i \in I}(A_n \cap C_i)\right)} = \frac{\mathbb{P}\left(\bigcup_{i \in I}(\{A_{n+1} \cap A_n\} \cap C_i)\right)}{\mathbb{P}\left(\bigcup_{i \in I}(A_n \cap C_i)\right)},$$

por ser $\mathbb{P}$ *una medida de probabilidad tenemos que la probabilidad de la unión de conjuntos es la suma de probabilidades de cada conjunto de la unión. Y, por la fórmula de las probabilidades condicionadas, tenemos que:*

$$\frac{\mathbb{P}\left(\bigcup_{i \in I}(\{A_{n+1} \cap A_n\} \cap C_i)\right)}{\mathbb{P}\left(\bigcup_{i \in I}(A_n \cap C_i)\right)} = \frac{\sum_{i \in I} \mathbb{P}(\{A_{n+1} \cap A_n\} \cap C_i)}{\sum_{i \in I} \mathbb{P}(A_n \cap C_i)} = \frac{\sum_{i \in I} \mathbb{P}(\{A_{n+1} \cap A_n\}/C_i)\mathbb{P}(C_i)}{\sum_{i \in I} \mathbb{P}(A_n/C_i)\mathbb{P}(C_i)},$$

ahora, por la Propiedad de Markov 2.2.1, tenemos que $\mathbb{P}(\{A_{n+1} \cap A_n\}/C_i) = \mathbb{P}(\{A_{n+1} \cap A_n\}/D_i)$ *y* $\mathbb{P}(A_n/C_i) = \mathbb{P}(A_n/D_i)$ *con* $D_i = \{x \in \Omega \,/\, \mathscr{X}_{n-1}(x) = i_{n-1}\}$. *Hemos podido aplicarla por hipótesis, ya que estamos demostrando la implicación de izquierda a derecha y* $\{\mathscr{X}_n\}$ *es una C.M.. Aplicando estas igualdades obtenemos:*

$$\frac{\sum_{i \in I} \mathbb{P}(\{A_{n+1} \cap A_n\}/C_i)\mathbb{P}(C_i)}{\sum_{i \in I} \mathbb{P}(A_n/C_i)\mathbb{P}(C_i)} = \frac{\sum_{i \in I} \mathbb{P}(\{A_{n+1} \cap A_n\}/D_i)\mathbb{P}(C_i)}{\sum_{i \in I} \mathbb{P}(A_n/D_i)\mathbb{P}(C_i)},$$

por la fórmula de las probabilidades condicionadas de nuevo, aplicada un par de veces, nos queda:

$$\frac{\sum_{i\in I}\mathbb{P}(\{A_{n+1}\cap A_n\}/D_i)\mathbb{P}(C_i)}{\sum_{i\in I}\mathbb{P}(A_n/D_i)\mathbb{P}(C_i)}=\frac{\sum_{i\in I}\frac{\mathbb{P}(A_{n+1}\cap A_n\cap D_i)}{\mathbb{P}(D_i)}\mathbb{P}(C_i)}{\sum_{i\in I}\frac{\mathbb{P}(A_n\cap D_i)}{\mathbb{P}(D_i)}\mathbb{P}(C_i)}=$$

$$\frac{\sum_{i\in I}\frac{\mathbb{P}(A_{n+1}/\{A_n\cap D_i\})}{\mathbb{P}(D_i)}\mathbb{P}(A_n\cap D_i)\mathbb{P}(C_i)}{\sum_{i\in I}\frac{\mathbb{P}(A_n\cap D_i)}{\mathbb{P}(D_i)}\mathbb{P}(C_i)}=\frac{\sum_{i\in I}\mathbb{P}(A_{n+1}/\{A_n\cap D_i\})\frac{\mathbb{P}(A_n\cap D_i)}{\mathbb{P}(D_i)}\mathbb{P}(C_i)}{\sum_{i\in I}\frac{\mathbb{P}(A_n\cap D_i)}{\mathbb{P}(D_i)}\mathbb{P}(C_i)}.$$

Aplicamos de nuevo la Propiedad de Markov a $\mathbb{P}(A_{n+1}/\{A_n\cap D_i\})$ y tenemos que: $\mathbb{P}(A_{n+1}/\{A_n\cap D_i\})=\mathbb{P}(A_{n+1}/A_n)$. Así pues:

$$\frac{\mathbb{P}(A_{n+1}/A_n)\sum_{i\in I}\frac{\mathbb{P}(A_n\cap D_i)}{\mathbb{P}(D_i)}\mathbb{P}(C_i)}{\sum_{i\in I}\frac{\mathbb{P}(A_n\cap D_i)}{\mathbb{P}(D_i)}\mathbb{P}(C_i)}=\mathbb{P}(A_{n+1}/A_n)=\mathbb{P}(E/F).$$

Como el n inicial fijado era arbitrario, vale para cualquier n. Por tanto, la Igualdad (A.7) se cumple siempre que $\mathbb{P}(F\cap G)=\mathbb{P}(A_n\cap B)>0$, tal y como queríamos ver.

- *Paso de inducción: Supongamos que es cierto para k y veamos si sigue siendo cierto para $k+1$:*

  Hay que ver que para $0\leq r<n<n+k+1$ con los sucesos definidos como: G identificado por $\mathscr{X}_0,\ldots,\mathscr{X}_r$, F identificado por $\mathscr{X}_{r+1},\ldots,\mathscr{X}_n$ y E identificado por $\mathscr{X}_{n+1},\ldots,\mathscr{X}_{n+k+1}$ se tiene que $\mathbb{P}(E/\{F\cap G\})=\mathbb{P}(E/F)$.

  Sea $n\in\mathbb{N}$ arbitrario pero fijado (como en la demostración del caso base anterior). Sean $A_n\cap B=F\cap G$, definidos como en el caso base ($F=A_n$ y $G=B$), y $E=A_{n+k+1}\cap\cdots\cap A_{n+1}$. Consideramos $r=n-1$ y suponemos de nuevo que $\mathbb{P}(F\cap G)=\mathbb{P}(A_n\cap B)>0$. Entonces, por la fórmula de las probabilidades condicionadas aplicada dos veces:

$$\mathbb{P}(E/\{F\cap G\})=\frac{\mathbb{P}(\{A_{n+k+1}\cap\cdots\cap A_{n+1}\}\cap\{A_n\cap B\})}{\mathbb{P}(A_n\cap B)}=$$

$$\frac{\mathbb{P}(A_{n+k+1}/\{A_{n+k}\cap\cdots\cap A_n\cap B\})\mathbb{P}(A_{n+k}\cap\cdots\cap A_n\cap B)}{\mathbb{P}(A_n\cap B)},$$

por la Propiedad de Markov aplicada a $\mathbb{P}(A_{n+k+1}/\{A_{n+k}\cap\cdots\cap A_n\cap B\})$ tenemos que: $\mathbb{P}(A_{n+k+1}/\{A_{n+k}\cap\cdots\cap A_n\cap B\})=\mathbb{P}(A_{n+k+1}/A_{n+k})$. Así pues:

$$\frac{\mathbb{P}(A_{n+k+1}/A_{n+k})\mathbb{P}(A_{n+k}\cap\cdots\cap A_n\cap B)}{\mathbb{P}(A_n\cap B)}=\mathbb{P}(A_{n+k+1}/A_{n+k})\mathbb{P}(\{A_{n+k}\cap\cdots\cap A_{n+1}\}/\{A_n\cap B\}).$$

Sea $E'=A_{n+k}\cap\cdots\cap A_{n+1}$ (notemos que $E=E'\cap A_{n+k+1}$). Aplicamos la hipótesis de inducción y tenemos que $\mathbb{P}(E'/\{A_n\cap B\})=\mathbb{P}(E'/\{F\cap G\})=\mathbb{P}(E'/F)$.

Así pues, sustituyendo en la fórmula de la derecha de la igualdad anterior y aplicando la fórmula de la probabilidad condicionada nos queda:

$$\mathbb{P}(A_{n+k+1}/A_{n+k})\mathbb{P}(E'/F) = \frac{\mathbb{P}(A_{n+k+1}/A_{n+k})\mathbb{P}(E'\cap F)}{\mathbb{P}(F)},$$

de nuevo, por la Propiedad de Markov aplicada al revés, tenemos que $\mathbb{P}(A_{n+k+1}/A_{n+k}) = \mathbb{P}(A_{n+k+1}/\{E'\cap F\})$ *y así obtenemos:*

$$\frac{\mathbb{P}(A_{n+k+1}/\{E'\cap F\})\mathbb{P}(E'\cap F)}{\mathbb{P}(F)} = \frac{\frac{\mathbb{P}(A_{n+k+1}\cap E'\cap F)}{\mathbb{P}(E'\cap F)}\mathbb{P}(E'\cap F)}{\mathbb{P}(F)} =$$

$$\frac{\mathbb{P}(A_{n+k+1}\cap E'\cap F)}{\mathbb{P}(F)} = \mathbb{P}(\{A_{n+k+1}\cap E'\}/F) = \mathbb{P}(E/F).$$

Por tanto, por inducción, tenemos que para cualquier n y para cualquier k enteros positivos tt.q. $0 \le r < n < n+k$, *con r también entero positivo, los* 3 *sucesos G, F y E (definidos como en el enunciado del teorema) cumplen la Igualdad (A.7) para* $\mathbb{P}(F\cap G) > 0$.

Para acabar la demostración de la implicación hacia la derecha, nos queda ver que $\mathbb{P}(\{E\cap F\}/G) = \mathbb{P}(E/F)\mathbb{P}(F/G)$. *Las igualdades que veremos a continuación se basan en aplicar repetidas veces la fórmula de las probabilidades condicionadas:*

$$\mathbb{P}(\{E\cap F\}/G) = \frac{\mathbb{P}(E\cap F\cap G)}{\mathbb{P}(G)} = \frac{\mathbb{P}(E\cap F\cap G)}{\mathbb{P}(F\cap G)}\frac{\mathbb{P}(F\cap G)}{\mathbb{P}(G)} = \mathbb{P}(E/\{F\cap G\})\mathbb{P}(F/G),$$

por la Igualdad (A.7), que ya hemos demostrado, tenemos que $\mathbb{P}(E/\{F\cap G\}) = \mathbb{P}(E/F)$, *por tanto:*

$$\mathbb{P}(E/\{F\cap G\})\mathbb{P}(F/G) = \mathbb{P}(E/F)\mathbb{P}(F/G),$$

y, en definitiva:

$$\mathbb{P}(\{E\cap F\}/G) = \mathbb{P}(E/F)\mathbb{P}(F/G),$$

siempre y cuando $\mathbb{P}(F\cap G) > 0$ *y* $\mathbb{P}(G) > 0$.

- *Veamos ahora la implicación hacia la izquierda. Queremos, por tanto, demostrar que* $\{\mathscr{X}_n\}$ *cumple la Propiedad de Markov 2.2.1, o sea, ver que para r, n tt.q. $r \le n$ y $i_0, \ldots, i_{n+1} \in S$,*

$$\mathbb{P}(A_{n+1}/\{A_n\cap\cdots\cap A_r\}) = \mathbb{P}(A_{n+1}/A_n).$$

Donde los A_n, para cada $n \in \mathbb{N}\cup\{0\}$, son los definidos al inicio de esta demostración del teorema.

Sea $n \in \mathbb{N}$ arbitrario pero fijado y $E = A_{n+1}$, $F = A_n$ y $G = A_{n-1} \cap \cdots \cap A_r$ con $r \leq n-1$. Entonces:

$$\mathbb{P}(A_{n+1}/\{A_n \cap \cdots \cap A_r\}) = \mathbb{P}(E/\{F \cap G\}).$$

Ahora, por hipótesis, se cumple la Igualdad (A.7): $\mathbb{P}(E/\{F \cap G\}) = \mathbb{P}(E/F)$ (siempre y cuando $\mathbb{P}(F \cap G) > 0$). Así pues:

$$\mathbb{P}(E/\{F \cap G\}) = \mathbb{P}(E/F) = \mathbb{P}(A_{n+1}/A_n).$$

Por tanto, $\forall r, n$ enteros positivos tt.q. $r \leq n$, y $i_0, \ldots, i_{n+1} \in S$, se tiene que

$$\mathbb{P}(A_{n+1}/\{A_n \cap \cdots \cap A_r\}) = \mathbb{P}(A_{n+1}/A_n).$$

Así $\{\mathscr{X}_n\}$ cumple la Propiedad de Markov 2.2.1 y es, por tanto, una C.M. por la Definición 2.2.5.

□

A raíz de este teorema, podemos deducir la Igualdad (A.9) que nos será muy útil en demostraciones del Apéndice A y B posteriores. Sea $\{\mathscr{X}_n\}$ una C.M., y sean los enteros r, k y n y los sucesos E, G definidos como en el Teorema A.9.1. Sea $\{F_i\}_{i\in I}$ una familia de sucesos de $\mathscr{E}$ disjuntos dos a dos e identificados por las vv.aa. $\mathscr{X}_{r+1}, \ldots, \mathscr{X}_n$, $F = \bigcup_{i\in I} F_i$, donde I es un conjunto de índices numerable.

Entonces, aplicando que la intersección de un conjunto con una unión de conjuntos es la unión de la intersección del conjunto con cada uno de los conjuntos de la unión, y la fórmula de las probabilidades condicionadas, obtenemos:

$$\mathbb{P}\left(\left\{E \cap \left\{\bigcup_{i\in I} F_i\right\}\right\}/G\right) = \mathbb{P}\left(\left\{\bigcup_{i\in I}(E \cap F_i)\right\}/G\right) = \frac{\mathbb{P}\left(G \cap \left\{\bigcup_{i\in I}(E \cap F_i)\right\}\right)}{\mathbb{P}(G)} = \frac{\mathbb{P}\left(\bigcup_{i\in I}(E \cap F_i \cap G)\right)}{\mathbb{P}(G)}.$$

Ahora, por ser $\mathbb{P}$ una medida de probabilidad y los sucesos disjuntos:

$$\frac{\mathbb{P}\left(\bigcup_{i\in I}(E \cap F_i \cap G)\right)}{\mathbb{P}(G)} = \frac{\sum_{i\in I}\mathbb{P}(E \cap F_i \cap G)}{\mathbb{P}(G)} = \frac{\sum_{i\in I}\mathbb{P}(\{E \cap F_i\}/G)\mathbb{P}(G)}{\mathbb{P}(G)} = \sum_{i\in I}\mathbb{P}(\{E \cap F_i\}/G).$$

Finalmente, por el Teorema A.9.1 tenemos que:

$$\sum_{i\in I}\mathbb{P}(\{E \cap F_i\}/G) = \sum_{i\in I}\mathbb{P}(E/F_i)\mathbb{P}(F_i/G). \tag{A.9}$$

A.10 Más sobre las CC.M.HH.

En esta sección explicaremos más conceptos sobre este tipo de cadenas, que pueden ser del interés del lector.

Primero de todo, veamos una propiedad muy interesante con la que cuentan las CC.M.HH.:

Proposición A.10.1 (Propiedad renovadora de las CC.M.HH.)**.** *Sea $\{\mathscr{X}_n\}$ una C.M.H., con S un espacio de estados, y $i_{n+k},\ldots,i_n \in S$. Entonces*

$$\mathbb{P}(\mathscr{X}_{n+k}=i_{n+k},\ldots,\mathscr{X}_{n+1}=i_{n+1}/\mathscr{X}_n=i_n)=\mathbb{P}(\mathscr{X}_k=i_{n+k},\ldots,\mathscr{X}_1=i_{n+1}/\mathscr{X}_0=i_n)$$

para cualesquiera $k,n\in\mathbb{N}\cup\{0\}$ enteros.

Demostración A.10.1. *Sea $A_j=\{x\in\Omega\,/\,\mathscr{X}_j(x)=i_j\}$ con $i_0,i_1,\cdots\in S$ y $j\in I$, donde I es un conjunto numerable de índices. Sean $E=A_{n+2}$, $F=A_{n+1}$, $G=A_n$ y $\mathbf{P}$ la matriz de transición para $\{\mathscr{X}_n\}$.*

Como $\{\mathscr{X}_n\}$ es una C.M.H., en particular, es un C.M., y por ello podemos aplicar la Igualdad (A.8) del Teorema A.9.1 y tenemos que:

$$\mathbb{P}(\{A_{n+2}\cap A_{n+1}\}/A_n)=\mathbb{P}(\{E\cap F\}/G)=\mathbb{P}(E/F)\mathbb{P}(F/G)=$$

$$\mathbb{P}(A_{n+2}/A_{n+1})\mathbb{P}(A_{n+1}/A_n),$$

ahora, por ser $\{\mathscr{X}_n\}$ una C.M.H.:

$$\mathbb{P}(A_{n+2}/A_{n+1})\mathbb{P}(A_{n+1}/A_n)=\mathbf{P}(n+2)^{i_{n+1}}_{i_{n+2}}\mathbf{P}(n+1)^{i_n}_{i_{n+1}}=\mathbf{P}^{i_{n+1}}_{i_{n+2}}\mathbf{P}^{i_n}_{i_{n+1}}.$$

Esto que acabamos de ver, lo podríamos aplicar de manera iterada para calcular

$$\mathbb{P}(A_{n+k},\ldots,A_{n+1}/A_n).$$

Para hacerlo, en lugar de tener el E definido de antes, definiríamos un $E_{n+2}=A_{n+k}\cap\cdots\cap A_{n+2}$, $F_{n+1}=A_{n+1}$ y $G_n=A_n$ para cada iteración en la que calculamos:

$$\mathbb{P}(\{E_{n+2}\cap F_{n+1}\}/G_n)=\mathbb{P}(E_{n+2}/F_{n+1})\mathbb{P}(F_{n+1}/G_n)=\mathbb{P}(E_{n+2}/F_{n+1})\mathbb{P}(A_{n+1}/A_n)=$$

$$\mathbb{P}(E_{n+2}/F_{n+1})\mathbf{P}(n+1)^{i_n}_{i_{n+1}}=\mathbb{P}(E_{n+2}/F_{n+1})\mathbf{P}^{i_n}_{i_{n+1}}.$$

Así pues, para $\mathbb{P}(E_{n+2}/F_{n+1})$, definimos $E_{n+3}=A_{n+k}\cap\cdots\cap A_{n+3},F_{n+2}=A_{n+2}$ y $G_{n+1}=A_{n+1}$ y tendremos que:

$$\mathbb{P}(\{E_{n+2}\cap F_{n+1}\}/G_n)=\mathbb{P}(E_{n+2}/F_{n+1})\mathbf{P}^{i_n}_{i_{n+1}}=\mathbb{P}(\{E_{n+3}\cap F_{n+2}\}/G_{n+1})\mathbf{P}^{i_n}_{i_{n+1}}=$$

$$\mathbb{P}(E_{n+3}/F_{n+2})\mathbb{P}(F_{n+2}/G_{n+1})\mathbf{P}^{i_n}_{i_{n+1}}=\mathbb{P}(E_{n+3}/F_{n+2})\mathbb{P}(A_{n+2}/A_{n+1})\mathbf{P}^{i_n}_{i_{n+1}}=$$

$$\mathbb{P}(E_{n+3}/F_{n+2})\mathbf{P}_{i_{n+2}}^{i_{n+1}}\mathbf{P}_{i_{n+1}}^{i_n}.$$

De esta forma, acabaríamos obteniendo:

$$\mathbb{P}(\mathscr{X}_{n+k}=i_{n+k},\dots,\mathscr{X}_{n+1}=i_{n+1}/\mathscr{X}_n=i_n)=\mathbf{P}_{i_{n+k}}^{i_{n+k-1}}\cdots\mathbf{P}_{i_{n+2}}^{i_{n+1}}\mathbf{P}_{i_{n+1}}^{i_n}. \tag{A.10}$$

Ahora, también tendríamos que

$$\mathbb{P}(\mathscr{X}_k=i_{n+k},\dots,\mathscr{X}_1=i_{n+1}/\mathscr{X}_0=i_n)=\mathbf{P}_{i_{n+k}}^{i_{n+k-1}}\cdots\mathbf{P}_{i_{n+2}}^{i_{n+1}}\mathbf{P}_{i_{n+1}}^{i_n},$$

con $i_{n+k},\dots,i_n\in S$. Ya que la matriz de transición es independiente de n en una C.M.H.. Por tanto, igualando las dos ecuaciones anteriores obtenemos que:

$$\mathbb{P}(\mathscr{X}_{n+k}=i_{n+k},\dots,\mathscr{X}_{n+1}=i_{n+1}/\mathscr{X}_n=i_n)=\mathbb{P}(\mathscr{X}_k=i_{n+k},\dots,\mathscr{X}_1=i_{n+1}/\mathscr{X}_0=i_n),$$

para cualesquiera $k,n\in\mathbb{N}\cup\{0\}$ enteros.

$\square$

Esta propiedad es conocida por el nombre de renovadora, debido a que si tuviéramos que $\mathbb{P}(\mathscr{X}_n=i_n)=\mathbb{P}(\mathscr{X}_0=i_n)$ (en general, para que se cumpla para cualquier $i_n\in S$, deberíamos tener que $\pi(0)=\pi(n)$), entonces:

$$\mathbb{P}(\mathscr{X}_{n+k}=i_{n+k},\dots,\mathscr{X}_{n+1}=i_{n+1}/\mathscr{X}_n=i_n)=\mathbb{P}(\mathscr{X}_k=i_{n+k},\dots,\mathscr{X}_1=i_{n+1}/\mathscr{X}_0=i_n)\Leftrightarrow$$

$$\frac{\mathbb{P}(\{\mathscr{X}_{n+k}=i_{n+k}\}\cap\cdots\cap\{\mathscr{X}_n=i_n\})}{\mathbb{P}(\mathscr{X}_n=i_n)}=\frac{\mathbb{P}(\{\mathscr{X}_k=i_{n+k}\}\cap\cdots\cap\{\mathscr{X}_0=i_n\})}{\mathbb{P}(\mathscr{X}_0=i_n)},$$

y si $\mathbb{P}(\mathscr{X}_n=i_n)=\mathbb{P}(\mathscr{X}_0=i_n)$, tenemos que

$$\frac{\mathbb{P}(\{\mathscr{X}_{n+k}=i_{n+k}\}\cap\cdots\cap\{\mathscr{X}_n=i_n\})}{\mathbb{P}(\mathscr{X}_n=i_n)}=\frac{\mathbb{P}(\{\mathscr{X}_k=i_{n+k}\}\cap\cdots\cap\{\mathscr{X}_0=i_n\})}{\mathbb{P}(\mathscr{X}_0=i_n)}\Leftrightarrow$$

$$\mathbb{P}(\{\mathscr{X}_{n+k}=i_{n+k}\}\cap\cdots\cap\{\mathscr{X}_n=i_n\})=\mathbb{P}(\{\mathscr{X}_k=i_{n+k}\}\cap\cdots\cap\{\mathscr{X}_0=i_n\}).$$

Por ello, si $\{\mathscr{X}_n\}$ es una C.M.H., la Proposición A.10.1 establece que la distribución conjunta de las k vv.aa. siguientes a $\mathscr{X}_n$ sería la misma que la de las k primeras vv.aa. (con $k<n$ donde $k\in\mathbb{N}\cup\{0\}$ y $n\in\mathbb{N}$) como acabamos de ver. De ahí que se diga que $\mathscr{X}_n$ reinicia o renueva la cadena.

A continuación, veremos un teorema relevante en este apartado que se conoce como propiedad fuerte de Markov.

Pero antes, debemos definir un concepto que nos hará falta:

Definición A.10.1. *Sea $\{\mathscr{X}_n\}$ una C.M. en el espacio de probabilidad $(\Omega,\mathscr{E},\mathbb{P})$ con un espacio de estados S.*

Una v.a. T definida como sigue

$$\begin{array}{rccc} T: & \Omega & \longrightarrow & \mathbb{N}\cup\{0\} \\ & x & \longrightarrow & n \end{array}$$

*es un **tiempo de parada** para $\{\mathscr{X}_n\}$, si para cualquier $n \in \mathbb{N} \cup \{0\}$ el conjunto*

$$\{x \in \Omega \,/\, T(x) = n\}$$

está identificado por las vv.aa. $\mathscr{X}_0, \ldots, \mathscr{X}_n$.

Sea $x \in \Omega$, $C \subseteq S$ un subconjunto de estados y $n \in \mathbb{N} \cup \{0\}$. Tenemos los dos ejemplos siguientes para ilustrar este concepto:

Ejemplo A.10.1. *La primera vez de paso de x por C, denotado por $t_C(x)$[8].*

Ejemplo A.10.2. *El número total de visitas de x a C en los primeros n pasos, denotado por $V_C^{(n)}(x)$[9].*

Estos ejemplos los veremos en la Subsección A.10.1 en profundidad. A continuación, el teorema:

Teorema A.10.1 (Propiedad fuerte de Markov)**.** *Sea $\{\mathscr{X}_n\}$ una C.M.H., con valores en un espacio de estados S, definida en un espacio de probabilidad $(\Omega, \mathscr{E}, \mathbb{P})$, T un tiempo de parada para $\{\mathscr{X}_n\}$ y $\mathbf{P}$ la matriz de transición de $\{\mathscr{X}_n\}$.*

Entonces, para cualquier $i, j \in S$ y $n \in \mathbb{N} \cup \{0\}$, tenemos que:

$$\mathbb{P}(\mathscr{X}_{T+n} = j \,/\, T < +\infty, \mathscr{X}_T = i) = \mathbb{P}(\mathscr{X}_n = j \,/\, \mathscr{X}_0 = i) = \mathbf{P}_j^i.$$

Demostración A.10.2. *Sea $T = h \in \mathbb{Z}^+$ t.q. $h < +\infty$ arbitrario pero fijado. Entonces:*

$$\mathbb{P}(\mathscr{X}_{T+n} = j,\, T = h,\, \mathscr{X}_T = i) = \mathbb{P}(\mathscr{X}_{h+n} = j,\, T = h,\, \mathscr{X}_h = i) =$$

$$\mathbb{P}(\mathscr{X}_{h+n} = j \,/\, T = h,\, \mathscr{X}_h = i)\mathbb{P}(T = h,\, \mathscr{X}_h = i).$$

Sabemos que $T = h$ y, por definición de tiempo de parada, que está identificado por las vv.aa. $\mathscr{X}_0, \ldots, \mathscr{X}_h$ (en $\mathscr{X}_h$ sabemos que vale $i \in S$, ya que, para un $x \in \Omega$, una v.a. no puede tener dos estados en el mismo paso). Por ello, como $\{\mathscr{X}_n\}$ es una C.M. en particular, aplicamos la Igualdad (A.7) del Teorema A.9.1, y tenemos que:

$$\mathbb{P}(\mathscr{X}_{h+n} = j \,/\, T = h,\, \mathscr{X}_h = i)\mathbb{P}(T = h,\, \mathscr{X}_h = i) = \mathbb{P}(\mathscr{X}_{h+n} = j \,/\, \mathscr{X}_h = i)\mathbb{P}(T = h,\, \mathscr{X}_h = i).$$

Al ser $\{\mathscr{X}_n\}$ una C.M.H., por la Proposición A.10.1: $\mathbb{P}(\mathscr{X}_{h+n} = j \,/\, \mathscr{X}_h = i) = \mathbb{P}(\mathscr{X}_n = j \,/\, \mathscr{X}_0 = i)$. Por tanto:

$$\mathbb{P}(\mathscr{X}_{T+n} = j,\, T = h,\, \mathscr{X}_h = i) = \mathbb{P}(\mathscr{X}_n = j \,/\, \mathscr{X}_0 = i)\mathbb{P}(T = h,\, \mathscr{X}_h = i) \Leftrightarrow$$

$$\frac{\mathbb{P}(\mathscr{X}_{T+n} = j,\, T = h,\, \mathscr{X}_h = i)}{\mathbb{P}(T = h,\, \mathscr{X}_h = i)} = \mathbb{P}(\mathscr{X}_n = j \,/\, \mathscr{X}_0 = i) \Leftrightarrow$$

$$\mathbb{P}(\mathscr{X}_{T+n} = j \,/\, T = h,\, \mathscr{X}_h = i) = \mathbb{P}(\mathscr{X}_n = j \,/\, \mathscr{X}_0 = i).$$

[8] Véase la Subsección A.10.1 del Apéndice A para más información.

[9] Consúltese la Subsección A.10.1 del Apéndice A para más información.

Como el h inicial elegido era arbitrario, esto se cumple para cualquier $h \in \mathbb{Z}^+$ *t.q.* $h < +\infty$.

En general, por la ecuación anterior, tendríamos que:

$$\mathbb{P}(\mathscr{X}_{T+n} = j \,/\, T < +\infty, \mathscr{X}_T = i) = \mathbb{P}(\mathscr{X}_n = j \,/\, \mathscr{X}_0 = i).$$

$\square$

Lo que este teorema nos viene a decir es que las vv.aa. $\{\mathscr{Y}_n\}$, definidas como $\mathscr{Y}_n(x) = \mathscr{X}_{T(x)+n}(x)$ y restringidas al suceso $\{x \in \Omega \,/\, T < +\infty\} \subset \Omega$, forman una C.M.H. con la misma matriz de transición que la cadena $\{\mathscr{X}_n\}$.

A.10.1 Parámetros característicos

En esta subsección del apéndice mostraremos una serie de conceptos o parámetros característicos de las CC.M.HH.. Algunos de ellos nos serán necesarios para realizar demostraciones del nudo de este trabajo.

Con el fin de no repetir lo mismo para cada concepto, de ahora en adelante (a no ser que digamos lo contrario en la definición de cada parámetro) estas suposiciones se conservarán a lo largo de esta subsección:

- Consideramos el espacio de probabilidad $(\Omega, \mathscr{E}, \mathbb{P})$.
- $\{\mathscr{X}_n\}$ será una C.M.H. con $\mathbf{P} = (p_{ij})$ su matriz de transición.
- S es el espacio de estados.
- Sea $C \subset S$. Si C es un conjunto unitario, o sea, $C = \{j\}$, como notación en los conceptos que describiremos no usaremos $\{j\}$, sino j directamente.
- Si en algún momento no especificamos el suceso $x \in \Omega$, lo hacemos por abuso de notación. Pero se sobreentiende que todos los conceptos están definidos para sucesos x.

Pasos hasta la primera visita:

Recordemos que si x visita C en el momento (o paso) n, quiere decir que $\mathscr{X}_n(x) \in C$. También decimos que x visita C si $\mathscr{X}_n \in C$ para algún $n \in \mathbb{N}$.

Definición A.10.2. *Sea*

$$t_C(x) = \text{mín}\{n \in \mathbb{N} \,/\, \mathscr{X}_n(x) \in C\},$$

el mínimo número de pasos, en el camino x, hasta visitar C. Entonces t_C *se conoce como* **1ž *vez de paso*** *o* ***tiempo de espera hasta llegar a* C.**

Si $t_C(x) = +\infty$ *es debido a que* $\mathscr{X}_n(x) \notin C$, $\forall n$.

Este parámetro, t_C, es una v.a. en el espacio de probabilidad en el que nos encontramos, debido a que asigna a cada resultado del experimento Ω un valor $n \in \mathbb{N}$:

$$\begin{array}{rccc} t_C: & \Omega & \longrightarrow & \mathbb{N} \\ & x & \longrightarrow & n \end{array}$$

Su densidad de masa, $\mathbb{P}_{t_C}$, la obtendríamos calculando $\mathbb{P}(t_C(x) = n)$ para cada $n \in \mathbb{N}$, que equivaldría verbalmente a calcular la probabilidad de que la primera vez de paso por C se de en el paso n para cada n.

Probabilidad de visitar C en el paso n desde i $\in$ S:

Definición A.10.3. *Sea $i \in S$. **La probabilidad de que el camino x visite C en el paso n empezando en el estado i** se define como:*

$$f_{iC}^{(n)}(x) = \mathbb{P}(t_C(x) = n \,/\, \mathscr{X}_0(x) = i) = \mathbb{P}(\mathscr{X}_n(x) \in C, \mathscr{X}_{n-1}(x) \notin C, \ldots, \mathscr{X}_1 \notin C \,/\, \mathscr{X}_0 = i).$$

*Si $i = j$, entonces $f_{jj}^{(n)}$ se conoce como **la probabilidad de retorno al estado j en el paso n**.*

Observación A.10.1. *Sea $B_i = \{\mathscr{X}_0 = i\}$. El conjunto de sucesos $\{B_i\}_{i \in S}$, forma una partición de Ω. Por ello, podemos aplicar el teorema de la probabilidad total y obtenemos la siguiente ecuación:*

$$\mathbb{P}(t_C = n) = \sum_{i \in S} \mathbb{P}(t_C = n \,/\, B_i)\mathbb{P}(B_i) = \sum_{i \in S} f_{iC}^{(n)} \mathbb{P}(\mathscr{X}_0 = i).$$

Fijémonos que el conjunto de sucesos $\{t_C = n\}_{n \in \mathbb{N}}$ forma también una partición de Ω.

Una vez conocemos $f_{iC}^{(n)}$, podemos definir el siguiente concepto que está estrechamente relacionado:

Definición A.10.4. *Definimos **la probabilidad de visitar C (al menos una vez) empezando en i** como sigue:*

$$f_{iC} = \mathbb{P}(t_C < +\infty \,/\, \mathscr{X}_0 = i) = \sum_{n=1}^{\infty} \mathbb{P}(t_C = n \,/\, \mathscr{X}_0 = i) = \sum_{n=1}^{\infty} f_{iC}^{(n)}. \tag{A.11}$$

*Como en la definición anterior: si $i = j$, entonces f_{jj} es conocido como **la probabilidad de retorno al estado j**.*

La siguiente proposición establece una forma práctica de calcular las probabilidades que hemos definido.

Proposición A.10.2. *La probabilidad $f_{iC}^{(n)}$, con $n \in \mathbb{N}$, satisface la igualdad recurrente:*

$$f_{iC}^{(n)} = \begin{cases} \sum\limits_{l \in C} p_{il} & si\ n = 1 \\ \sum\limits_{l \notin C} p_{il} \cdot f_{lC}^{(n-1)} & si\ n \geq 2 \end{cases}$$

Donde $p_{il} = p_{il}^{(1)} = \mathbb{P}(\mathscr{X}_1 = l \,/\, \mathscr{X}_0 = i)$.

Demostración A.10.3. *Primero realizaremos la demostración para el caso $n = 1$ y luego para el caso $n \geq 2$:*

- *Si* $n = 1$ *tenemos que* $f_{iC}^{(1)}$ *es la probabilidad de visitar C desde i en un paso:*

$$f_{iC}^{(1)} = \sum_{l \in C} \mathbb{P}(\mathscr{X}_1 = l \,/\, \mathscr{X}_0 = i) = \sum_{l \in C} p_{il}.$$

- *Sea* $n \geq 2$, $l \in S$ *y*

$$E_C^n = \{x \in \Omega \,/\, \mathscr{X}_1(x) \notin C, \ldots, \mathscr{X}_{n-1}(x) \notin C, \mathscr{X}_n(x) \in C\} = \{x \in \Omega \,/\, t_C(x) = n\},$$

$$E = \{x \in \Omega \,/\, \mathscr{X}_2(x) \notin C, \ldots, \mathscr{X}_{n-1}(x) \notin C, \mathscr{X}_n(x) \in C\},$$

$$F_l = \{x \in \Omega \,/\, \mathscr{X}_1(x) = l\}, \text{ con } l \notin C.$$

Por cómo hemos definido estos conceptos, tenemos que

$$E_C^n = E \cap \left\{ \bigcup_{l \notin C} F_l \right\}.$$

Ahora, aplicando la Igualdad (A.9), con $E = E$, $F = \bigcup_{l \notin C} F_l$ *y* $G = \{\mathscr{X}_0 = i\}$, *obtenemos lo siguiente:*

$$f_{iC}^{(n)} = \mathbb{P}(E_C^n \,/\, \mathscr{X}_0 = i) = \mathbb{P}(E \cap \left\{ \bigcup_{l \notin C} F_l \right\} \,/\, \mathscr{X}_0 = i) = \sum_{l \notin C} \mathbb{P}(E/F_l)\mathbb{P}(F_l/\mathscr{X}_0 = i) \Rightarrow$$

$$f_{iC}^{(n)} = \sum_{l \notin C} \mathbb{P}(E/F_l)\mathbb{P}(F_l/\mathscr{X}_0 = i) \tag{A.12}$$

Identificaremos por separado a $\mathbb{P}(E/F_l)$ *y* $\mathbb{P}(F_l/\mathscr{X}_0 = i)$*:*

$$\mathbb{P}(E/F_l) = \mathbb{P}(\mathscr{X}_2 \notin C, \ldots, \mathscr{X}_{n-1} \notin C, \mathscr{X}_n \in C \,/\, \mathscr{X}_1 = l),$$

por ser $\{\mathscr{X}_n\}$ *una C.M.H.:*

$$\mathbb{P}(\mathscr{X}_2 \notin C, \ldots, \mathscr{X}_{n-1} \notin C, \mathscr{X}_n \in C/\mathscr{X}_1 = l) = \mathbb{P}(\mathscr{X}_1 \notin C, \ldots, \mathscr{X}_{n-2} \notin C, \mathscr{X}_{n-1} \in C/\mathscr{X}_0 = l) = f_{lC}^{(n-1)}.$$

Por otro lado,

$$\mathbb{P}(F_l/\mathscr{X}_0 = i) = \mathbb{P}(\mathscr{X}_1 = l \,/\, \mathscr{X}_0 = i) = p_{il}.$$

En definitiva, la Igualdad (A.12) anterior resulta ser:

$$f_{iC}^{(n)} = \sum_{l \notin C} \mathbb{P}(E/F_l)\mathbb{P}(F_l/\mathscr{X}_0 = i) = \sum_{l \notin C} f_{lC}^{(n-1)} \cdot p_{il}$$

tal y como queríamos ver.

□

Número de visitas de x a C en los primeros n pasos:

Sea $n \in \mathbb{N}$, $E_n = \{x \in \Omega \,/\, \mathscr{X}_n(x) \in C\}$ y $\mathbb{1}_{E_n}$ su función característica. O sea, $\mathbb{1}_{E_n} = 1$ si x visita a C en el paso n, y $\mathbb{1}_{E_n} = 0$ sino.

Definición A.10.5. *Definimos el* ***número de visitas de*** **x** ***a*** **C** ***en los primeros*** **n** ***pasos*** *como:*

$$V_C^{(n)}(x) = \sum_{k=1}^{n} \mathbb{1}_{E_k}. \tag{A.13}$$

Así pues,

$$V_C^{(n)}(x) = \sum_{j \in C} V_j^{(n)}(x).$$

También contamos con la definición para $n = +\infty$:

Definición A.10.6. *El* ***número total de visitas de*** **x** ***a*** **C** *viene dado por la siguiente ecuación:*

$$V_C(x) = \sum_{k=1}^{+\infty} \mathbb{1}_{E_k}. \tag{A.14}$$

Notemos, también aquí, que

$$V_C(x) = \sum_{j \in C} V_j(x),$$

y que $V_C(x)$ puede llegar a valer $+\infty$.

Observación A.10.2. *Gracias a estas últimas definiciones, contamos con las siguientes igualdades que relacionan los conceptos ya vistos:*

$$\{x \in \Omega \,/\, V_C^{(n)}(x) \geq 1\} = \{x \in \Omega \,/\, t_C(x) \leq n\},$$

$$\{x \in \Omega \,/\, V_C(x) \geq 1\} = \{x \in \Omega \,/\, t_C(x) \leq +\infty\}. \tag{A.15}$$

Demostración A.10.4. *Primero demostraremos la primera igualdad de la Observación A.10.2 y después la Igualdad A.15:*

- *Demostraremos la igualdad:*

$$\{x \in \Omega \,/\, V_C^{(n)}(x) \geq 1\} = \{x \in \Omega \,/\, t_C(x) \leq n\}.$$

Antes de nada, sea $A = \{x \in \Omega \,/\, V_C^{(n)}(x) \geq 1\}$ y $B = \{x \in \Omega \,/\, t_C(x) \leq n\}$ para abreviar la escritura de ahora en adelante.

Para demostrarla, veremos las dos inclusiones. Empecemos con la inclusión de izquierda a derecha (o sea: $\subseteq$):

$$x \in A \Rightarrow V_C^{(n)}(x) \geq 1 \Rightarrow \sum_{k=1}^{n} \mathbb{1}_{E_k} \geq 1,$$

ahora, si $\sum_{k=1}^{n} \mathbb{1}_{E_k} \geq 1$*, significa que* x *visita* C *al menos una vez en los primeros* n *pasos. Por tanto:*

$$\sum_{k=1}^{n} \mathbb{1}_{E_k} \geq 1 \Rightarrow t_C(x) \leq n \Rightarrow x \in B.$$

Veamos ahora la inclusión de derecha a izquierda ($\supseteq$*) y así ya tendremos la igualdad de los conjuntos:*

$$x \in B \Rightarrow t_C(x) \leq n,$$

entonces x *visita* C *por primera vez en el paso* n *o antes. Recordemos que si* $t_C(x) = i$*, entonces* $i = \inf\{n \in \mathbb{N} \,/\, \mathscr{X}_n(x) \in C\}$*. Sea* $1 \leq i \leq n$ *el paso en que* x *visita a* C *por primera vez. Entonces*

$$\mathbb{1}_{E_i} = 1 \Rightarrow \sum_{k=1}^{n} \mathbb{1}_{E_k} \geq 1 \Rightarrow V_C^{(n)}(x) \geq 1 \Rightarrow x \in A.$$

Por tanto, $A = B$*.*

- *Ahora demostraremos la Igualdad A.15. Sean* $D = \{x \in \Omega \,/\, V_C(x) \geq 1\}$ *y* $H = \{x \in \Omega \,/\, t_C(x) < +\infty\}$*.*

  Empecemos con la inclusión de izquierda a derecha:

$$x \in D \Rightarrow V_C(x) \geq 1,$$

entonces $\exists n \in \mathbb{N}$ *t.q.* x *visita* C *en el paso* n*. Por tanto*

$$V_C(x) \geq 1 \Rightarrow t_C(x) < +\infty \Rightarrow x \in H.$$

Vayamos ahora con la inclusión de derecha a izquierda:

$$x \in H \Rightarrow t_C(x) < +\infty,$$

$\exists n \in \mathbb{N}$ *t.q.* x *visita* C *por primera vez en el paso* n*. Por tanto*

$$t_C(x) < +\infty \Rightarrow \mathbb{1}_{E_n} = 1 \Rightarrow V_C(x) \geq 1 \Rightarrow x \in D.$$

Así $D = H$*.*

$\square$

Una vez contamos con estos parámetros, podemos definir la siguiente probabilidad a la que haremos alusión en el próximo teorema:

Definición A.10.7. *Sean $i, j \in S$ y $r \in \mathbb{N}$. La* ***probabilidad de visitar* j *al menos* r *veces empezando desde el estado* i** *se calcula como sigue:*

$$\mathbb{P}(V_j \geq r \,/\, \mathscr{X}_0 = i).$$

Teorema A.10.2. *Para cualquier entero positivo r y para cualquier par de estados $i, j \in S$ tenemos que:*

$$\mathbb{P}(V_j \geq r \,/\, \mathscr{X}_0 = i) = f_{ij} f_{jj}^{r-1}.$$

Y, en particular, la probabilidad de visitar r veces j desde j es:

$$\mathbb{P}(V_j \geq r \,/\, \mathscr{X}_0 = j) = f_{jj}^{r}. \tag{A.16}$$

Demostración A.10.5. *Realizaremos la demostración para $r = 1$ y para $r \geq 2$ por separado:*

- *Sea $r = 1$. Por la Igualdad (A.15) y por la Definición A.10.3 tenemos que:*

$$\mathbb{P}(V_j \geq 1 \,/\, \mathscr{X}_0 = i) = \mathbb{P}(t_j < +\infty \,/\, \mathscr{X}_0 = i) = fij.$$

  Por lo que la primera ecuación que anuncia el teorema se cumpliría.

- *Sea $r \geq 2$ y $h \in \mathbb{N}$. Definimos*

$$E_n = \{x \in \Omega \,/\, \mathscr{X}_n(x) = j\} \quad y \quad R_h(x) = \sum_{k=h+1}^{+\infty} \mathbb{1}_{E_k}(x).$$

  Vamos, primero de todo, a demostrar la siguiente igualdad:

$$\{x \in \Omega \,/\, V_j(x) \geq r\} = \bigcup_{h=1}^{+\infty} \{x \in \Omega \,/\, R_h(x) \geq r-1,\ t_j(x) = h\}, \quad \forall r \geq 2. \tag{A.17}$$

  Veremos que se cumplen las dos inclusiones. Para abreviar escritura en la demostración llamaremos a los conjuntos de la igualdad: $A = \{x \in \Omega \,/\, V_j(x) \geq r\}$ y $B = \bigcup_{h=1}^{+\infty} \{x \in \Omega \,/\, R_h(x) \geq r-1,\ t_j(x) = h\}$.

  Empecemos con la inclusión de izquierda a derecha ($\subseteq$):

  Sea $x \in A$. Por ser de A, x visita j al menos r veces, o sea que

$$V_j(x) = \sum_{k=1}^{+\infty} \mathbb{1}_{E_k}(x) \geq r.$$

  Por ser $V_j(x) \geq r \geq 2$, existe $h \in \mathbb{N}$ t.q. $t_j(x) = h$ (por la Igualdad A.15). Por tanto, el primer paso en que visitamos j sería en $\mathscr{X}_h(x) = j$. Esto quiere decir que

$$\sum_{k=1}^{+\infty} \mathbb{1}_{E_k}(x) \geq r \Rightarrow \sum_{k=h+1}^{+\infty} \mathbb{1}_{E_k}(x) + \mathbb{1}_{E_h}(x) \geq r \Rightarrow \sum_{k=h+1}^{+\infty} \mathbb{1}_{E_k}(x) + 1 \geq r \Rightarrow \sum_{k=h+1}^{+\infty} \mathbb{1}_{E_k}(x) \geq r-1 \Rightarrow$$

$$R_h(x) \geq r-1.$$

Como $t_j(x) = h$ y $R_h(x) \geq r-1$ para el $x \in A$ que habíamos elegido, entonces $x \in B$ y se cumple la inclusión.

Vamos ahora con la inclusión de derecha a izquierda ($\supseteq$):

Sea $x \in B$ arbitrario pero fijado. Por ser de B, existe un $h \in \mathbb{N}$ t.q. $t_j(x) = h$ y $R_h(x) \geq r-1$ (notemos que si $x \in B$ entonces x solo pertenece a uno de los conjuntos de la unión, no puede pertenecer a más de uno ya que la variable aleatoria t_j tiene un valor único para cada x). Por tanto:

$$t_j(x) = h \text{ y } R_h(x) \geq r-1 \Rightarrow \mathbb{1}_{E_h}(x) = 1 \text{ y } \sum_{k=h+1}^{+\infty} \mathbb{1}_{E_k}(x) \geq r-1 \Rightarrow \sum_{k=h+1}^{+\infty} \mathbb{1}_{E_k}(x) + \mathbb{1}_{E_h}(x) \geq r,$$

como $t_j(x) = h$, la única vez que x está en j entre los pasos 1 y h es en h, por lo que:

$$\sum_{k=h+1}^{+\infty} \mathbb{1}_{E_k}(x) + \mathbb{1}_{E_h}(x) \geq r \Rightarrow \sum_{k=1}^{+\infty} \mathbb{1}_{E_k}(x) \geq r \Rightarrow V_j(x) \geq r \Rightarrow x \in A.$$

Por tanto la Igualdad (A.17) se mantiene cierta.

Una vez tenemos demostrada esta ecuación, continuemos con la demostración original. Sea

$$E = \{x \in \Omega \,/\, R_h(x) \geq r-1\},$$

$$F = \{x \in \Omega \,/\, t_j(x) = h\},$$

$$G = \{x \in \Omega \,/\, \mathscr{X}_0(x) = i\}.$$

Notemos que E está identificada por las vv.aa. de $h+1$ en adelante y F está identificada por las vv.aa. de 1 a h. Exigimos que $\mathbb{P}(F \cap G) > 0$ y así cumplimos las hipótesis para aplicar la Igualdad (A.8) del Teorema A.9.1:

$$\mathbb{P}(R_h \geq r-1, t_j = h \,/\, \mathscr{X}_0 = i) = \mathbb{P}(\{E \cap F\}/G) = \mathbb{P}(E/F)\mathbb{P}(F/G) =$$

$$\mathbb{P}(R_h \geq r-1 \,/\, t_j = h)\mathbb{P}(t_j = h \,/\, \mathscr{X}_0 = i). \tag{A.18}$$

Veamos ahora qué es exactamente $\mathbb{P}(R_h \geq r-1 \,/\, t_j = h)$: como $\{\mathscr{X}_n\}$ es una C.M. en particular, y F está identificada por las vv.aa. de 1 a h tenemos que:

$$\mathbb{P}(R_h \geq r-1 \,/\, t_j = h) = \mathbb{P}(R_h \geq r-1 \,/\, \mathscr{X}_h = j)$$

por la propiedad de Markov. *Ahora, por ser $\{\mathscr{X}_n\}$ una C.M.H., podemos aplicar la Proposición A.10.1, y sabemos que E está identificada por las vv.aa. de $h+1$ en adelante, por ello:*

$$\mathbb{P}(R_h \geq r-1 \,/\, \mathscr{X}_h = j) = \mathbb{P}(V_j \geq r-1 \,/\, \mathscr{X}_0 = j).$$

Por otro lado, claramente $\mathbb{P}(t_j = h \,/\, \mathscr{X}_0 = i) = f_{ij}^{(h)}$. Sustituyendo estos valores en la Igualdad (A.18) nos queda:

$$\mathbb{P}(R_h \geq r-1 \,/\, t_j = h)\mathbb{P}(t_j = h \,/\, \mathscr{X}_0 = i) = \mathbb{P}(V_j \geq r-1 \,/\, \mathscr{X}_0 = j) f_{ij}^{(h)}. \tag{A.19}$$

Aplicaremos ahora las Igualdades (A.17) y (A.19) respectivamente para obtener:

$$\mathbb{P}(V_j \geq r/\mathscr{X}_0 = i) = \sum_{h=1}^{+\infty} \mathbb{P}(R_h \geq r-1, t_j = h/\mathscr{X}_0 = i) = \left(\sum_{h=1}^{+\infty} f_{ij}^{(h)}\right) \mathbb{P}(V_j \geq r-1/\mathscr{X}_0 = j) =$$

$$f_{ij} \cdot \mathbb{P}(V_j \geq r-1 \,/\, \mathscr{X}_0 = j). \tag{A.20}$$

Pasemos finalmente a demostrar por inducción que $\mathbb{P}(V_j \geq r/\mathscr{X}_0 = i) = f_{ij} f_{jj}^{r-1}, \forall r \geq 2$:

- *Sea $r = 2$. Por la Igualdad (A.20) que acabamos de ver, y por la primera parte de esta demostración (hemos visto que para $r = 1$ se cumple que $\mathbb{P}(V_j \geq 1 \,/\, \mathscr{X}_0 = i) = f_{ij}$) tenemos que:*

$$\mathbb{P}(V_j \geq 2 \,/\, \mathscr{X}_0 = i) = f_{ij} \cdot \mathbb{P}(V_j \geq 1 \,/\, \mathscr{X}_0 = j) = f_{ij} \cdot f_{jj} = f_{ij} f_{jj}^{r-1}.$$

- *Supongamos que se cumple para r y veamos si se cumple para $r+1$. Por la Igualdad (A.20) de nuevo:*

$$\mathbb{P}(V_j \geq r+1 \,/\, \mathscr{X}_0 = i) = f_{ij} \cdot \mathbb{P}(V_j \geq r \,/\, \mathscr{X}_0 = j),$$

  ahora, por hipótesis de inducción:

$$\mathbb{P}(V_j \geq r \,/\, \mathscr{X}_0 = j) = f_{jj} \cdot f_{jj}^{r-1} = f_{jj}^{r}.$$

  Por tanto:

$$f_{ij} \cdot \mathbb{P}(V_j \geq r \,/\, \mathscr{X}_0 = j) = f_{ij} \cdot f_{jj}^{r},$$

  tal y como queríamos ver.

$\square$

Número esperado de visitas al estado j empezando desde el estado i:

Sea $V_j(x)$ el número total de visitas de x a j. Este parámetro es una v.a. debido a que asigna un valor numérico a los resultados del experimento Ω. Podemos calcular el valor esperado de una v.a., como es V_j, de la siguiente manera:

$$\mathbb{E}[V_j] = \frac{1}{\mathbb{P}(\Omega)} \int_\Omega V_j(x)\mathbb{P}(dx),$$

Esta notación puede sorprendernos, por ello realizaremos una demostración para ver de donde sale:

Demostración A.10.6. *Vamos a ver qué es*

$$\int_\Omega \mathscr{X}(x)\mathbb{P}(dx).$$

Sea $\mathscr{X} : \Omega \longrightarrow \mathbb{R}$ una función $\mathscr{E}$-medible [10] *con valores discretos dados por $\{a_k\}_{k\in\mathbb{Z}^+}$, un conjunto numerable de valores no-negativos, y $\{E_k\}_{k\in\mathbb{Z}^+}$ un conjunto de sucesos tt.q. $E_k = \{x \in \Omega \,/\, \mathscr{X}(x) = a_k\}$ para cada $k \in \mathbb{Z}^+$. Notemos que los sucesos E_k son disjuntos dos a dos.*

Podemos definir la función de la siguiente manera:

$$\mathscr{X}(x) = \sum_{j=1}^{+\infty} a_j \cdot \mathbb{1}_{E_j}(x),$$

ya que, al ser los sucesos E_k disjuntos dos a dos, si $x \in E_p$ (con $p \in \mathbb{Z}^+$) tenemos que:

$$\mathscr{X}(x) = a_1 \cdot 0 + \cdots a_{p-1} \cdot 0 + a_p \cdot 1 + a_{p+1} \cdot 0 + a_{p+2} \cdot 0 \cdots \Rightarrow \mathscr{X}(x) = a_p,$$

que es lo que debería valer. Por tanto la definición está bien construida. Distinguiremos dos casos: cuando $\mathscr{X}$ tiene rango finito y cuando no lo tiene:

- *Supongamos que $\mathscr{X}$ tiene rango finito, o sea, que $+\infty \notin \mathscr{X}(x)$, $\forall x \in \Omega$.*

  Entonces $\mathscr{X}$ es una función simple definida de la siguiente manera:

  $$\mathscr{X}(x) = \sum_{j=1}^{k} a_j \cdot \mathbb{1}_{E_j}(x), \tag{A.21}$$

  donde $k \in \mathbb{Z}^+$ (este valor k viene determinado por el camino x y el suceso E_k al que pertenece) y $\mathbb{1}_{E_j}(x)$ es la función característica, que vale 1 cuando $\mathscr{X}(x) = a_j$ y 0 sino (para cada $j \in \{1, \ldots, k\}$). O sea, es una combinación lineal finita de funciones características.

  Ahora, por definición de las integrales con respecto a las medidas tenemos que:

  $$\int_\Omega \mathscr{X}(x)\mathbb{P}(dx) = \sum_{j=1}^{k} a_j \mathbb{P}(E_j) = \sum_{j=1}^{k} a_j \mathbb{P}(\mathscr{X}(x) = a_j). \tag{A.22}$$

[10] Véase la Definición A.4.4 del Apéndice A para más detalle.

Como $\mathscr{X}(\Omega) = \{a_k\}$ *es un conjunto numerable, uniendo las Igualdades (A.21) y (A.22), obtenemos:*

$$\int_\Omega \sum_{j=1}^{k} a_j \cdot \mathbb{1}_{E_j}(x)\mathbb{P}(dx) = \sum_{j=1}^{k} a_j \mathbb{P}(\mathscr{X}(x) = a_j), \tag{A.23}$$

para cualquier $k \in \mathbb{N}$.

Sea $\{f_k\}$ *una sucesión de funciones tt.q.*

$$f_k(x) = \sum_{j=1}^{k} a_j \mathbb{1}_{E_j}(x),$$

para cada $k \in \mathbb{Z}^+$ *. Notemos 2 cosas:*

- *Por la Igualdad (A.21):*

$$\mathscr{X}(x) = f_k(x), \forall k \in \mathbb{Z}^+.$$

  Por ello los $f_k(x)$ *son funciones no-negativas* $\mathscr{E}$*-medibles.*

- *La sucesión* $\{f_k\}$ *es no-decreciente. O sea:*

$$f_k(x) \leq f_{k+1}(x), \forall k \in \mathbb{Z}^+$$

  para $x \in \Omega$. *Debido a que, para cada* k*:*

  - *Si* $x \in E_1 \cup \cdots \cup E_k$, *entonces* $f_k(x)$ *es igual a* $f_{k+1}(x)$.
  - *Si* $x \in E_{k+1}$, *entonces* $f_k(x) = 0$, *y si* $a_{k+1} > 0$, *tenemos que* $f_k(x) < a_{k+1} = f_{k+1}(x)$.
  - *Si* $x \in E_{k+2} \cup E_{k+3} \cup \cdots$, *entonces* $f_k(x) = 0 = f_{k+1}(x)$.

Fijémonos que también se cumple lo siguiente:

$$\lim_{k\to+\infty} \sum_{j=1}^{k} a_j \mathbb{1}_{E_j}(x) = \sum_{j=1}^{+\infty} a_j \mathbb{1}_{E_j}(x) = \mathscr{X}(x).$$

Por tanto, cumplimos las hipótesis del teorema de Beppo Levi [11] *para poder aplicarlo, siendo* $f(x) = \mathscr{X}(x)$, *y, por lo que acabamos de ver:*

$$\int_\Omega \mathscr{X}(x)\mathbb{P}(dx) = \int_\Omega \lim_{k\to+\infty} \sum_{j=1}^{k} a_j \mathbb{1}_{E_j}(x)\mathbb{P}(dx) = \lim_{k\to+\infty} \int_\Omega \sum_{j=1}^{k} a_j \mathbb{1}_{E_j}(x)\mathbb{P}(dx).$$

Ahora, por la Igualdad (A.23) tenemos que:

[11] Véase el Teorema A.4.1 del Apéndice A para consultar el enunciado.

$$\lim_{k\to+\infty}\int_{\Omega}\sum_{j=1}^{k} a_j \mathbb{1}_{E_j}(x)\mathbb{P}(dx) = \lim_{k\to+\infty}\sum_{j=1}^{k} a_j\mathbb{P}(\mathscr{X}(x)=a_j) = \sum_{j=1}^{+\infty} a_j\mathbb{P}(\mathscr{X}(x)=a_j),$$

y así:

$$\int_{\Omega}\mathscr{X}(x)\mathbb{P}(dx) = \sum_{j=1}^{+\infty} a_j\mathbb{P}(\mathscr{X}(x)=a_j). \tag{A.24}$$

- *Supongamos ahora que el rango de $\mathscr{X}$ no tiene porque ser finito. O sea, $+\infty\in\mathscr{X}(\Omega)$. Sea $\{a_k\}\cup\{+\infty\}$ el rango de $\mathscr{X}$ (con $\{a_k\}$ definido como antes). En este caso nos conviene definir la integral que queremos ver lo que vale como:*

$$\int_{\Omega}\mathscr{X}(x)\mathbb{P}(dx) = \int_{\{\mathscr{X}<+\infty\}}\mathscr{X}(x)\mathbb{P}(dx) + \int_{\{\mathscr{X}=+\infty\}}\mathscr{X}(x)\mathbb{P}(dx),$$

  y calcular cada sumando por separado:

  - *El primer sumando lo conocemos por la Igualdad (A.24) ya que es el caso en que el rango de $\mathscr{X}$ es finito: $\{\mathscr{X}<+\infty\}$.*
  - *El segundo sumando debemos calcularlo:*

$$\int_{\{\mathscr{X}=+\infty\}}\mathscr{X}(x)\mathbb{P}(dx) = \lim_{k\to+\infty} k\cdot\mathbb{P}(\mathscr{X}=+\infty) = \begin{cases} 0 & si\ \mathbb{P}(\mathscr{X}=+\infty)=0 \\ +\infty & si\ \mathbb{P}(\mathscr{X}=+\infty)>0 \end{cases}$$

  En definitiva:

$$\int_{\Omega}\mathscr{X}(x)\mathbb{P}(dx) = \begin{cases} 0+\sum_{j=1}^{+\infty} a_j\mathbb{P}(\mathscr{X}(x)=a_j) & si\ \mathbb{P}(\mathscr{X}=+\infty)=0 \\ \sum_{j=1}^{+\infty} a_j\mathbb{P}(\mathscr{X}(x)=a_j)+\infty & si\ \mathbb{P}(\mathscr{X}=+\infty)>0 \end{cases} =$$

$$\begin{cases} \sum_{j=1}^{+\infty} a_j\mathbb{P}(\mathscr{X}(x)=a_j) & si\ \mathbb{P}(\mathscr{X}=+\infty)=0 \\ +\infty & si\ \mathbb{P}(\mathscr{X}=+\infty)>0 \end{cases}$$

$\square$

Definimos $\mathbb{P}_i[A]$ como

$$\mathbb{P}_i[A] = \mathbb{P}(A\,/\,\mathscr{X}_0=i),$$

para $i\in S$ y $A\in\mathscr{E}$.

Lo que estamos calculando aquí es el número esperado de visitas que haremos a j en particular (aunque podría ser a un conjunto de estados $C\subset S$ en general), de entre todos los caminos x que empiezan en el estado i.

A continuación expondremos un teorema que nos proporcionará una serie de igualdades que se mantienen ciertas y que nos serán de gran ayuda en las Subsecciones siguientes:

Teorema A.10.3. *Las siguientes igualdades se mantienen ciertas:*

$$\mathbb{E}_i[V_j] = \begin{cases} +\infty & si\ \mathbb{P}(V_j = +\infty \,/\, \mathscr{X}_0 = i) > 0 \\ \sum_{k=0}^{+\infty} k \cdot \mathbb{P}(V_j = k \,/\, \mathscr{X}_0 = i) & si\ \mathbb{P}(V_j = +\infty \,/\, \mathscr{X}_0 = i) = 0 \end{cases} \tag{A.25}$$

$$\mathbb{E}_i[V_j] = \sum_{n=0}^{+\infty} p_{ij}^{(n)} \tag{A.26}$$

$$\mathbb{E}_i[V_j] = \begin{cases} +\infty & si\ f_{jj} = 1 \\ \frac{f_{ij}}{1-f_{jj}} & si\ f_{jj} < 1 \end{cases} \tag{A.27}$$

En particular, juntando las dos ecuaciones anteriores:

$$\sum_{n=0}^{+\infty} p_{ij}^{(n)} = \begin{cases} +\infty & si\ f_{jj} = 1 \\ \frac{f_{ij}}{1-f_{jj}} & si\ f_{jj} < 1 \end{cases} \tag{A.28}$$

Demostración A.10.7. *Veamos cuál es el valor de:*

$$\mathbb{E}_i[V_j] = \frac{1}{\mathbb{P}(\mathscr{X}_0 = i)} \int_{\{\mathscr{X}_0 = i\}} V_j(x)\mathbb{P}(dx).$$

Recordemos que $V_j(x) = \sum_{k=1}^{+\infty} \mathbb{1}_{E_k}(x)$*, con* $E_k = \{x \in \Omega \,/\, \mathscr{X}_k(x) = j\}$*, es una v.a. t.q.* $\mathscr{X} : \Omega \longrightarrow S$ *es una función* $\mathscr{E}$*-medible no negativa y con rango* S*. Por tanto, podemos aplicar lo explicado en la Demostración A.10.6 anterior y así:*

$$\int_{\{\mathscr{X}_0 = i\}} V_j(x)\mathbb{P}(dx) = \begin{cases} \sum_{k=1}^{+\infty} k \cdot \mathbb{P}(V_j(x) = k, \mathscr{X}_0 = i) & si\ \mathbb{P}(V_j = +\infty, \mathscr{X}_0 = i) = 0 \\ +\infty & si\ \mathbb{P}(V_j = +\infty, \mathscr{X}_0 = i) > 0 \end{cases}$$

Multiplicado en ambas partes de la igualdad por $\frac{1}{\mathbb{P}(\mathscr{X}_0 = i)}$ *obtenemos:*

$$\mathbb{E}_i[V_j] = \frac{1}{\mathbb{P}(\mathscr{X}_0 = i)} \int_{\{\mathscr{X}_0 = i\}} V_j(x)\mathbb{P}(dx) = \begin{cases} \sum_{k=1}^{+\infty} \frac{k \cdot \mathbb{P}(V_j(x) = k, \mathscr{X}_0 = i)}{\mathbb{P}(\mathscr{X}_0 = i)} & si\ \frac{\mathbb{P}(V_j = +\infty, \mathscr{X}_0 = i)}{\mathbb{P}(\mathscr{X}_0 = i)} = 0 \\ +\infty & si\ \frac{\mathbb{P}(V_j = +\infty, \mathscr{X}_0 = i)}{\mathbb{P}(\mathscr{X}_0 = i)} > 0 \end{cases} =$$

$$\begin{cases} \sum_{k=1}^{+\infty} k \cdot \mathbb{P}(V_j(x) = k \,/\, \mathscr{X}_0 = i) & si\ \mathbb{P}(V_j = +\infty \,/\, \mathscr{X}_0 = i) = 0 \\ +\infty & si\ \mathbb{P}(V_j = +\infty \,/\, \mathscr{X}_0 = i) > 0 \end{cases}$$

Y así resulta la Igualdad (A.25). Vamos ahora a demostrar la Igualdad (A.26):

Por definición de V_j *y propiedades de la esperanza matemática* [12] *tenemos que:*

$$\mathbb{E}_i[V_j] = \mathbb{E}_i[\sum_{n=0}^{+\infty} \mathbb{1}_{E_n}(x)] = \sum_{n=0}^{+\infty} \mathbb{E}_i[\mathbb{1}_{E_n}(x)].$$

[12] Véase la Propiedad A.3.1 del Apéndice A para más información sobre las propiedades de la esperanza.

Ahora, $\mathbb{1}_{E_n}(x)$ es la función característica del suceso $E_n \in \mathscr{E}$. Esta función sería nuestra v.a. discreta. Si aplicamos la fórmula expuesta al principio de esta sección a E_n, junto con la definición de esperanza de una v.a. discreta [13]*, obtenemos:*

$$\mathbb{E}_i[\mathbb{1}_{E_n}(x)] = 1 \cdot \mathbb{P}(\mathscr{X}_n = j \,/\, \mathscr{X}_0 = i) + 0 \cdot \mathbb{P}(\mathscr{X}_n \neq j \,/\, \mathscr{X}_0 = i) = \mathbb{P}(\mathscr{X}_n = j \,/\, \mathscr{X}_0 = i).$$

Así pues:

$$\sum_{n=0}^{+\infty} \mathbb{E}_i[\mathbb{1}_{E_n}(x)] = \sum_{n=0}^{+\infty} \mathbb{P}(\mathscr{X}_n = j \,/\, \mathscr{X}_0 = i) = \sum_{n=0}^{+\infty} p_{ij}^{(n)}$$

y por tanto:

$$\mathbb{E}_i[V_j] = \sum_{n=0}^{+\infty} p_{ij}^{(n)},$$

tal y como queríamos ver. Por último, veamos la Igualdad (A.27):

Sabemos que $V_j(x)$ es una función no negativa $\mathscr{E}$-medible y que toma valores en $S \subset \overline{\mathbb{R}}_+$. Por ello, podemos aplicar la fórmula de Cavalieri [14] *a $V_j(x)$ (que es una v.a. discreta) y sobre $\{\mathscr{X}_0 = i\}$, en lugar de Ω, y nos queda:*

$$\mathbb{E}_i[V_j] = \frac{1}{\mathbb{P}(\mathscr{X}_0 = i)} \int_{\{\mathscr{X}_0 = i\}} V_j(x)\mathbb{P}(dx) = \frac{1}{\mathbb{P}(\mathscr{X}_0 = i)} \sum_{k=0}^{+\infty} \mathbb{P}(V_j > k, \mathscr{X}_0 = i) = \sum_{k=0}^{+\infty} \mathbb{P}(V_j > k / \mathscr{X}_0 = i) =$$

$$\sum_{k=1}^{+\infty} \mathbb{P}(V_j \geq k \,/\, \mathscr{X}_0 = i),$$

por el Teorema A.10.2:

$$\sum_{k=1}^{+\infty} \mathbb{P}(V_j \geq k \,/\, \mathscr{X}_0 = i) = \sum_{k=1}^{+\infty} f_{ij} \cdot f_{jj}^{k-1} = f_{ij} \sum_{k=1}^{+\infty} f_{jj}^{k-1} = \begin{cases} +\infty & si\ f_{jj} = 1 \\ 0 & si\ f_{jj} = 0 \\ f_{ij} \displaystyle\sum_{k=1}^{+\infty} f_{jj}^{k-1} & si\ 0 < f_{jj} < 1 \end{cases}$$

Ahora, en el caso en que $0 < f_{jj} < 1$ tenemos que $\sum_{k=1}^{+\infty} f_{jj}^{k-1}$ es una progresión geométrica de razón f_{jj}. Por ello

$$f_{ij} \sum_{k=1}^{+\infty} f_{jj}^{k-1} = f_{ij} \frac{1}{1 - f_{jj}} = \frac{f_{ij}}{1 - f_{jj}}.$$

En definitiva, aplicando esto a la igualdad anterior nos queda:

$$\mathbb{E}_i[V_j] = \begin{cases} +\infty & si\ f_{jj} = 1 \\ 0 & si\ f_{jj} = 0 \\ \frac{f_{ij}}{1-f_{jj}} & si\ 0 < f_{jj} < 1 \end{cases} = \begin{cases} +\infty & si\ f_{jj} = 1 \\ \frac{f_{ij}}{1-f_{jj}} & si\ 0 \leq f_{jj} < 1 \end{cases}$$

tal y como queríamos demostrar.

□

[13] Véase la Definición A.3.10 del Apéndice A para más información sobre la esperanza de una v.a..

[14] Más información sobre esta fórmula en el Teorema A.4.2 del Apéndice A.

Estados recurrentes y transitorios:

En esta sección hablaremos sobre unas proposiciones estrechamente relacionados con los estados recurrentes y transitorios definidos en el nudo de este trabajo.

Proposición A.10.3. *Son equivalentes:*

i) *j es un estado recurrente.*

ii) *Los caminos desde j visitan j casi seguro, o sea: $f_{jj} = 1$.*

iii) *Los caminos desde j visitan j un número infinito de veces casi seguro.*

Demostración A.10.8. *Por la Igualdad (A.27), tenemos que i) y ii) son equivalentes. Además, iii) implica obviamente ii). Veamos si ii) implica iii):*

Si ii) se cumple, tenemos que $f_{jj} = 1$. Queremos comprobar si $\mathbb{P}(V_j = +\infty / \mathscr{X}_0 = j) = 1$ para ver que se cumple iii). Veamos:

$$\mathbb{P}(V_j = +\infty \,/\, \mathscr{X}_0 = j) = \lim_{r\to+\infty} \mathbb{P}(V_j \geq r \,/\, \mathscr{X}_0 = j),$$

por el Teorema A.10.2:

$$\lim_{r\to+\infty} \mathbb{P}(V_j \geq r \,/\, \mathscr{X}_0 = j) = \lim_{r\to+\infty} f_{jj}^r = \lim_{r\to+\infty} 1 = 1.$$

□

Nota A.10.1. *Si se cumple esta proposición, entonces los caminos que empiezan en i visitan j un número infinito de veces con probabilidad f_{ij} ya que:*

$$\mathbb{P}(V_j = +\infty \,/\, \mathscr{X}_0 = i) = \lim_{r\to+\infty} \mathbb{P}(V_j \geq r \,/\, \mathscr{X}_0 = i) = \lim_{r\to+\infty} f_{ij} f_{jj}^{r-1} = \lim_{r\to+\infty} f_{ij} = f_{ij},$$

debido a que $f_{jj} = 1$.

Esta probabilidad, además, es la misma que la de visitar j al menos una vez (desde i).

Veamos la proposición equivalente para los estados transitorios:

Proposición A.10.4. *Son equivalentes:*

i) *j es un estado transitorio.*

ii) *La probabilidad de que los caminos visiten j desde j es menor que 1, o sea: $f_{jj} < 1$.*

iii) *La probabilidad de que los caminos desde j visiten j un número infinito de veces es 0.*

iv) *Los caminos desde j visitan j un número finito de veces casi seguro.*

Demostración A.10.9. *De nuevo, por la Igualdad (A.27), tenemos que i) y ii) son equivalentes.*

Por otro lado, de manera semejante a como hemos realizado la demostración de la propiedad anterior, tenemos que:

$$\mathbb{P}(V_j = +\infty \,/\, \mathcal{X}_0 = j) = \lim_{r\to+\infty} \mathbb{P}(V_j \geq r \,/\, \mathcal{X}_0 = j) = \lim_{r\to+\infty} f_{jj}^r.$$

Así que iii) se mantendrá cierto sii $f_{jj} < 1$, o sea, sii ii) es cierto.

Por último, iii) y iv) son equivalentes ya que uno es el contra-recíproco del otro.

Nota A.10.2. *Notemos que si esta proposición se cumple, entonces los caminos desde i visitan j un número infinito de veces con probabilidad 0 ya que:*

$$\mathbb{P}(V_j = +\infty \,/\, \mathcal{X}_0 = i) = \lim_{r\to+\infty} \mathbb{P}(V_j \geq r \,/\, \mathcal{X}_0 = i) = \lim_{r\to+\infty} f_{ij} f_{jj}^{r-1} = 0,$$

debido a que, en este caso, $f_{jj} < 1$.

La siguiente proposición nos da una serie de resultados interesantes sobre los dos tipos de estados y qué ocurre si otros estados están comunicados con ellos:

Proposición A.10.5. *Lo siguiente se mantiene cierto:*

i) Si $i \longrightarrow j$ y j es recurrente, entonces

$$\sum_{k=1}^{+\infty} p_{ij}^{(k)} = +\infty.$$

O sea, si $i \longrightarrow j$ y $f_{jj} = 1$, entonces $f_{ij} = 1$.

ii) Si $i \longleftrightarrow j$, entonces o i y j son transitorios, o i y j son recurrentes.

O sea, si $i \longleftrightarrow j$, entonces $f_{ii} = 1 \Leftrightarrow f_{jj} = 1$.

Demostración A.10.10. *Veamos las demostraciones de cada apartado por orden:*

i) Como $i \longrightarrow j$ tenemos que existe $k \in \mathbb{N}$ t.q. $p_{ij}^{(k)} > 0$.

Por otro lado, contamos con la siguiente desigualdad:

$$p_{ij}^{(k+n)} = \sum_{\alpha\in S} p_{i\alpha}^{(k)} p_{\alpha j}^{(n)} \geq p_{ij}^{(k)} p_{jj}^{(n)},$$

para cada $n \in \mathbb{N} \cup \{0\}$. Para demostrarla basta que cojamos $\alpha = j$ en particular. Por lo que

$$p_{ij}^{(k)} p_{jj}^{(n)} \leq p_{ij}^{(k+n)},$$

con k fijado y para cada $n \in \mathbb{N} \cup \{0\}$.

Por tanto:

$$\sum_{n=1}^{+\infty} p_{ij}^{(k)} p_{jj}^{(n)} \leq \sum_{n=1}^{+\infty} p_{ij}^{(k+n)}.$$

Ahora, por ser j recurrente, sabemos que $\sum_{n=1}^{+\infty} p_{jj}^{(n)} = +\infty$. *Aplicando esto a la desigualdad anterior nos queda que:*

$$+\infty = \sum_{n=1}^{+\infty} p_{ij}^{(k)} p_{jj}^{(n)} \leq \sum_{n=1}^{+\infty} p_{ij}^{(k+n)} \Rightarrow \sum_{n=1}^{+\infty} p_{ij}^{(k+n)} = +\infty.$$

Y, como $\{\mathscr{X}_n\}$ es una C.M.H., tenemos que $p_{ij}^{(k+n)} = p_{ij}^{(n)}$. Así pues,

$$\sum_{n=1}^{+\infty} p_{ij}^{(k+n)} = \sum_{n=1}^{+\infty} p_{ij}^{(n)} = +\infty,$$

tal y como queríamos ver.

ii) Si $i \longleftrightarrow j$, entonces existen h y k enteros positivos tt.q.

$$p_{ij}^{(h)} > 0 \text{ y } p_{ji}^{(k)} > 0.$$

De manera análoga a lo que hemos hecho antes, la siguiente desigualdad se mantiene cierta:

$$p_{ii}^{(h+n+k)} = \sum_{\beta,\alpha \in S} p_{i\alpha}^{(h)} p_{\alpha\beta}^{(n)} p_{\beta i}^{(k)} \geq p_{ij}^{(h)} p_{jj}^{(n)} p_{ji}^{(k)},$$

con k, h fijados y para $n \in \mathbb{N} \cup \{0\}$. En este caso, basta coger $\alpha = \beta = j$ en particular. Por lo que:

$$\sum_{n=1}^{+\infty} p_{ij}^{(h)} p_{jj}^{(n)} p_{ji}^{(k)} \leq \sum_{n=1}^{+\infty} p_{ii}^{(h+n+k)},$$

y por ser $\{\mathscr{X}_n\}$ una C.M.H., tenemos que $p_{ii}^{(h+n+k)} = p_{ii}^{(n)}$, por ello:

$$\sum_{n=1}^{+\infty} p_{ij}^{(h)} p_{jj}^{(n)} p_{ji}^{(k)} \leq \sum_{n=1}^{+\infty} p_{ii}^{(n)}. \tag{A.29}$$

De manera similar podríamos obtener la siguiente desigualdad:

$$\sum_{n=1}^{+\infty} p_{ji}^{(k)} p_{jj}^{(n)} p_{ij}^{(h)} \leq \sum_{n=1}^{+\infty} p_{jj}^{(n)} \tag{A.30}$$

para los mismos k, h enteros positivos y $n \in \mathbb{N} \cup \{0\}$.

Finalmente, veamos lo que queríamos demostrar:

- *Si j es recurrente entonces $\sum_{n=1}^{+\infty} p_{jj}^{(n)} = +\infty$, y si aplicamos la Desigualdad (A.29) obtenemos que $\sum_{n=1}^{+\infty} p_{ii}^{(n)} = +\infty$ y por tanto i también sería recurrente.*

  Si, por lo contrario, fuera i recurrente, con un argumento simétrico al que acabamos de hacer, aplicando la Desigualdad (A.30) en lugar de la Desigualdad (A.29), obtendríamos que j también sería recurrente.

- *Si j es transitorio entonces $\sum_{n=1}^{+\infty} p_{jj}^{(n)} < +\infty$, y como $p_{ij}^{(h)} > 0, p_{ji}^{(k)} > 0$, por la Desigualdad (A.30) obtendríamos que $\sum_{n=1}^{+\infty} p_{ii}^{(n)} < +\infty$, y así i sería transitorio.*

  De manera simétrica, usando la Desigualdad (A.29) en lugar de la Desigualdad (A.30), tendríamos que si i es transitorio, entonces j también lo es

En definitiva, si se cumple la hipótesis de este apartado de la proposición, tenemos que si j es recurrente entonces i también, y si j es transitorio i también lo es, tal y como queríamos demostrar.

□

A.11 Ejemplo de una representación canónica

Sea

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0,5 & 0 & 0,5 & 0 \\ 0 & 0,3 & 0 & 0,7 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,5 & 0 & 0 & 0,5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0,1 & 0,1 & 0,1 & 0,1 & 0 & 0 & 0 & 0 & 0,6 \\ 0,1 & 0 & 0 & 0 & 0,9 & 0 & 0 & 0 & 0 & 0 \\ 0,2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,8 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0,6 & 0 & 0,4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

una matriz de transición de una C.M.H. con $S = \{1,2,3,4,5,6,7,8,9,10\}$ como conjunto de estados.

Primero debemos identificar las clases cerradas minimales. Para hacerlo, calcularemos primero la matriz $\mathbf{B}$, o sea:

$$\mathbf{B} = \sum_{n=1}^{10} (\mathbf{P}^n).$$

Para hacer este cálculo hemos utilizado el programa *Octave* que es idóneo para realizar operaciones con matrices. El resultado ha sido el siguiente:

$$\mathbf{B} = \begin{pmatrix} 0,68 & 0 & 0 & 0 & 0 & 0 & 3,79 & 0 & 5,53 & 0 \\ 0 & 0,43 & 0 & 9,57 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0,17 & 1,18 & 1,48 & 1,18 & 0 & 0 & 1,06 & 0 & 4,93 \\ 0 & 0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0,18 & 0,24 & 2,17 & 0,24 & 0 & 0 & 0,12 & 0 & 7,05 \\ 0,16 & 0,16 & 0,2 & 1,74 & 1,11 & 0 & 0,34 & 0,1 & 0,5 & 5,68 \\ 0,82 & 0 & 0 & 0 & 0 & 0 & 3,47 & 0 & 5,71 & 0 \\ 0 & 0,16 & 2,13 & 1,25 & 1,18 & 0 & 0 & 1,06 & 0 & 4,22 \\ 0,7 & 0 & 0 & 0 & 0 & 0 & 3,85 & 0 & 5,45 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 \end{pmatrix}$$

Ahora aplicaremos la Proposición 2.4.3 para cada estado $i \in S$ e identificaremos las diferentes clases cerradas minimales:

- Sea $i = 1$. Tenemos que $\mathbf{B}_7^i > 0$ y $\mathbf{B}_9^i > 0$. Ahora debemos comprobar si $\mathbf{B}_i^7$ y $\mathbf{B}_i^9$ son también mayores que 0. Como $\mathbf{B}_i^7 = 0,82 > 0$ y $\mathbf{B}_i^9 = 0,7 > 0$, podemos concluir por la proposición que $i = 1 \in C_1$, donde C_1 es una clase cerrada minimal.

- Sea $i = 7$. Ya sabemos que está conjuntamente comunicado con el estado 1, veamos si lo está con el estado 9. Tenemos que $\mathbf{B}_9^i > 0$ y, además, $\mathbf{B}_i^9 > 0$. Esto quiere decir que $i = 7$ también está conjuntamente comunicado con 9, y por lo tanto concluimos que la clase cerrada minimal C_1 está formada por los estados $1,7$ y 9.

- Sea $i = 2$. Tenemos que $\mathbf{B}_4^i > 0$, pero $\mathbf{B}_i^4 = 0$. Esto quiere decir que el estado $i = 2$ no pertenece a ninguna clase cerrada minimal.

- Sea $i = 3$. El valor $\mathbf{B}_2^i > 0$, pero $\mathbf{B}_i^2 = 0$. Esto quiere decir que el estado $i = 3$ no pertenece a ninguna clase cerrada minimal tampoco.

- Sea $i = 4$. El único valor diferente de 0 de $\mathbf{B}^i$ (fila i-ésima de la matriz $\mathbf{B}$) es el $\mathbf{B}_i^i$. Esto significa que el estado $i = 4$ pertenece a una clase cerrada minimal de la que es él el único elemento del conjunto: $C_2 = \{4\}$.

- Sea $i = 5$. El valor $\mathbf{B}_2^i > 0$, pero $\mathbf{B}_i^2 = 0$. Esto quiere decir que el estado $i = 5$ no pertenece a ninguna clase cerrada minimal.

- Sea $i = 6$. El valor $\mathbf{B}_1^i > 0$, pero $\mathbf{B}_i^1 = 0$. Esto quiere decir que el estado $i = 6$ tampoco pertenece a ninguna clase cerrada minimal.

- Sea $i = 8$. El valor $\mathbf{B}_2^i > 0$, pero $\mathbf{B}_i^2 = 0$. Esto quiere decir que el estado $i = 8$ tampoco pertenece a ninguna clase cerrada minimal.

- Sea $i = 10$. El único valor diferente de 0 de $\mathbf{B}^i$ es el propio $\mathbf{B}_i^i$. Esto significa, al igual que ha ocurrido con el estado 4, que el estado $i = 10$ pertenece a una clase cerrada minimal de la que es él el único elemento del conjunto: $C_3 = \{10\}$.

Por tanto tenemos 3 clases cerradas minimales: $C_1 = \{1,7,9\}$, $C_2 = \{4\}$ y $C_3 = \{10\}$ de cardinales $k_1 = 3$, $k_2 = 1$ y $k_3 = 1$ respectivamente. Sea ahora $C = C_1 \cup C_2 \cup C_3$ y $T = S \setminus C$, por la Proposición 2.4.4 tenemos que los estados $1,4,7,9$ y 10 son recurrentes y los estados $2,3,5,6$ y 8 transitorios.

Para obtener la forma canónica lo que hacemos es re-etiquetar algunos estados en $\mathbf{P}$, de forma que los primeros sean los estados recurrentes, agrupándolos con sus respectivas clases cerradas minimales, y los últimos sean los transitorios. Los intercambios de filas y columnas que haremos serán:[15]

$$\mathbf{P}_2 \leftrightarrow \mathbf{P}_7 \text{ y } \mathbf{P}^2 \leftrightarrow \mathbf{P}^7, \quad \mathbf{P}_3 \leftrightarrow \mathbf{P}_9 \text{ y } \mathbf{P}^3 \leftrightarrow \mathbf{P}^9, \quad \mathbf{P}_5 \leftrightarrow \mathbf{P}_{10} \text{ y } \mathbf{P}^5 \leftrightarrow \mathbf{P}^{10}.$$

Es decir, hemos reetiquetado 3 estados. Los estados que tenían la etiqueta $2,3$ y 5 en la matriz original $\mathbf{P}$, ahora tendrán la etiqueta $7,9$ y 10 respectivamente en la representación canónica de $\mathbf{P}$. De esta forma, la matriz $\mathbf{P}$ en su forma canónica sería:

$$\mathbf{P} = \begin{pmatrix} 0 & 0,5 & 0,5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0,2 & 0 & 0,8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0,6 & 0,4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0,1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,9 \\ 0 & 0 & 0 & 0,7 & 0 & 0 & 0,3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,5 & 0 & 0,5 \\ 0 & 0 & 0 & 0,1 & 0,6 & 0 & 0,1 & 0 & 0,1 & 0,1 \end{pmatrix}.$$

Notemos que la estructura de la matriz es semejante a la que exponemos en el Teorema 2.4.1, con:

$$(\mathbf{B}_1) = \begin{pmatrix} 0 & 0,5 & 0,5 \\ 0,2 & 0 & 0,8 \\ 0 & 0,6 & 0,4 \end{pmatrix}, \quad (\mathbf{B}_2) = 1, \text{ y } (\mathbf{B}_3) = 1,$$

las respectivas matrices irreducibles.

A.12 Más sobre la convergencia de las potencias de P

En esta sección mostraremos más lemas y proposiciones que deberían estar en el Capítulo 3, pero, por espacio, las hemos trasladado aquí.

Proposición A.12.1. *Sean: $p_1, \dots, p_N > 0$ con $\sum_{i=1}^{N} p_i = 1$, $z_1, \dots, z_n \in \mathbb{C}$ con $|z_1|, \dots, |z_N| \leq r$ y $|\sum_{i=1}^{N} p_i z_i| = r$. Entonces $z_1 = \cdots = z_N$.*

Demostración A.12.1. *Supongamos, sin pérdida de generalidad, que $1 = r = \sum_{i=1}^{N} p_i z_i$. Ahora, si se cumpliera esta implicación que pretendemos demostrar, tendríamos que $z_1 = \cdots = z_N$ y, por tanto, $|z_i| = 1$, $\forall i \in \{1, \dots, N\}$. Ya que, sea $i_1 \in \{1, \dots, N\}$, arbitrario pero fijado, se cumple lo siguiente:*

[15] Recordemos que denotamos por $\mathbf{P}^i$ a la fila i-ésima de la matriz $\mathbf{P}$, y por $\mathbf{P}_j$ a la columna j-ésima.

$$|\sum_{i=1}^{N} p_i z_i| = |z_{i_1} \sum_{i=1}^{N} p_i| = |z_{i_1}| = 1.$$

Realizaremos la demostración por reducción al absurdo. Supondremos que $|z_{i_1}| \neq 1$ *y llegaremos a contradicción.*

Por ser $|z_{i_1}| \neq 1$ *y* $|z_{i_1}| \leq r = 1$ *(por hipótesis), tenemos que el módulo de* z_{i_1} *es como mucho* $1-\epsilon$*, con* $\epsilon > 0$*. Ahora,*

$$|\sum_{i=1}^{N} p_i z_i| = |p_{i_1} z_{i_1} + \sum_{\substack{i=1\\ i\neq i_1}}^{N} p_i z_i|,$$

por la propiedad b) de la Definición A.6.2, tenemos que:

$$|p_{i_1} z_{i_1} + \sum_{\substack{i=1\\ i\neq i_1}}^{N} p_i z_i| \leq |p_{i_1} z_{i_1}| + |\sum_{\substack{i=1\\ i\neq i_1}}^{N} p_i z_i|,$$

sabemos que $p_{i_1} > 0$ *por hipótesis, así pues:*

$$|p_{i_1} z_{i_1}| + |\sum_{\substack{i=1\\ i\neq i_1}}^{N} p_i z_i| = p_{i_1}|z_{i_1}| + |\sum_{\substack{i=1\\ i\neq i_1}}^{N} p_i z_i|,$$

como $|\sum_{i=1}^{N} p_i z_i| = r = 1$*, también por hipótesis, nos queda:*

$$p_{i_1}|z_{i_1}| + |\sum_{\substack{i=1\\ i\neq i_1}}^{N} p_i z_i| \leq p_{i_1}|z_{i_1}| + 1 - p_{i_1}|z_{i_1}|,$$

y, por ser $|z_{i_1}| \leq 1-\epsilon < 1$*, finalmente nos resulta:*

$$p_{i_1}|z_{i_1}| + 1 - p_{i_1}|z_{i_1}| < p_{i_1}(1-\epsilon) + 1 - p_{i_1} = 1 - p_{i_1}\epsilon < 1.$$

Esto quiere decir que $|\sum_{i=1}^{N} p_i z_i| < 1 = r$*, que es una contradicción.*

En definitiva, la implicación se cumple.

□

A.13 Ejemplo de cálculo de un punto fijo de P

Sea

$$\mathbf{P} = \begin{pmatrix} 0{,}5 & 0{,}45 & 0{,}05 \\ 0{,}25 & 0{,}5 & 0{,}25 \\ 0{,}3 & 0{,}3 & 0{,}4 \end{pmatrix}$$

la matriz de transición de una C.M.H. con espacio de estados finito $S = \{1,2,3\}$. Queremos encontrar el punto fijo atractivo para la aplicación $\mathrm{P}(\mathbf{x}) = \mathbf{xP}$ con $\mathbf{x} \in \mathscr{T}$. Sabemos que existe por el Teorema 3.2.1, debido a que $\mathbf{P}$ es obviamente regular. Con el punto fijo atractivo averiguaremos las distribuciones a largo plazo de las vv.aa.. Sea $\mathbf{w} = (w_1, w_2, w_3)$ el punto fijo, vamos a resolver el siguiente sistema para encontrar el valor de las componentes de $\mathbf{w}$:

$$(w_1, w_2, w_3)\begin{pmatrix} 0{,}5 & 0{,}45 & 0{,}05 \\ 0{,}25 & 0{,}5 & 0{,}25 \\ 0{,}3 & 0{,}3 & 0{,}4 \end{pmatrix} = (w_1, w_2, w_3) \Rightarrow$$

$$\begin{cases} 0{,}5w_1 + 0{,}25w_2 + 0{,}3w_3 = w_1 \\ 0{,}45w_1 + 0{,}5w_2 + 0{,}3w_3 = w_2 \\ 0{,}05w_1 + 0{,}25w_2 + 0{,}4w_3 = w_3 \end{cases} \Rightarrow$$

$$\begin{cases} -0{,}5w_1 + 0{,}25w_2 + 0{,}3w_3 = 0 \\ 0{,}45w_1 - 0{,}5w_2 + 0{,}3w_3 = 0 \\ 0{,}05w_1 + 0{,}25w_2 - 0{,}6w_3 = 0 \end{cases} \Rightarrow$$

Si resolvemos por el método de Gauss (ver Definición A.1.7) este sistema de ecuaciones, obtenemos lo siguiente:

$$w_1 = \frac{90}{55}w_3, \quad w_2 = \frac{114}{55}w_3.$$

Ahora, sabemos que

$$\mathbf{w} \in \mathcal{T} \Rightarrow \|\mathbf{w}\|_1 = 1,\ w_i \geq 0,\ \forall i \in S \Rightarrow w_1 + w_2 + w_3 = 1.$$

Añadiendo esta condición a las igualdades anteriores, el resultado final es:

$$w_1 = \frac{90}{259}, \quad w_2 = \frac{114}{259}, \quad w_3 = \frac{55}{259}.$$

Así que:

$$\mathbf{w} = (\frac{90}{259}, \frac{114}{259}, \frac{55}{259}) \simeq (0{,}35, 0{,}44, 0{,}21).$$

Esto quiere decir que, a partir de un paso n suficientemente grande, independientemente del estado $i \in S$ inicial, la probabilidad de, en el paso $m \geq n$, estar en el estado 1 es 0,35, en el estado 2 es 0,44 y en el estado 3 es $0,21$.

APÉNDICE B

DEMOSTRACIONES PENDIENTES DE LOS CAPÍTULOS

En este apéndice vamos a realizar las demostraciones que no han podido ser incluidos en el nudo del trabajo. Dividiremos las demostraciones por las diferentes secciones donde se encuentran los enunciados de las mismas en el nudo.

B.1 Cadenas de Markov

B.1.1 Introducción

Demostración B.1.1. *Queremos ver que la aplicación está bien definida. Sea* $\mathbf{x} \in \ell_1$ *arbitrario pero fijado. Para ver que está bien definida, comprobaremos si el vector*

$$\mathrm{P}(\mathbf{x}) = \mathbf{x}\mathbf{P} = (\sum_{i\in S} x_i \mathbf{P}_1^i, \sum_{i\in S} x_i \mathbf{P}_2^i, \dots)$$

pertenece al conjunto ℓ_1*. Para ello, debemos ver si la suma de las componentes del vector anterior es absolutamente convergente.*

Tenemos que:

$$\sum_{j\in S}\sum_{i\in S} |x_i \mathbf{P}_j^i| = |\mathbf{x}\mathbf{P}_1 + \mathbf{x}\mathbf{P}_2 + \cdots|,$$

aplicando la desigualdad triangular:

$$|\mathbf{x}\mathbf{P}_1 + \mathbf{x}\mathbf{P}_2 + \cdots| \le |\mathbf{x}\mathbf{P}_1| + |\mathbf{x}\mathbf{P}_2| + \cdots,$$

realizamos el producto de dentro de los valores absolutos de cada sumando:

$$|\mathbf{x}\mathbf{P}_1| + |\mathbf{x}\mathbf{P}_2| + \cdots = |x_1\mathbf{P}_1^1 + x_2\mathbf{P}_1^2 + \cdots| + |x_1\mathbf{P}_2^1 + x_2\mathbf{P}_2^2 + \cdots| + \cdots,$$

de nuevo, por la desigualdad triangular:

$$|x_1\mathbf{P}_1^1 + x_2\mathbf{P}_1^2 + \cdots| + |x_1\mathbf{P}_2^1 + x_2\mathbf{P}_2^2 + \cdots| + \cdots \leq |x_1\mathbf{P}_1^1| + |x_2\mathbf{P}_1^2| + \cdots + |x_1\mathbf{P}_2^1| + |x_2\mathbf{P}_2^2| + \cdots =$$

$$|x_1||\mathbf{P}_1^1| + |x_2||\mathbf{P}_1^2| + \cdots + |x_1||\mathbf{P}_2^1| + |x_2||\mathbf{P}_2^2| + \cdots,$$

sacando factor común, obtenemos:

$$|x_1||\mathbf{P}_1^1| + |x_2||\mathbf{P}_1^2| + \cdots + |x_1||\mathbf{P}_2^1| + |x_2||\mathbf{P}_2^2| + \cdots = \sum_{j\in S}\sum_{i\in S}|x_i||\mathbf{P}_j^i|.$$

Ahora, como hemos comprobado en la Igualdad (B.4) de la Demostración B.1.2, sumaremos esta doble serie de otra manera, siendo el resultado el mismo, para poder aplicar nuestras hipótesis:

$$\sum_{j\in S}\sum_{i\in S}|x_i||\mathbf{P}_j^i| = \sum_{i\in S}|x_i|\sum_{j\in S}|\mathbf{P}_j^i|. \tag{B.1}$$

Por hipótesis sabemos que $\mathbf{x} \in \ell_1$ *y* $\|\mathbf{P}\| < +\infty$*, o sea que*

$$\sum_{j\in S}|x_j| < +\infty \quad y \quad \sup_{i\in S}\{\sum_{j\in S}|\mathbf{P}_j^i|\} < +\infty.$$

Aplicando esto a la Ecuación B.1, nos queda que:

$$\sum_{i\in S}|x_i|\sum_{j\in S}|\mathbf{P}_j^i| \leq \|\mathbf{P}\|\sum_{i\in S}|x_i| < +\infty.$$

En definitiva:

$$\sum_{j\in S}\sum_{i\in S}|\mathrm{P}(\mathbf{x})| = \sum_{j\in S}\sum_{i\in S}|x_i\mathbf{P}_j^i| < +\infty.$$

Por lo que

$$\mathrm{P}(\mathbf{x}) \in \ell_1, \qquad \forall \mathbf{x} \in \ell_1,$$

ya que el $\mathbf{x}$ *inicial elegido era arbitrario pero fijado.*

$\square$

Demostración B.1.2. *Primero veremos la implicación de izquierda a derecha (⇒) y después de derecha a izquierda (⇐):*

- *Sean* $\mathbf{P}$ *una matriz estocástica y* $\mathbf{x} \in \mathcal{T}$ *arbitrarios pero fijados. Lo que queremos probar es que al aplicar* P *a* $\mathbf{x}$ *obtenemos un vector estocástico. Por tanto, veamos si* $\mathrm{P}(\mathbf{x})$ *cumple las condiciones para ser un vector estocástico:*

  a) *El vector* $\mathrm{P}(\mathbf{x})$ *tiene la siguiente forma:*

  $$\mathrm{P}(\mathbf{x}) = \mathbf{xP} = (\sum_{i\in S}x_i\mathbf{P}_1^i, \sum_{i\in S}x_i\mathbf{P}_2^i, \ldots, \sum_{i\in S}x_i\mathbf{P}_N^i). \tag{B.2}$$

  Sea $j \in S$ *arbitrario pero fijado. La componente* j*-ésima es:*

$$\sum_{i \in S} x_i \mathbf{P}_j^i = x_1 \mathbf{P}_j^1 + \cdots + x_N \mathbf{P}_j^N. \tag{B.3}$$

Por ser $\mathbf{P}$ *una matriz estocástica y* $\mathbf{x} \in \mathscr{T}$ *(lo sabemos por hipótesis) tenemos que*

$$x_i \geq 0 \quad y \quad \mathbf{P}_j^i \geq 0, \forall\, i, j \in S.$$

Por tanto, en particular, tendríamos que cada sumando de la Igualdad (B.3) sería igual o mayor que 0. *Así que:*

$$\sum_{i \in S} x_i \mathbf{P}_j^i \geq 0.$$

Como el j *elegido era arbitrario pero fijado, esto se cumple para todo* $j \in S$*, o sea, todas las componentes del Vector (B.2) son mayores o iguales que* 0.

b) Vamos a comprobar si las componentes del Vector (B.2) suman 1.

Notemos que es igual sumar los componentes de esta forma:

$$\begin{gathered} x_1 \mathbf{P}_1^1 + \cdots + x_N \mathbf{P}_1^N \\ x_1 \mathbf{P}_2^1 + \cdots + x_N \mathbf{P}_2^N \\ \vdots \\ x_1 \mathbf{P}_N^1 + \cdots + x_N \mathbf{P}_N^N \end{gathered}$$

que de esta otra, sacando factor común:

$$\begin{gathered} x_1 (\mathbf{P}_1^1 + \cdots + x_N \mathbf{P}_N^1) \\ x_2 (\mathbf{P}_1^2 + \cdots + x_N \mathbf{P}_N^2) \\ \vdots \\ x_N (\mathbf{P}_1^N + \cdots + x_N \mathbf{P}_N^N). \end{gathered}$$

Por ello tenemos la siguiente igualdad:

$$\sum_{j \in S} \sum_{i \in S} x_i \mathbf{P}_j^i = \sum_{i \in S} \left(x_i \sum_{j \in S} \mathbf{P}_j^i \right). \tag{B.4}$$

Como $\mathbf{P}$ *es una matriz estocástica, tenemos que* $\sum_{j \in S} \mathbf{P}_j^i = 1$ *para cada* $i \in S$.

Por tanto:

$$\sum_{j \in S} \sum_{i \in S} x_i \mathbf{P}_j^i = \sum_{i \in S} \left(x_i \sum_{j \in S} \mathbf{P}_j^i \right) = \sum_{i \in S} x_i.$$

Y, por ser $\mathbf{x} \in \mathscr{T}$*, tenemos que* $\sum_{i \in S} x_i = 1$*. En definitiva:*

$$\sum_{j \in S} \sum_{i \in S} x_i \mathbf{P}_j^i = 1.$$

Como hemos completado la demostración para una matriz estocástica $\mathbf{P}$ *y* $\mathbf{x} \in \mathcal{T}$ *arbitrarios pero fijados, tenemos que esto se cumple para cualquier matriz estocástica y cualquier vector estocástico. Y, por tanto, tenemos que:*

$$\mathrm{P}(\mathcal{T}) \subset \mathcal{T}.$$

- *Por otro lado, supongamos que* $\mathrm{P}(\mathcal{T}) \subset \mathcal{T}$. *Queremos demostrar que la matriz* $\mathbf{P}$ *es estocástica. Para hacerlo, veremos que cada una de las filas de la matriz* $\mathbf{P}$ *es estocástica:*

  Sea $i \in S$ *arbitrario pero fijado y sea* $\mathbf{e^i} \in \mathbb{R}^N$ *el vector* i*-ésimo de la base canónica de* $\mathbb{R}^N$. *Claramente el vector* $\mathbf{e^i} \in \mathcal{T}$ *(ver Ejemplo 2.1.1) y, por tanto, tenemos que:*

$$\mathrm{P}(\mathbf{e^i}) \in \mathcal{T}, \tag{B.5}$$

  por hipótesis. Ahora bien:

$$\mathrm{P}(\mathbf{e^i}) = \mathbf{e^i P} = (\mathbf{P}^i_1, \mathbf{P}^i_2, \dots, \mathbf{P}^i_N).$$

  O sea, $\mathrm{P}(\mathbf{e^i})$ *es la fila* i*-ésima de la matriz* $\mathbf{P}$. *Entonces, por la Relación (B.5), tenemos que la fila* i*-ésima de* $\mathbf{P}$ *es un vector estocástico.*

  Como el i *elegido era arbitrario pero fijado, tenemos que lo demostrado se cumple para cualquier* $i \in S$, *lo que implica que la matriz* $\mathbf{P}$ *sea estocástica.*

□

Demostración B.1.3. *Sean* $\mathbf{P}$ *y* $\mathbf{Q}$ *dos matrices indexadas por S. Entonces definimos las aplicaciones lineales asociadas a estas matrices:*

$$\begin{array}{rccc} \mathrm{P}: & \ell_1 & \longrightarrow & \ell_1 \\ & \mathbf{x} & \longrightarrow & \mathbf{xP} \end{array} \quad y \quad \begin{array}{rccc} \mathrm{Q}: & \ell_1 & \longrightarrow & \ell_1 \\ & \mathbf{x} & \longrightarrow & \mathbf{xQ} \end{array}$$

Ambas aplicaciones están bien definidas y además cumplen que:

$$(\mathrm{Q} \circ \mathrm{P})(\mathbf{x}) = \mathrm{Q}(\mathrm{P}(\mathbf{x})) = \mathrm{Q}(\mathbf{xP}),$$

y por ser $\mathbf{xP} \in \ell_1$*:*

$$\mathrm{Q}(\mathbf{xP}) = (\mathbf{xP})\mathbf{Q} = \mathbf{x}(\mathbf{PQ}).$$

Por tanto:

$$(\mathrm{Q} \circ \mathrm{P})(\mathbf{x}) = \mathbf{x}(\mathbf{PQ}).$$

De esta forma definiríamos la composición de las aplicaciones Q *y* P *respectivamente. Esta composición está bien definida, así como el producto de matrices*

$$(\mathbf{PQ})^i_j = \sum_{k \in S} \mathbf{P}^i_k \mathbf{Q}^k_j.$$

Y $\mathbf{x}(\mathbf{PQ}) \in \ell_1$ *ya que sabemos que* $\|\mathbf{P}\|, \|\mathbf{Q}\| < +\infty$, *por tanto:*

$$\|\mathbf{PQ}\| \leq \|\mathbf{P}\|\|\mathbf{Q}\| < +\infty.$$

(Continuaríamos la demostración de que $\mathbf{x}(\mathbf{PQ}) \in \ell_1$ *de manera análoga a como lo hemos hecho en la Demostración B.1.1).*

Por último, decir que podríamos definir de igual manera, a como hemos definido esta composición de aplicaciones, la aplicación $(\mathrm{P} \circ \mathrm{Q})$.

Demostración B.1.4. *Para ver que* $\mathbf{P}$ *es una matriz estocástica, demostraremos que cumple con las dos condiciones para serlo:*

a) Como $t_k \geq 0$ *y* $\mathbf{P}_k$ *es matriz estocástica* $\forall\, k \in \{1, \dots, M\}$, *tenemos que:*

$$\mathbf{P}^i_j = \sum_{k=1}^{M} t_k (\mathbf{P}_k)^i_j \geq 0 \qquad \forall\, i, j \in S.$$

b) Sea $i \in S$ *arbitrario pero fijado, queremos ver si* $\sum_{j \in S} \mathbf{P}^i_j = 1$.

$$\sum_{j \in S} \mathbf{P}^i_j = \sum_{j \in S} \sum_{k=1}^{M} t_k (\mathbf{P}_k)^i_j = \sum_{k=1}^{M} t_k \sum_{j \in S} (\mathbf{P}_k)^i_j.$$

Ahora, con i *fijado, tenemos que* $\sum_{j \in S} (\mathbf{P}_k)^i_j = 1$ *para cada* $k \in \{1, \dots, M\}$. *Por tanto:*

$$\sum_{k=1}^{M} t_k \sum_{j \in S} (\mathbf{P}_k)^i_j = \sum_{k=1}^{M} t_k,$$

y como $\sum_{k=1}^{M} t_k = 1$ *por hipótesis, obtenemos finalmente que:*

$$\sum_{j \in S} \mathbf{P}^i_j = 1.$$

Como $i \in S$ *era arbitrario pero fijado, esto se cumple* $\forall\, i \in S$.

□

Demostración B.1.5. *Para ver que* $i \longleftrightarrow j$ *es una relación de equivalencia debo ver que cumple las propiedades: reflexiva, simétrica y transitiva* $\forall\, i, j \in S$. *Sean* $i, j, k \in S$ *t.q.* $i \neq j, j \neq k$ *e* $i \neq k$ *arbitrarios pero fijados:*

1) Se cumple por la Observación 2.1.2. (Propiedad reflexiva).

2) Veamos que si $i \longleftrightarrow j$, *entonces* $j \longleftrightarrow i$.

Como $i \longleftrightarrow j$ *tenemos que* $i \longrightarrow j$ *y* $j \longrightarrow i$, *por tanto,* $j \longrightarrow i$ *e* $i \longrightarrow j$, *lo que quiere decir que* $j \longleftrightarrow i$. *(Propiedad simétrica).*

3) *Queremos ver que si* $i \longleftrightarrow j$ *y* $j \longleftrightarrow k$*, entonces* $i \longleftrightarrow k$.

   Como $i \longleftrightarrow j$ *y* $j \longleftrightarrow k$ *tenemos que, en particular,* $i \longrightarrow j$ *y* $j \longrightarrow k$*, por tanto, existe un camino de* i *a* j *y de* j *a* k*. Si unimos ambos caminos obtenemos un camino de* i *a* k *y, en definitiva,* $i \longrightarrow k$.

   De manera análoga veríamos que $k \longrightarrow i$*. Por tanto,* $i \longleftrightarrow k$*.(Propiedad transitiva).*

□

Demostración B.1.6. *Lo que queremos demostrar es que todo camino de* $i \in S$ *a* $j \in S$*, está formado por mínimo* 1 *paso (*1 *arco) y máximo* N *pasos.*

Supongamos que existe un camino dirigido de i *a* j *de* $N+1$ *arcos:*

$$(i, i_1)(i_1, i_2)\cdots(i_{N-1}, i_N)(i_N, j).$$

Ahora bien, cada 2 *nodos tenemos* 1 *arco, y cada* 3 *nodos* 2 *arcos, y así sucesivamente. Por lo que si tenemos* $N+1$ *arcos en este camino dirigido, significaría que hay* $N+2$ *nodos en el camino. Como solo hay* N *nodos (*$|S| = N$*) en el grafo, tenemos que al menos hay un ciclo en este camino.*

Por hipótesis, sabemos que el camino de i *a* j *existe, por lo que debemos poder identificar los nodos que se repiten. Tenemos dos casos:*

- *Si se repite uno de los nodos intermedios del camino (*$\{i_1, \dots, i_N\}$*), al estar considerando caminos dirigidos simples, debemos poder quitar al menos* 2 *aristas del ciclo que se forme para evitar este ciclo. De este modo, nos quedará un camino de* $N-1$ *arcos máximo.*

- *Si únicamente se repiten el nodo inicial (*i*) con el nodo final (*j*) y no se repite ningún nodo intermedio, entonces tenemos que el camino entero es un ciclo, y tendríamos* N *aristas en total, lo cual es una contradicción.*

  Por lo que únicamente puede darse el primer caso.

Si el camino inicial hubiera sido de más de N *arcos, para realizar la demostración hubiéramos aplicado este razonamiento de forma iterada hasta obtener un camino de máximo* $N-1$ *arcos.*

□

B.1.2 Cadenas de Markov

Demostración B.1.7. *Primero de todo, veamos a qué es igual el vector fila* $\pi(n+1)$ *componente a componente:*

Sea $j \in S$ *arbitrario pero fijado:*

$$\pi(n+1)_j = \mathbb{P}(\mathscr{X}_{n+1} = j).$$

Definimos $E_i = \{\mathscr{X}_n = i\}$ *para cada* $i \in S$*. El conjunto de los* E_i *forma una partición sobre* Ω*. Por el teorema de la probabilidad total,* [1] *tenemos que:*

[1] Véase el Teorema A.3.1 del Apéndice A para consultar qué dice el teorema.

$$\mathbb{P}(\mathscr{X}_{n+1} = j) = \sum_{i \in S} \mathbb{P}(\mathscr{X}_{n+1} = j / E_i)\mathbb{P}(E_i).$$

Ahora, para los sumandos tt.q. $\mathbb{P}(E_i) > 0$*, tenemos que* $\mathbb{P}(E_i) = \mathbb{P}(\mathscr{X}_n = i) = \pi(n)_i$ *y, por otro lado,* $\mathbb{P}(\mathscr{X}_{n+1} = j / E_i) = \mathbf{P}(n+1)^i_j$.

Nota B.1.1. *Notemos que no todos los sumandos* $\mathbb{P}(E_i)$ *pueden valer* 0*, ya que los* E_i *forman una partición de* Ω *y* $\mathbb{P}$ *es una medida de probabilidad, por ello*

$$1 = \mathbb{P}(\Omega) = \mathbb{P}\left(\bigcup_{i \in S} E_i\right) = \sum_{i \in S} \mathbb{P}(E_i).$$

Siguiendo con la demostración:

$$\sum_{i \in S} \mathbb{P}(\mathscr{X}_{n+1} = j / E_i)\mathbb{P}(E_i) = \sum_{i \in S} \mathbf{P}(n+1)^i_j \pi(n)_i.$$

En definitiva,

$$\pi(n+1)_j = \sum_{i \in S} \mathbf{P}(n+1)^i_j \pi(n)_i, \ \forall j \in S.$$

Como j *era arbitrario pero fijado, tenemos que:*

$$\pi(n+1) = \mathbf{P}(n+1)\pi(n),$$

para cada $n \in \mathbb{N} \cup \{0\}$*. En particular, si vamos sustituyendo para cada valor de* n*, desde* $n = 0$*, obtendremos lo siguiente:*

$$\begin{gathered} \pi(1) = \mathbf{P}(1)\pi(0) \\ \pi(2) = \mathbf{P}(2)\pi(1) = \mathbf{P}(2)\mathbf{P}(1)\pi(0) \\ \vdots \\ \pi(n) = \mathbf{P}(n)\mathbf{P}(n-1)\cdots\mathbf{P}(1)\pi(0) \\ \pi(n+1) = \mathbf{P}(n+1)\mathbf{P}(n)\cdots\mathbf{P}(1)\pi(0). \end{gathered}$$

(Fácilmente podríamos demostrar esta fórmula, para n*, mediante el método de inducción).*

Por lo que:

$$\pi(n) = \mathbf{P}(n)\mathbf{P}(n-1)\cdots\mathbf{P}(1)\pi(0), \ \forall n \in \mathbb{N} \cup \{0\}.$$

Ahora queremos ver que para cualquier $k \in \mathbb{N} \cup \{0\}$*:*

$$\pi(n+k) = \mathbf{P}(n+k)\cdots\mathbf{P}(2+k)\mathbf{P}(1+k)\pi(k).$$

Veremos esto por inducción (esta inducción sería similar a la que no hemos hecho antes):

- *Caso base: para* $k = 0$ *acabamos de ver que se cumple.*
- *Paso de inducción: Supongamos que es cierto para* k *y veamos si sigue siendo cierto para* $k+1$*:*

  Sabemos que:

$$\pi(n+k+1) = \mathbf{P}(n+k+1)\pi(n+k).$$

Por hipótesis de inducción tenemos que:

$$\pi(n+k) = \mathbf{P}(n+k)\cdots\mathbf{P}(2+k)\mathbf{P}(1+k)\pi(k).$$

Y sustituyendo en la ecuación anterior:

$$\pi(n+k+1) = \mathbf{P}(n+k+1)\pi(n+k) = \mathbf{P}(n+k+1)\mathbf{P}(n+k)\cdots\mathbf{P}(2+k)\mathbf{P}(1+k)\pi(k).$$

Por tanto, por inducción hemos acabado de demostrar que

$$\pi(n+k) = \mathbf{P}(n+k)\cdots\mathbf{P}(2+k)\mathbf{P}(1+k)\pi(k),\ \forall n,k \in \mathbb{N}\cup\{0\},$$

tal y como queríamos ver.

$\square$

B.1.3 Representación canónica de P

Demostración B.1.8. *Vamos a demostrar que si C y D son clases cerradas del conjunto de estados S, entonces $C \cup D$ y $C \cap D$ también lo son:*

- *Veamos si $C \cup D$ es una clase cerrada. Sea $i \in C \cup D$ un estado arbitrario pero fijado. Vamos a realizar la demostración por reducción al absurdo. Supongamos que $\exists j \notin C \cup D$ t.q. $i \longrightarrow j$. Tenemos 2 subcasos, ya que $i \in C$ p $i \in D$ (el caso en que $i \in C \cap D$ lo obviamos porque es un caso particular de cualquiera de los dos anteriores):*

  - *Sea $i \in C$. Como $j \notin C \cup D$, entonces $j \notin C$ y sabemos que $i \longrightarrow j$, por lo tanto C no sería clase cerrada, que es una contradicción.*
  - *Sea $i \in D$. De manera similar al subcaso anterior tendríamos que D no sería clase cerrada, que sería también una contradicción.*

  Por lo que $C \cup D$ sería una clase cerrada.

- *Ahora comprobaremos que $C \cap D$ es clase cerrada. Sea $i \in C \cap D$ un estado arbitrario pero fijado y $j \in S$ t.q. $i \longrightarrow j$:*

  $$i \in C \cap D \Rightarrow \left\{ \begin{matrix} i \in C,\ i \longrightarrow j \Rightarrow j \in C \\ i \in D,\ i \longrightarrow j \Rightarrow j \in D \end{matrix} \right\} \Rightarrow j \in C \cap D.$$

  Por tanto, por definición de clase cerrada, $C \cap D$ sería una clase cerrada.

$\square$

Demostración B.1.9. *Empecemos con la implicación de derecha a izquierda ($\Leftarrow$):*

Sea C una clase cerrada con todos sus elementos conjuntamente comunicados dos a dos. Supongamos que C no es minimal y lleguemos a contradicción. Al no ser C minimal, existe una clase cerrada D t.q. $D \subseteq C$ y $D \neq C$.

Si C contuviera solo 1 único elemento tendríamos que $D = C$ y por tanto ya tendríamos la contradicción. Supongamos que C tiene al menos 2 elementos. Como $D \subseteq C$ y $D \neq C$, $\exists i \in C$ t.q. $i \notin D$. Sea $j \in D$ (este j existe porque D tiene al menos 1 elemento), como C tiene todos sus elementos comunicados dos a dos por hipótesis, tenemos que $i \longleftrightarrow j$, por tanto $j \longrightarrow i$, $j \in D$ e $i \notin D$. Esto sería una contradicción, debido a que D era clase cerrada.

La contradicción proviene de suponer que C no era clase cerrada minimal, por lo que C es clase cerrada minimal.

Vayamos ahora con la implicación de izquierda a derecha ($\Rightarrow$):

Sea C una clase cerrada minimal y $j \in C$. Definimos los conjuntos de estados $D_j = \{i \in C \,/\, i \nrightarrow j\}$. Queremos ver que $D_j = \emptyset$, de esta forma todos los estados estarían conjuntamente comunicados dos a dos ya que no habría ningún estado $i \in C$ t.q. $i \nrightarrow j$.

Primero de todo, tenemos que $j \notin D_j$, debido a que sabemos que $j \longleftrightarrow j$, $\forall j \in S$. Por tanto $D_j \subset C$. Lo que haremos ahora será suponer que $D_j \neq \emptyset$ y como $D_j \subset C$, tendríamos que, por cómo está definida, sería clase cerrada y además C no sería minimal, lo que nos llevaría a una contradicción.

Entonces suponemos que $D_j \neq \emptyset$. Veremos que es clase cerrada. Para probarlo debemos ver que si $k \notin D_j$ e $i \in D_j$ entonces $i \nrightarrow k$. Distinguimos dos casos:

- *Si $k \notin C$ tenemos que $i \nrightarrow k$ ya que C es clase cerrada e $i \in C$.*
- *Si $k \in C \setminus D_j$ entonces $k \longrightarrow j$ por definición de D_j. Si $i \longrightarrow k$, tendríamos un camino de i a j, pero es una contradicción porque $i \in D_j$. Por lo que $i \nrightarrow k$.*

Por lo que D_j sería clase cerrada y $D_j \subset C$, entonces C no sería clase cerrada minimal, que es una contradicción.

Esta contradicción proviene de suponer que $D_j \neq \emptyset$, por lo que $D_j = \emptyset$ para $j \in C$ y esto implica (ya lo hemos explicado antes) que todos los estados de C están comunicados dos a dos.

□

Demostración B.1.10. *Haremos las demostraciones de los apartados por orden.*

- *Empecemos primero demostrando i):*

  Sea $j \in C_s$ para algún $s \in \{1, \ldots, r\}$. Lo que queremos ver ahora es que j es recurrente.

  Al ser C_s una clase cerrada, tenemos que $p_{jh}^{(n)} = 0$ para cualquier $h \notin C_s$ y cualquier $n \in \mathbb{N}$. Esto quiere decir que, por ser la matriz $\mathbf{P}$ estocástica, $\sum_{h \in C_s} p_{jh}^{(n)} = 1$ para cada $n \in \mathbb{N}$. Por tanto:

$$\sum_{h \in C_s} \sum_{n=1}^{+\infty} p_{jh}^{(n)} = \sum_{n=1}^{+\infty} \sum_{h \in C_s} p_{jh}^{(n)} = +\infty.$$

En particular existe un $j_0 \in C_s$ *t.q.* $\sum_{n=1}^{+\infty} p_{jj_0}^{(n)} = +\infty$. *Como* C_s *es minimal tenemos que por la Proposición 2.4.2* $j_0 \longleftrightarrow j$, *por ello existe un* $k \in \mathbb{N}$ *t.q.* $p_{j_0 j}^{(k)} > 0$. *Así pues:*

$$p_{jj}^{(n+k)} = \sum_{\alpha \in S} p_{j\alpha}^{(n)} \cdot p_{\alpha j}^{(k)} \geq p_{jj_0}^{(n)} \cdot p_{j_0 j}^{(k)} \Rightarrow \sum_{n=1}^{+\infty} p_{jj}^{(n+k)} \geq \sum_{n=1}^{+\infty} p_{jj_0}^{(n)} \cdot p_{j_0 j}^{(k)} \Rightarrow \sum_{n=1}^{+\infty} p_{jj}^{(n+k)} = +\infty,$$

como $\{\mathscr{X}_n\}$ *es una C.M.H. tenemos que* $p_{jj}^{(n+k)} = p_{jj}^{(n)}$. *Por lo que:*

$$\sum_{n=1}^{+\infty} p_{jj}^{(n)} = +\infty,$$

que significa que j *es recurrente, tal y como queríamos ver.*

- *Veamos ahora* ii*):*

  Sea $i \in T$. *Definimos el conjunto* $D = \{j \in S \,/\, i \longrightarrow j\}$ *que es una clase cerrada.*[2] *Por ser una clase cerrada, tenemos que alguna clase cerrada minimal estará contenida en* D *o será igual a* D, *por lo que* $C \cap D \neq \emptyset$.

  Por otro lado, sabemos que $p_{iC}^{(n)} = \sum_{j \in C} p_{ij}^{(n)}$ *y como* $C \cap D \neq \emptyset$, $\exists j_0 \in C \cap D$ *t.q.* $p_{ij_0}^{(n)} > 0$ *para algún* $n \in \mathbb{N}$. *Sea* $\overline{k} = \overline{k}(i) \in \mathbb{N}$ *este* n *que hemos nombrado, o sea:* $p_{iC}^{(\overline{k})} > 0$ *(lo denotamos así,* $\overline{k}(i)$*, porque existe uno para cada* $i \in T$ *y pueden ser diferentes entre sí).*

  Vamos a ver que $\{p_{iC}^{(k)}\}$ *es una sucesión monótona no-decreciente con respecto a* k*:*

  Primero de todo, para cualquier $\alpha \in C$ *tenemos que* $p_{\alpha j} = 0, \forall j \notin C$ *ya que* C *es una clase cerrada al ser unión de clases cerradas.*[3] *Por tanto, como* $\mathbf{P}$ *es estocástica, tenemos que* $1 = \sum_{j \in C} p_{\alpha j} + \sum_{j \notin C} p_{\alpha j} = \sum_{j \in C} p_{\alpha j}$.

  Una vez sabemos esto, sea $k \in \mathbb{N}$ *arbitrario pero fijado:*

$$p_{iC}^{(k+1)} = \sum_{j \in C} p_{ij}^{(k+1)} = \sum_{j \in C} \sum_{\alpha=1}^{N} p_{i\alpha}^{(k)} p_{\alpha j} = \sum_{\alpha=1}^{N} p_{i\alpha}^{(k)} \left(\sum_{j \in C} p_{\alpha j} \right) \geq \sum_{\alpha \in C} p_{i\alpha}^{(k)} \cdot 1 = \sum_{\alpha \in C} p_{i\alpha}^{(k)} = p_{iC}^{(k)}.$$

  Por lo que $p_{iC}^{(k+1)} \geq p_{iC}^{(k)}, \forall k \in \mathbb{N}$.

  Esto quiere decir que para cualquier $i \in T$, *se cumple que* $p_{iC}^{(k)} \geq p_{iC}^{(\overline{k})} > 0$ *si* $k \geq \overline{k}$ *(que* $p_{iC}^{(\overline{k})} > 0$ *lo hemos visto antes).*

  Lo que veremos ahora es que, para cada $k \in \mathbb{N}$, *podemos acotar* $p_{iT}^{(k)}$ *superiormente por un valor entre* 0 *y* 1*:*

[2] Véase la explicación justo posterior a la Definición 2.4.1 para saber porqué.

[3] Por la Proposición 2.4.1.

Al ser T finito (ya que C y S lo son) y por lo que acabamos de ver, para cada $i \in T$, existen $p_0 > 0$ ($p_0 \in (0,1)$) y $k_0 \in \mathbb{N}$ tt.q. $p_{iC}^{(k)} \geq p_0$, $\forall k \geq k_0$. En particular:

$$p_{iT}^{(k_0)} = \sum_{j \in T} p_{ij}^{(k_0)},$$

y como $\mathbf{P}$ *es una matriz estocástica y $T = S \setminus C \Rightarrow C \cup T = S$, tenemos que:*

$$1 = \sum_{j \in S} p_{ij}^{(k_0)} = \sum_{j \in C} p_{ij}^{(k_0)} + \sum_{j \in T} p_{ij}^{(k_0)} = p_{iC}^{(k_0)} + \sum_{j \in T} p_{ij}^{(k_0)},$$

por tanto:

$$\sum_{j \in T} p_{ij}^{(k_0)} = p_{iT}^{(k_0)} = 1 - p_{iC}^{(k_0)} \leq 1 - p_0,$$

ya que $p_{iC}^{(k_0)} \geq p_0$. En definitiva:

$$p_{iT}^{(k_0)} \leq 1 - p_0.$$

Veamos ahora si podemos acotar $p_{iT}^{(2k_{min})}$, con $k_{min} = \text{mín}\{k_0, k_1\}$ y $k_1 \in \mathbb{N}$ (ahora veremos de donde surge este k_1):

$$p_{iT}^{(2k_{min})} = \sum_{j \in T} p_{ij}^{(2k_{min})} = \sum_{j \in T} \sum_{\alpha=1}^{N} p_{i\alpha}^{(k_{min})} p_{\alpha j}^{(k_{min})} = \sum_{\alpha=1}^{N} p_{i\alpha}^{(k_{min})} \sum_{j \in T} p_{\alpha j}^{(k_{min})},$$

como hemos visto ya antes en esta misma demostración, $p_{\alpha j}^{(k)} = 0$ para $\alpha \in C$ y $j \notin C$. Por ello, basta que en el sumatorio $\sum_{\alpha=1}^{N} p_{i\alpha}^{(k_{min})}$ sumemos los $\alpha \notin C$, o sea, los $\alpha \in T$, ya que los demás valdrán 0. Así pues:

$$\sum_{\alpha=1}^{N} p_{i\alpha}^{(k_{min})} \sum_{j \in T} p_{\alpha j}^{(k_{min})} = \sum_{\alpha \in T} p_{i\alpha}^{(k_{min})} \sum_{j \in T} p_{\alpha j}^{(k_{min})} = p_{iT}^{(k_{min})} p_{\alpha T}^{(k_{min})} \leq (1 - p_0)(1 - p_1),$$

donde este p_1 surge de que para $\alpha \in T$ también existe un $p_1 > 0$ ($p_1 \in (0,1)$) y $k_1 \in \mathbb{N}$ tt.q. $p_{\alpha C}^{(k)} \geq p_1$, $\forall k \geq k_1$. Sea $p_{min} = \text{mín}\{p_0, p_1\}$, entonces:

$$(1 - p_0)(1 - p_1) \leq (1 - p_{min})^2.$$

En definitiva:

$$p_{iT}^{(2k_{min})} \leq (1 - p_{min})^2.$$

Y, con un argumento de inducción, veríamos fácilmente que:

$$p_{iT}^{(q\cdot k_{min})} \leq (1-p_{min})^q, \quad \forall q \in \mathbb{N}.$$

(Considerando el p_{min} y k_{min} de entre los q diferentes posibles).

Recordemos que hemos visto esto porque queríamos demostrar que para cada $k \in \mathbb{N}$, podíamos acotar $p_{iT}^{(k)}$ superiormente por un valor entre 0 y 1. Ahora, para cada entero positivo k, dividimos por k_{min} (que recordemos que existe y es menor o igual que k) y obtenemos: $k = q \cdot k_{min} + r$, para $0 \leq r < k_{min}$ y q entero positivo (este q puede cambiar dependiendo del entero k, al igual que el valor r).

Por otro lado, recordemos que hemos visto que $\{p_{iC}^{(k)}\}$ es una sucesión monótona no-decreciente con respecto a k y que $p_{iT}^{(k)} = 1 - p_{iC}^{(k)}$. Esto quiere decir que $\{p_{iT}^{(k)}\}$ es monótona no-creciente. Por tanto:

$$\sum_{k=1}^{+\infty} p_{iT}^{(k)} \leq \sum_{q=1}^{+\infty} p_{iT}^{(q\cdot k_{min})},$$

ya que $\{p_{iT}^{(k)}\}$ es monótona no-creciente y $k \geq k_{min}$. Y, por ser $0 \leq r < k_{min}$, tenemos que $k_{min} \in \mathbb{N}$ (porque r es un entero positivo) y por ello:

$$\sum_{q=1}^{+\infty} p_{iT}^{(q\cdot k_{min})} \leq k_{min} + k_{min} \sum_{q=1}^{+\infty} p_{iT}^{(q\cdot k_{min})} \leq k_{min} + k_{min} \sum_{q=1}^{+\infty} (1-p_{min})^q.$$

La serie $\sum_{q=1}^{+\infty} (1-p_{min})^q$ es una serie geométrica con $(1-p_{min}) \in (0,1)$ como razón, así que el valor de la serie es

$$\frac{1-p_{min}}{1-(1-p_{min})} = \frac{1-p_{min}}{p_{min}},$$

ya que $(1-p_{min})^{+\infty} = 0$. Aplicando esto a la desigualdad anterior tenemos que:

$$k_{min} + k_{min} \sum_{q=1}^{+\infty} (1-p_{min})^q \leq k_{min} + k_{min} \frac{1-p_{min}}{p_{min}} = \frac{k_0}{p_{min}}.$$

En definitiva,

$$\sum_{k=1}^{+\infty} p_{iT}^{(k)} \leq \frac{k_{min}}{p_{min}},$$

para cualquier $i \in T$. Esto quiere decir que, en particular:

$$\sum_{k=1}^{+\infty} p_{ii}^{(k)} < +\infty.$$

Por tanto i es transitorio, tal y como queríamos ver.

□

Demostración B.1.11. *Primero explicaremos porque las matrices de la diagonal de la potencia n-ésima de* $\mathbf{P}$ *son la potencia n-ésima de las matrices de la diagonal de la matriz* $\mathbf{P}$ *y luego explicaremos porque hemos denotado por* $(\mathbf{R}_n)$ *a los valores de* $\mathbf{R}$ *en* $(\mathbf{P}^n)$*:*

- *Sea* $(\mathbf{B}_l)$ *con* $l \in \{1, \ldots, r\}$ *una de las matrices de la diagonal de* $\mathbf{P}$ *y sean* $i, j \in C_l$ *tt.q.* $i \neq j$ *arbitrarios pero fijados. Definimos* $L = k_1 + \cdots + k_{l-1}$*. Lo que haremos será calcular* $(\mathbf{P}^2)^i_j$ *y ver si se corresponde con* $((\mathbf{B}_l)^2)^i_j$*:*

  El valor $(\mathbf{P}^2)^i_j$ *se obtiene de multiplicar la fila* i*-ésima de* $\mathbf{P}$*:*

  $$(0, \ldots, 0, p_{iL+1}, \ldots, p_{ij}, \ldots, p_{iL+k_l}, 0, \ldots, 0)$$

  por la columna j*-ésima de* $\mathbf{P}$*:*

  $$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ p_{L+1j} \\ \vdots \\ p_{ij} \\ \vdots \\ p_{L+k_l j} \\ 0 \\ \vdots \\ 0 \\ p_{K+1j} \\ \vdots \\ p_{Nj} \end{pmatrix}$$

  donde $L+1, \ldots, L+k_l$ *son los estados de S reetiquetados de esta forma debido a que son recurrentes y forman la clase cerrada minimal* C_l*:* $\{L+1, \ldots, L+k_l\} = C_l$*. Las probabilidades* $p_{K+1j}, \ldots, p_{Nj}$ *se corresponden con los elementos de la columna* j*-ésima de* $\mathbf{R}$*.*

  El resultado de este producto es:

  $$p_{iL+1} \cdot p_{L+1j} + \cdots + p_{ij}^2 + \cdots + p_{iL+k_l} \cdot p_{L+k_l j} = \sum_{l=1}^{k_l} p_{iL+l} \cdot p_{L+lj}$$

  que es exactamente el valor de multiplicar la fila i*-ésima de* $(\mathbf{B}_l)$ *con su columna* j*-ésima. Ya que* $(\mathbf{B}_l)$ *es:*

$$(\mathbf{B}_l) = \begin{pmatrix} p_{L+1L+1} & p_{L+1L+2} & \cdots & p_{L+1j} & \cdots & p_{L+1L+k_l} \\ p_{L+2L+1} & p_{L+2L+2} & \cdots & p_{L+2j} & \cdots & p_{L+2L+k_l} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{iL+1} & p_{iL+2} & \cdots & p_{ij} & \cdots & p_{iL+k_l} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{L+k_lL+1} & p_{NK+2} & \cdots & p_{L+k_lj} & \cdots & p_{L+k_lL+k_l} \end{pmatrix}$$

Por lo tanto $(\mathbf{P}^2)^i_j = ((\mathbf{B}_l)^2)^i_j$. *Como esto lo hemos comprobado para* l, i, j *arbitrarios pero fijados, se cumple que* $(\mathbf{P}^2)_{i,j\in C_l} = ((\mathbf{B}_l)^2)$, $\forall\, l \in \{1,\ldots,r\}$.

Y, con un argumento de inducción simple, veríamos que $(\mathbf{P}^n)_{i,j\in C_l} = ((\mathbf{B}_l)^n)$, $\forall\, l \in \{1,\ldots,r\}$.

- *La matriz* $\mathbf{V}$ *tiene la siguiente forma:*

$$\mathbf{V} = \begin{pmatrix} p_{K+1K+1} & p_{K+1K+2} & \cdots & p_{K+1N-1} & p_{K+1N} \\ p_{K+2K+1} & p_{K+2K+2} & \cdots & p_{K+2N-1} & p_{K+2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{N-1K+1} & p_{N-1K+2} & \cdots & p_{N-1N-1} & p_{N-1N} \\ p_{NK+1} & p_{NK+2} & \cdots & p_{NN-1} & p_{NN} \end{pmatrix}$$

Sean $i, j \in \{K+1,\ldots,N\}$. *La fila* i*-ésima, columna* j*-ésima de la matriz* $\mathbf{V}^2$ *vendría dada por:*

$$\sum_{l=1}^{N-K} p_{iK+l} \cdot p_{K+lj}.$$

Como antes, lo que queremos ver ahora es si $(\mathbf{P}^2)^i_j = (\mathbf{V}^2)^i_j$. *Para ello calcularemos* $(\mathbf{P}^2)^i_j$. *Multiplicaremos la fila* i*-ésima de* $\mathbf{P}$*:*

$$(p_{i1},\ldots,p_{iK},p_{iK+1},\ldots,p_{iN})$$

por la columna j*-ésima:*

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ p_{K+1j} \\ \vdots \\ p_{Nj} \end{pmatrix}$$

donde $p_{i1},\ldots,p_{iK}$ *son los valores de la fila* i*-ésima de* $\mathbf{R}$. *El resultado del producto sería:*

$$p_{iK+1} \cdot p_{K+1j} + \cdots + p_{iN} \cdot p_{Nj} = \sum_{l=1}^{N-K} p_{iK+l} \cdot p_{K+lj}.$$

Por tanto, $(\mathbf{P}^2)^i_j = (\mathbf{V}^2)^i_j$ *tal y como queríamos ver. Como* i, j *eran arbitrarios pero fijados, tenemos que* $(p^{(2)}_{ij}) = (\mathbf{V}^2)$, $\forall i, j \in \{K+1, \ldots, N\}$.

Con un argumento de inducción similar al que hubiéramos hecho antes, veríamos que $(p^{(n)}_{ij}) = (\mathbf{V}^n)$, $\forall i, j \in \{K+1, \ldots, N\}$.

- *Por último, la matriz* **R** *vale:*

$$\mathbf{R} = \begin{pmatrix} p_{K+11} & p_{K+12} & \cdots & p_{K+1K-1} & p_{K+1K} \\ p_{K+21} & p_{K+22} & \cdots & p_{K+2K-1} & p_{K+2K} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{N-11} & p_{N-12} & \cdots & p_{N-1K-1} & p_{N-1K} \\ p_{N1} & p_{N2} & \cdots & p_{NK-1} & p_{NK} \end{pmatrix}$$

Notemos que es la única matriz no-cuadrada de las que hemos analizado hasta ahora. Sean $i \in \{K+1, \ldots, N\}$ *y* $j \in \{1, \ldots, K\}$ *arbitrarios pero fijados. Vamos a ver que* $(\mathbf{P}^2)^i_j \neq (\mathbf{R}^2)$*:*

Para obtener $(\mathbf{P}^2)^i_j$*, como hemos hecho ya en los dos casos anteriores, debemos multiplicar la fila* i*-ésima de* **P**

$$(p_{i1}, \ldots, p_{iK}, p_{iK+1}, \ldots, p_{iN})$$

por la columna j*-ésima*

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ p_{M+1j} \\ \vdots \\ p_{M+k_l j} \\ 0 \\ \vdots \\ 0 \\ p_{K+1j} \\ \vdots \\ p_{Nj} \end{pmatrix}$$

donde las probabilidades de transición $p_{iK+1}, \ldots, p_{iN}$ *son elementos de la matriz* **V**, *y* $p_{M+1j}, \ldots, p_{M+k_l j}$ *son las de la matriz* $(\mathbf{B}_l)$, *con* $M = k_1 + \cdots + k_{l-1}$ *y* $l \in \{1, \ldots, r\}$. *El producto de estos dos vectores sería:*

$$p_{iM+1} \cdot p_{M+1j} + \cdots + p_{iM+k_l} \cdot p_{M+k_lj} + p_{iK+1} \cdot p_{K+1j} + \cdots + p_{iN} \cdot p_{Nj} =$$

$$\sum_{t=1}^{k_l} p_{iM+t} \cdot p_{M+tj} + \sum_{l=1}^{N-K} p_{iK+l} \cdot p_{K+lj}.$$

Por otro lado, el producto de la fila i-ésima de $\mathbf{R}$ con la columna j-ésima resulta:

$$p_{i1} \cdot p_{K+1j} + \cdots + p_{iK} \cdot p_{Nj},$$

que es diferente, en general, al resultado del producto anterior. Por ello podemos concluir que $(\mathbf{P}^2)^i_j \neq (\mathbf{R}^2)$. Por tanto, en general, tendríamos que $(p^{(2)}_{ij}) \neq (\mathbf{R}^2)$, $\forall i \in \{K+1,\ldots,N\}$ y $j \in \{1,\ldots,K\}$ y por ello (con un argumento de inducción): $(p^{(n)}_{ij}) \neq (\mathbf{R}^n)$, $\forall i \in \{K+1,\ldots,N\}$ y $j \in \{1,\ldots,K\}$.

Por esto denotamos por $(\mathbf{R}_n)$, y no por $(\mathbf{R}^n)$, a la matriz que ocupa las posiciones de los elementos de $\mathbf{R}$ en $(\mathbf{P}^n)$.

$\square$

Demostración B.1.12. *Consideremos la siguiente igualdad:*

$$(\mathbf{Id}-\mathbf{V})(\mathbf{V}+\cdots+(\mathbf{V}^n)) = (\mathbf{V}+\cdots+(\mathbf{V}^n)) - ((\mathbf{V}^2)+\cdots+(\mathbf{V}^{n+1})) \Rightarrow$$

$$(\mathbf{Id}-\mathbf{V})(\mathbf{V}+\cdots+(\mathbf{V}^n)) = \mathbf{V} - (\mathbf{V}^{n+1}). \tag{B.6}$$

Los estados i,j correspondientes a las filas y columnas de los elementos de $\mathbf{V}$ son de T, por ello, los elementos tt.q. $i \longrightarrow j$, tienden a 0 así como $n \to +\infty$, debido al apartado ii) de la Proposición 2.4.4. Esto significa que

$$\lim_{n\to+\infty} (p^{(n)}_{ij}) = \mathbf{0} \Rightarrow \lim_{n\to+\infty} (\mathbf{V}^n) = \mathbf{0}.$$

Ahora, sabiendo esto, si aplicamos límites cuando $n \to +\infty$ a cada lado de la Igualdad (B.6) obtenemos:

$$(\mathbf{Id}-\mathbf{V}) \lim_{n\to+\infty} \sum_{k=1}^{n} (\mathbf{V}^k) = \mathbf{V} - \lim_{n\to+\infty} (\mathbf{V}^{n+1}) \Rightarrow (\mathbf{Id}-\mathbf{V}) \sum_{k=1}^{+\infty} (\mathbf{V}^k) = \mathbf{V} - \mathbf{0} \Rightarrow$$

$$(\mathbf{Id}-\mathbf{V}) \sum_{k=1}^{+\infty} (\mathbf{V}^k) = \mathbf{V} \Rightarrow \mathbf{V}^{-1}(\mathbf{Id}-\mathbf{V}) \sum_{k=1}^{+\infty} (\mathbf{V}^k) = \mathbf{Id}.$$

Por lo que $\sum_{k=1}^{+\infty} (\mathbf{V}^k)$ tiene inversa y vale $\mathbf{V}^{-1}(\mathbf{Id}-\mathbf{V})$. Por tanto,

$$\sum_{n=0}^{+\infty} (\mathbf{V}^n) = (\mathbf{Id}-\mathbf{V})^{-1}\mathbf{V},$$

tal y como queríamos ver.

$\square$

Demostración B.1.13. *Sea i un estado transitorio y $x \in \Omega$ un camino t.q. $\mathscr{X}_0(x) = i$, arbitrarios pero fijados,y definimos $K = k_1 + \cdots + k_r$.*

Este camino estará en un estado recurrente j concreto (al ser recurrente, $j \in \{1,\ldots,K\}$) en el primer paso (o sea, en $\mathscr{X}_1$) con probabilidad p_{ij}, o en más de un paso, yendo por otros estados transitorios, con probabilidad $\sum_{l=1}^{N-K} p_{iK+l} \cdot p_{K+lj}$. Recordemos que los estados $K+1,\ldots,N$ de la matriz $\mathbf{P}$ son transitorios.

Por tanto, la probabilidad de que x esté en j vendría dada por:

$$a_{ij} = p_{ij} + \sum_{l=1}^{N-K} p_{iK+l} \cdot p_{K+lj} = p_{ij} + \sum_{l=1}^{N-K} p_{iK+l} \cdot a_{lj}. \tag{B.7}$$

Notar que si los estados $K+1,\ldots,N$ son transitorios, entonces tenemos $N-K$ estados transitorios en total, por ello la probabilidad p_{K+lj} es igual a a_{lj}, con $l \in \{1,\ldots,N-K\}$ y j recurrente fijado, identificando las $N-K$ filas de la matriz $\mathbf{A}$ con las últimas $N-K$ filas de la matriz $\mathbf{P}$, ya que ambos valores representan lo mismo.

Ahora, fijémonos que p_{ij} se corresponde con el elemento $(i-K, j)$ de la matriz $\mathbf{R}$ (con i transitorio y j recurrente) y que p_{iK+l}, con $l \in \{1,\ldots,N-K\}$, se corresponde con el valor en la posición $(i-K, l)$ de la matriz $\mathbf{V}$.

Como la Ecuación B.7 se cumple para cualquier estado transitorio i y cualquier estado recurrente j, tenemos que su forma matricial viene dada por:

$$\mathbf{A} = \mathbf{R} + \mathbf{VA} \Rightarrow \mathbf{A} - \mathbf{VA} = \mathbf{R} \Rightarrow (\mathbf{Id} - \mathbf{V})\mathbf{A} = \mathbf{R} \Rightarrow \mathbf{A} = (\mathbf{Id} - \mathbf{V})^{-1}\mathbf{R} \Rightarrow \mathbf{A} = \mathbf{NR},$$

tal y como queríamos ver.

□

B.2 Convergencia de las Matrices regulares estocásticas

Demostración B.2.1. *Sea $(\mathscr{T}, \ell_1)$ el espacio normado. Considerando la distancia d_1 del Ejemplo A.1.4, tenemos que $(\mathscr{T}, d_1)$ es un espacio métrico. Vamos a ver que $\mathscr{T}$ está acotado y que es cerrado:*

- *Recordemos que si $\mathbf{x} \in \mathscr{T}$, entonces $\|\mathbf{x}\|_1 = 1$ (ver Definición 2.1.4). Por tanto:*

  $$diam(\mathscr{T}) = \sup_{\mathbf{x},\mathbf{y} \in \mathscr{T}} \{\|\mathbf{x} - \mathbf{y}\|_1\} \leq \sup_{\mathbf{x},\mathbf{y} \in \mathscr{T}} \{\|\mathbf{x}\|_1 + \|\mathbf{y}\|_1\} = 2 < +\infty,$$

  que quiere decir que $\mathscr{T}$ está acotado.

- *Podemos demostrar que $\mathscr{T}$ es cerrado gracias a que $\mathscr{T} \subset \mathbb{R}^N$ con N un entero positivo finito. O sea, gracias a que los vectores de $\mathscr{T}$ tienen un número finito de componentes. Porque si no fuera así, $\mathscr{T}$ no tiene porque ser cerrado. Veámoslo:*

  Sea $\mathbf{x}^{(n)} = \{x_j^{(n)}\} \subset [0,1]^\infty$. Esta notación quiere decir que tenemos $n \in \mathbb{N}$ vectores, con infinitas componentes cada uno, y cada componente j pertenece al intervalo $[0,1]$.

  Sea $\mathbf{x}^{(n)} = \delta_{\cdot,n}$ el vector n-ésimo y $x_j^{(n)} = \delta_{j,n}$ la componente j-ésima del vector n. Definimos este vector como sigue:

$$\delta_{j,n} = \begin{cases} 1 & si\ j = n \\ 0 & si\ j \neq n \end{cases}$$

Claramente $\delta_{\cdot,n} \in \mathcal{T}$, $\forall n$. *Ahora:*

$$\lim_{n\to+\infty} \delta_{\cdot,n} = \mathbf{0} \Rightarrow \{\delta_{\cdot,n}\} \to \mathbf{0},$$

donde $\mathbf{0}$ *es el vector de infinitas componentes que valen* 0. *Y, obviamente* $\mathbf{0} \notin \mathcal{T}$. *Esto querría decir que* $\mathcal{T}$ *no sería cerrado.*

Pero, como hemos dicho al principio, $\mathcal{T} \subset \mathbb{R}^N$ *con* N *un entero positivo finito. En este caso sí se cumple lo siguiente:*

Sea $\{\mathbf{x}_n\} \subset \mathcal{T}$ *una sucesión de elementos de* $\mathcal{T}$ *que converge a un vector* $\mathbf{x}$, *entonces* $\mathbf{x} \in \mathcal{T}$.

Y esto demuestra que $\mathcal{T}$ *es cerrado.*

La idea que demuestra esto se basa en la definición de límite de una sucesión, que dice así:

Si $\{\mathbf{x}_n\} \to \mathbf{x}$, *entonces* $\forall \epsilon > 0, \exists m \in \mathbb{N}$ *t.q.* $\mathbf{x}_n \in B(\mathbf{x},\epsilon)$, $\forall n \geq m$.

Donde $B(\mathbf{x},\epsilon)$ *es la bola de centro* $\mathbf{x}$ *y radio* ϵ *que se define así:*

$$B(\mathbf{x},\epsilon) = \{\mathbf{y} \in \mathbf{R}_*^N \ / \ d_1(\mathbf{x},\mathbf{y}) < \epsilon\}.$$

Esta definición de límite lo que nos quiere decir es que sea cual sea el valor de ϵ *que elijamos (todo lo pequeño que queramos), siempre existirán vectores de la sucesión que se encuentren dentro de la bola de centro* $\mathbf{x}$ *y radio el* ϵ *elegido.*

Por lo que, suponiendo que $\mathbf{x} \notin \mathcal{T}$ *y utilizando esta definición de límite, podríamos ver que siempre existen vectores de la sucesión que están en esta bola, y que no serían de* $\mathcal{T}$, *y esto sería una contradicción porque* $\{\mathbf{x}_n\} \subset \mathcal{T}$, $\forall n$ *(sería una demostración por reducción al absurdo).*

Ahora, como $\mathcal{T}$ *es cerrado y es un subespacio de* $(\mathbb{R}^N, d_1)$, *que es completo*[4], *tenemos que* $\mathcal{T}$ *es también completo, y, como* $\mathcal{T}$ *es también acotado, tenemos que es compacto.*

□

Demostración B.2.2. *Demostraremos cada uno de los apartados de la proposición por orden:*

i) Tenemos que $\mathbf{R}^i \in \mathcal{T}$, $\forall i \in S$, *y que* $\|\cdot\|_1$ *es una norma.*[5] *Por tanto:*

$$\|\mathbf{R}^i - \mathbf{R}^j\|_1 \leq \|\mathbf{R}^i\|_1 + \|\mathbf{R}^j\|_1 = 2, \forall i,j \in S \Rightarrow \max_{i,j\in S}\{\|\mathbf{R}^i - \mathbf{R}^j\|_1\} \leq 2 \Rightarrow 2C \leq 2 \Rightarrow C \leq 1,$$

tal y como queríamos ver.

[4] Véase el Ejemplo A.4.2 del Apéndice A.

[5] Véase la Definición A.1.1 del Apéndice A.

ii) *Vamos a realizar la demostración por contrarrecíproco, o sea, veremos que si $C = 1$, entonces hay valores de $\mathbf{P}$ que son 0:*

Si $C = 1$ tenemos que $\max_{i,j\in S}\{\|\mathbf{R}^i - \mathbf{R}^j\|_1\} = 2$. Sean $i, j \in S$ las filas correspondientes al máximo y $\mathbf{R}^i = \mathbf{x}$ y $\mathbf{R}^j = \mathbf{y}$ los vectores estocásticos que componen dichas filas. Así pues, tenemos que $\|\mathbf{x} - \mathbf{y}\|_1 = 2$.

Y, por otro lado, como se cumple la desigualdad $\|\mathbf{x} - \mathbf{y}\|_1 \leq \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1$ y $\mathbf{x}, \mathbf{y} \in \mathcal{T}$, tenemos que:

$$2 = \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1 = \|\mathbf{x} - \mathbf{y}\|_1. \tag{B.8}$$

Recordemos que $\mathbf{x} = (x_1, \ldots, x_N)$ e $\mathbf{y} = (y_1, \ldots, y_N)$. Sean $A = \{i \in S \,/\, x_i > y_i\}$ y $B = S \setminus A$ dos conjuntos de índices.

Vamos a ver que ni A ni B pueden ser vacíos:

- *Supongamos que $A = \emptyset$. Por definición, $B = S$ y esto quiere decir que $y_i \geq x_i, \forall i \in S$. Supongamos ahora que existe algún $j \in S$ t.q. $y_j > x_j$. Esto significa que*

  $$\sum_{i\in S} y_i > \sum_{i\in S} x_i,$$

  pero por ser $\mathbf{x}, \mathbf{y} \in \mathcal{T}$ tenemos que

  $$\sum_{i\in S} y_i = \sum_{i\in S} x_i = 1,$$

  por lo que tendríamos que $1 < 1$, que es una contradicción. Entonces $y_i = x_i, \forall i \in S$. Pero, si este fuera el caso, tendríamos que $\|\mathbf{x} - \mathbf{y}\|_1 = 0$ que es otra contradicción, ya que $\|\mathbf{x} - \mathbf{y}\|_1 = 2$.

  Por tanto $A \neq \emptyset$.

- *Supongamos ahora que $B = \emptyset$:*

  $$y_i < x_i, \forall i \in S \Rightarrow 1 = \sum_{i\in S} y_i < \sum_{i\in S} x_i = 1 \Rightarrow 1 < 1,$$

  que sería de nuevo una contradicción.

  Por tanto, $B \neq \emptyset$.

Entonces $A, B \neq \emptyset$. Veamos ahora que hay componentes de $\mathbf{x}$ e $\mathbf{y}$ que valen 0 y habremos acabado:

$$\sum_{i\in A} x_i + \sum_{i\in B} x_i + \sum_{i\in A} y_i + \sum_{i\in B} y_i = \sum_{i\in S} x_i + \sum_{i\in S} y_i = 2 = \|\mathbf{x} - \mathbf{y}\|_1 = \sum_{i\in S} |x_i - y_i|. \tag{B.9}$$

Ahora, por cómo están definidos los conjuntos de índices tenemos que

$$\sum_{i\in S}|x_i-y_i|=\sum_{i\in A}(x_i-y_i)+\sum_{i\in B}(y_i-x_i),$$

debido a que $i\in A$ o $i\in B$, $\forall i\in S$. Entonces, si $i\in A$, tenemos que $x_i>y_i$ y por ello $|x_i-y_i|=(x_i-y_i)$. Por otro lado, si $i\in B$, se tiene que $y_i\geq x_i$ y por ello $|y_i-x_i|=(y_i-x_i)$.

Así pues, aplicando esto último a la Igualdad (B.9) anterior, nos queda que:

$$\sum_{i\in A}x_i+\sum_{i\in B}x_i+\sum_{i\in A}y_i+\sum_{i\in B}y_i=\sum_{i\in S}|x_i-y_i|=\sum_{i\in A}(x_i-y_i)+\sum_{i\in B}(y_i-x_i)\Rightarrow$$

$$\sum_{i\in A}x_i+\sum_{i\in B}x_i+\sum_{i\in A}y_i+\sum_{i\in B}y_i=\sum_{i\in A}x_i-\sum_{i\in A}y_i+\sum_{i\in B}y_i-\sum_{i\in B}x_i\Rightarrow 2\sum_{i\in A}y_i+2\sum_{i\in B}x_i=0\Rightarrow$$

$$\sum_{i\in A}y_i+\sum_{i\in B}x_i=0,$$

y, como $\mathbf{x},\mathbf{y}\in\mathscr{T}$, tenemos que $x_i\geq 0$ e $y_i\geq 0$, $\forall i\in S$, por lo que para que la suma anterior sea 0, todos los sumandos deben valer 0. Entonces $x_i=0$, $\forall i\in B$ e $y_i=0$, $\forall i\in A$. Esto prueba la existencia de valores de $\mathbf{P}$ que son iguales a 0.

En definitiva, hemos demostrado el apartado ii) por contrarrecíproco.

iii) *Sea $\mathbf{R}$ el conjunto formado por los vectores $\mathbf{R}^1,\dots,\mathbf{R}^N$ y $\delta=diam(\mathbf{K})$. Claramente, $\mathbf{R}\subset\mathscr{T}$ y, de manera análoga a como hemos demostrado que $\mathscr{T}$ era cerrado, podríamos demostrar que $\mathbf{R}$ es también cerrado. Ahora, como $\mathscr{T}$ es compacto y $\mathbf{R}\subset\mathscr{T}$ cerrado, tenemos que $\mathbf{R}$ es también compacto. Por tanto $\mathbf{K}$ es la envoltura convexa de un compacto, esto quiere decir que $\mathbf{K}$ es también compacto y por ello:*

$$\delta=diam(\mathbf{K})=\sup_{\mathbf{x},\mathbf{y}\in\mathbf{K}}\{\|\mathbf{x}-\mathbf{y}\|_1\}=\max_{\mathbf{x},\mathbf{y}\in\mathbf{K}}\{\|\mathbf{x}-\mathbf{y}\|_1\}=2C,$$

tal y como queríamos ver.

iv) *Vamos a demostrar que si $\mathbf{a}=(a_1,\dots,a_N)$ es t.q. $\sum_{i\in S}a_i=0$, entonces:*

$$\|\mathrm{P}(\mathbf{a})\|_1\leq C\|\mathbf{a}\|_1, \tag{B.10}$$

y así habremos demostrado el apartado iv). Debido a que si consideramos $\mathbf{x},\mathbf{y}\in\mathscr{T}$, entonces

$$1=\sum_{i=1}^N x_i=\sum_{i=1}^N y_i\Rightarrow\sum_{i=1}^N x_i-\sum_{i=1}^N y_i=0\Rightarrow\sum_{i=1}^N(x_i-y_i)=0,$$

y, si $a_i=x_i-y_i$, $\forall i\in S$, aplicamos la Desigualdad (B.10) y por ser P lineal ya tenemos demostrado el apartado iv).

Por tanto, vamos a demostrar primero la Desigualdad (B.10):

Asumamos, sin pérdida de generalidad, que $\|\mathbf{a}\|_1 = 2$ *(no está mal considerado ya que si luego queremos hacer el cambio* $\mathbf{a} = \mathbf{x} - \mathbf{y}$*, sabemos que* $\|\mathbf{x} - \mathbf{y}\|_1 \le \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1 = 2$ *por ser* $\mathbf{x}, \mathbf{y} \in \mathcal{T}$*). Para cualquier* $\mathbf{x} \in \mathbb{R}^N$*, definimos los conjuntos* $A = \{i \in S \,/\, x_i > 0\}$ *y* $B = \{i \in S \,/\, x_i \le 0\}$ *tt.q.:*

$$\sum_{i \in S} x_i = \sum_{i \in A} x_i + \sum_{i \in B} x_i.$$

Ahora,

$$\sum_{i \in A} a_i + \sum_{i \in B} a_i = \sum_{i \in S} a_i = 0,$$

por hipótesis, y

$$\|\mathbf{a}\|_1 = \sum_{i \in S} |a_i| = 2 \Rightarrow \sum_{i \in S} |a_i| = \sum_{i \in A} |a_i| + \sum_{i \in B} |a_i| = 2 \Rightarrow \sum_{i \in A} a_i - \sum_{i \in B} a_i = 2.$$

Sea $f = \sum_{i \in A} a_i$ *y* $g = \sum_{i \in B} a_i$*. Por las igualdades que acabamos de ver, se plantea el siguiente sistema a resolver:*

$$\begin{cases} f + g = 0 \\ f - g = 2 \end{cases},$$

donde la solución es $f = 1$ *y* $g = -1$*. Esto significa que*

$$\sum_{i \in A} a_i = -\sum_{i \in B} a_i = 1.$$

Definimos $\mathbf{x}' = \sum_{i \in A} a_i \mathbf{R}^i$ *y* $\mathbf{x}'' = -\sum_{i \in B} a_i \mathbf{R}^i$*, que son combinaciones convexas de los vectores de* $\mathbf{R}$*.* [6] *Como* $\mathbf{K}$ *es la envoltura convexa de* $\mathbf{R}$*, tenemos que* $\mathbf{x}', \mathbf{x}'' \in \mathbf{K}$*.*

Denotamos por $\sum_{i \in S} a_i \mathbf{R}^i$ *al vector que se obtiene de la siguiente forma:*

$$\mathbf{aP} = (\sum_{i \in S} a_i (R^i)_1, \ldots, \sum_{i \in S} a_i (R^i)_N).$$

Así pues:

$$\|\mathrm{P}(\mathbf{a})\|_1 = \|\sum_{i \in S} a_i \mathbf{R}^i\|_1 = \|\sum_{i \in A} a_i \mathbf{R}^i + \sum_{i \in B} a_i \mathbf{R}^i\|_1 = \|\mathbf{x}' - \mathbf{x}''\|_1,$$

como $\mathbf{x}', \mathbf{x}'' \in \mathbf{K}$*, por el apartado* $iii)$ *tenemos que* $\|\mathbf{x}' - \mathbf{x}''\|_1 \le 2C$*, por tanto:*

$$\|\mathrm{P}(\mathbf{a})\|_1 \le 2C = C\|\mathbf{a}\|_1,$$

[6] Debido a que $\sum_{i \in A} a_i = -\sum_{i \in B} a_i = 1$.

tal y como queríamos ver.

Notemos que si hubiéramos elegido $\|\mathbf{a}\|_1 = d$*, con* $d \in (0,2)$*, el resultado hubiera sido el mismo. Debido a que en el sistema de ecuaciones hubiéramos obtenido* $f = \frac{d}{2}$ *y* $g = -\frac{d}{2}$*, por ello* $\mathbf{x}' = \frac{2}{d}\sum_{i\in A} a_i\mathbf{R}^i$ *y* $\mathbf{x}'' = -\frac{2}{d}\sum_{i\in B} a_i\mathbf{R}^i$*, con* $\frac{2}{d}\sum_{i\in A} a_i = 1 = -\dfrac{2}{d}\sum_{i\in B} a_i$*. De esta forma,* $\mathbf{x}', \mathbf{x}'' \in \mathbf{K}$ *igual que antes, y:*

$$\|\sum_{i\in A} a_i\mathbf{R}^i + \sum_{i\in B} a_i\mathbf{R}^i\|_1 = \|\frac{d}{2}\mathbf{x}' - \frac{d}{2}\mathbf{x}''\|_1 = \frac{d}{2}\|\mathbf{x}' - \mathbf{x}''\|_1 \le \frac{d}{2}2C = dC = C\|\mathbf{a}\|_1.$$

En definitiva, al demostrar la Desigualdad (B.10), si elegimos $\mathbf{a} = \mathbf{x} - \mathbf{y}$*, con* $\mathbf{x}, \mathbf{y} \in \mathcal{T}$*, tenemos que:*

$$\|\mathrm{P}(\mathbf{x} - \mathbf{y})\|_1 \le C\|\mathbf{x} - \mathbf{y}\|_1,$$

y por ser P *una función lineal:* $\mathrm{P}(\mathbf{x} - \mathbf{y}) = \mathrm{P}(\mathbf{x}) - \mathrm{P}(\mathbf{y})$*. Por tanto:*

$$\|\mathrm{P}(\mathbf{x}) - \mathrm{P}(\mathbf{y})\|_1 \le C\|\mathbf{x} - \mathbf{y}\|_1.$$

□

Demostración B.2.3. *Vamos a realizar la demostración por reducción al absurdo. Supongamos que* $\lambda = 0$*. Sea* $j \in S$ *arbitrario pero fijado:*

Por ser **P** *irreducible, tenemos que para cada* $i \in S$ *existe un* $k_0 \in \{1, 2, \ldots, N-1\}$*, al que llamaremos* $k_0(i)$*, tal que* $p_{ij}^{(k_0(i))} > 0$*.*[7] *Ahora,*

$$p_{jj}^{(k+k_0(i))} = \sum_{a=1}^{N} p_{ja}^{(k)} p_{aj}^{(k_0(i))} \ge p_{ji}^{(k)} p_{ij}^{(k_0(i))},$$

para cada $i \in S$*. Por hipótesis, por la suposición que hemos hecho y por la desigualdad que acabamos de ver, tenemos lo siguiente:*

$$\lim_{n\to+\infty} \frac{1}{n}\left(\sum_{k=1}^{n} p_{ji}^{(k)}\right) p_{ij}^{(k_0(i))} \le \lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} p_{jj}^{(k+k_0(i))} = 0 \Rightarrow$$

$$\lim_{n\to+\infty} \frac{1}{n}\left(\sum_{k=1}^{n} p_{ji}^{(k)}\right) p_{ij}^{(k_0(i))} = 0,$$

y como $p_{ij}^{(k_0(i))} > 0$*:*

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

Pero para que el límite de esta suma sea 0*, cada sumando debe valer* 0 *ya que todos los sumandos son mayores o iguales a* 0*. Por ello tenemos que* $p_{ji}^{(k)} = 0$*. Esto se cumple para*

[7] Esto es debido a que por ser **P** irreducible, tenemos que $\mathbf{B}_j^i > 0$, $\forall i, j \in S$. Esto implica que para cada $i, j \in S$ fijados, existe al menos un $k_0 \in \{1, 2, \ldots, N-1\}$ tal que $p_{ij}^{(k_0)} > 0$, porque sino $\mathbf{B}_j^i$ no sería mayor que 0.

cada $i \in S$, por ello $p_{ji}^{(k)} = 0, \forall i \in S$. Pero esto es una contradicción, debido a que para un $j \in S$ fijado, sabemos que $\sum_{i=1}^{N} p_{ji}^{(k)} = 1$, por ser **P** *estocástica.*

Esta contradicción proviene de haber supuesto que $\lambda = 0$. Así que

$$\frac{1}{n}\sum_{k=1}^{n} p_{jj}^{(k)} \to \lambda \Rightarrow \lambda > 0, \forall j \in S,$$

y, en particular:

$$p_{jj}^{(n)} \to \lambda \Rightarrow \lambda > 0, \forall j \in S.$$

□

Demostración B.2.4. *Como ya sabemos que $\mathbf{W} = \mathbf{WP}$, basta que demostremos las igualdades $\mathbf{W} = (\mathbf{W}^2)$ y $\mathbf{WP} = \mathbf{PW}$:*

- *Veamos primero la igualdad $\mathbf{W} = (\mathbf{W}^2)$:*

$$(\mathbf{W}^2) = \begin{pmatrix} w_1 & \cdots & w_N \\ \vdots & \ddots & \vdots \\ w_1 & \cdots & w_N \end{pmatrix} \begin{pmatrix} w_1 & \cdots & w_N \\ \vdots & \ddots & \vdots \\ w_1 & \cdots & w_N \end{pmatrix} = \begin{pmatrix} w_1 \sum\limits_{i \in S} w_i & \cdots & w_N \sum\limits_{i \in S} w_i \\ \vdots & \ddots & \vdots \\ w_1 \sum\limits_{i \in S} w_i & \cdots & w_N \sum\limits_{i \in S} w_i \end{pmatrix}$$

  Ahora, por ser **w** *un vector estocástico, tenemos que $\sum_{i \in S} w_i = 1$. Por tanto:*

$$\begin{pmatrix} w_1 \sum\limits_{i \in S} w_i & \cdots & w_N \sum\limits_{i \in S} w_i \\ \vdots & \ddots & \vdots \\ w_1 \sum\limits_{i \in S} w_i & \cdots & w_N \sum\limits_{i \in S} w_i \end{pmatrix} = \begin{pmatrix} w_1 & \cdots & w_N \\ \vdots & \ddots & \vdots \\ w_1 & \cdots & w_N \end{pmatrix} = \mathbf{W}$$

  En definitiva, $(\mathbf{W}^2) = \mathbf{W}$ tal y como queríamos ver.

- *A continuación, demostraremos que $\mathbf{WP} = \mathbf{PW}$:*

  Sabemos que $\mathbf{WP} = \mathbf{W}$. Por ello bastará que veamos que $\mathbf{PW} = \mathbf{W}$ para que la igualdad quede demostrada:

$$\mathbf{PW} = \begin{pmatrix} p_{11} & \cdots & p_{1N} \\ \vdots & \ddots & \vdots \\ p_{N1} & \cdots & p_{NN} \end{pmatrix} \begin{pmatrix} w_1 & \cdots & w_N \\ \vdots & \ddots & \vdots \\ w_1 & \cdots & w_N \end{pmatrix} = \begin{pmatrix} w_1 \sum\limits_{i \in S} p_{1i} & \cdots & w_N \sum\limits_{i \in S} p_{1i} \\ \vdots & \ddots & \vdots \\ w_1 \sum\limits_{i \in S} p_{Ni} & \cdots & w_N \sum\limits_{i \in S} p_{Ni} \end{pmatrix}$$

  Por ser **P** *una matriz estocástica, tenemos que todas las filas de la matriz son vectores estocásticos y por ello $\sum_{i \in S} p_{ji} = 1, \forall j \in S$. Por tanto:*

$$\begin{pmatrix} w_1 \sum_{i\in S} p_{1i} & \cdots & w_N \sum_{i\in S} p_{1i} \\ \vdots & \ddots & \vdots \\ w_1 \sum_{i\in S} p_{Ni} & \cdots & w_N \sum_{i\in S} p_{Ni} \end{pmatrix} = \begin{pmatrix} w_1 & \cdots & w_N \\ \vdots & \ddots & \vdots \\ w_1 & \cdots & w_N \end{pmatrix} = \mathbf{W}$$

Esto quiere decir que $\mathbf{PW} = \mathbf{W} = \mathbf{WP}$.

La conclusión es:

$$\mathbf{W}^2 = \mathbf{W} = \mathbf{PW} = \mathbf{WP}.$$

□

Demostración B.2.5. *Primero veremos que* $\mathbf{v} = (1,\dots,1)$ *es un vector propio para la matriz* $\mathbf{P}$, *y* $\lambda_v = 1$ *el valor propio asociado:*

$$\mathbf{Pv}^t = \begin{pmatrix} p_{11} & \cdots & p_{1N} \\ \vdots & \ddots & \vdots \\ p_{N1} & \cdots & p_{NN} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{i\in S} p_{1i} \\ \vdots \\ \sum_{i\in S} p_{Ni} \end{pmatrix}.$$

Como $\mathbf{P}$ *es una matriz estocástica, tenemos que todas las filas de la matriz son vectores estocásticos y por ello* $\sum_{i\in S} p_{ji} = 1$, $\forall j \in S$. *Por tanto:*

$$\begin{pmatrix} \sum_{i\in S} p_{1i} \\ \vdots \\ \sum_{i\in S} p_{Ni} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \mathbf{v}^t.$$

Esto significa que $\mathbf{Pv}^t = \mathbf{v}^t$. *Así pues,* $\mathbf{v} = (1,\dots,1)$ *es un vector propio para la matriz* $\mathbf{P}$, *y* $\lambda_v = 1$ *el valor propio asociado.*

Ahora veremos que dado $\lambda \in \mathbb{C}$, *un valor propio asociado al vector propio*

$$\mathbf{z}^t = \begin{pmatrix} z^1 \\ \vdots \\ z^N \end{pmatrix} \neq \mathbf{0}^t$$

de $\mathbf{P}$, *se cumple que* $|\lambda| \le 1$:

Sea $i_0 \in S$ *t.q.* $|z^{i_0}| = \max_{i\in S}\{|z^i|\}$. *Notemos que* $|z^{i_0}| > 0$, *debido a que* $\mathbf{z} \neq \mathbf{0}$ *y* $|z^{i_0}|$ *es el módulo del máximo de los módulos de las componentes de* $\mathbf{z}$.

Entonces, por ser $\mathbf{z}$ *vector propio de* $\mathbf{P}$, *tenemos que:*

$$|\lambda||z^{i_0}| = |\lambda z^{i_0}| = |(\mathbf{Pz})^{i_0}| = |\sum_{j\in S} (\mathbf{P})^{i_0}_j z^j|.$$

Ahora, por ser $\mathbf{P}$ *estocástica, sabemos que* $\mathbf{P}^{i_0}_j \ge 0$, $\forall j \in S$. *Por ello* $|\mathbf{P}^{i_0}_j| = \mathbf{P}^{i_0}_j$, $\forall j \in S$. *Y, como* $|z + w| \le |z| + |w|$, $\forall z, w \in \mathbb{C}$,[8] *nos queda que:*

[8] Consúltese la Definición A.6.2 para más información sobre las propiedades de este conjunto.

$$|\sum_{j\in S}\mathbf{P}^{i_0}_j z^j| \leq \sum_{j\in S}\mathbf{P}^{i_0}_j |z^j|,$$

y por ser $|z^{i_0}| = \max_{i\in S}\{|z^i|\}$*:*

$$\sum_{j\in S}\mathbf{P}^{i_0}_j |z^j| \leq |z^{i_0}|\sum_{j\in S}\mathbf{P}^{i_0}_j,$$

como $\mathbf{P}$ *es estocástica:* $\sum_{j\in S}\mathbf{P}^{i_0}_j = 1$*. Por lo que:*

$$|z^{i_0}|\sum_{j\in S}\mathbf{P}^{i_0}_j = |z^{i_0}|.$$

En definitiva, $|\lambda||z^{i_0}| \leq |z^{i_0}|$*. Y, al ser* $|z^{i_0}| > 0$*, resulta que*

$$|\lambda| \leq \frac{|z^{i_0}|}{|z^{i_0}|} \Rightarrow |\lambda| \leq 1.$$

□

Demostración B.2.6. *Por ser* $\mathbf{P}$ *una matriz regular,* $\exists k_0 \in \mathbb{N}$ *t.q.* $(\mathbf{P}^{k_0})^i_j > 0$*,* $\forall i, j \in S$*. Sea* $\lambda \in \mathbb{C}$ *t.q.* $|\lambda| = 1$ *es un valor propio para* $\mathbf{P}$*. Queremos ver que* $\lambda = 1$*.*

- *Supongamos, primero de todo, que no solo* $(\mathbf{P}^{k_0})$ *tiene todas sus componentes estrictamente positivas, sino que* $\mathbf{P}$ *también las tiene.*

  Sea $\mathbf{x}$ *el vector propio asociado a* λ *(así que* $\mathbf{x} \neq \mathbf{0}$*) y* $i_0 \in S$ *t.q.* $|x^{i_0}| = \max_{i\in S}\{|x^i|\}$*.*

  Entonces $|x^{i_0}| > 0$*.*

  La fila i_0*-ésima de* $\mathbf{P}\mathbf{x}^t = \lambda\mathbf{x}^t$[9] *sería:*

  $$p_{i_0 1}x^1 + \cdots p_{i_0 N}x^N = \lambda x_{i_0},$$

  con $\sum_{j\in S} p_{i_0 j} = 1$ *por ser* $\mathbf{P}$ *estocástica. Al suponer que* $\mathbf{P}$ *tiene todas sus componentes positivas, tenemos que* $p_{i_0 1}, \ldots p_{i_0 N} > 0$*. Por otro lado,*

  $$|\sum_{j\in S} p_{i_0 j}x^j| = |p_{i_0 1}x^1 + \cdots p_{i_0 N}x^N| = |\lambda x^{i_0}| = |\lambda||x^{i_0}| = |x^{i_0}| > 0,$$

  y, por cómo hemos elegido x^{i_0}*:* $|x^1|, \ldots, |x^N| \leq |x^{i_0}|$*. Por tanto, cumplimos las hipótesis para poder aplicar la Proposición A.12.1, y así resulta que* $x^1 = \cdots = x^N$*. Esto quiere decir que:*

  $$x^{i_0} = x^{i_0}\sum_{j\in S} p_{i_0 j} = \sum_{j\in S} p_{i_0 j}x^j = \lambda x^{i_0} \Rightarrow \lambda = 1,$$

  ya que $|x^{i_0}| > 0$*.*

[9] Esta igualdad se cumple por ser λ el valor propio de $\mathbf{P}$ y $\mathbf{x}$ su vector propio asociado.

- *Veamos ahora el caso general, o sea,* $\mathbf{P}$ *no tiene porque tener todos sus elementos positivos.*

  Pero, por ser regular, ya hemos dicho al principio de esta demostración que $\exists k_0 \in \mathbb{N}$ *t.q.* $(\mathbf{P}^{k_0})^i_j > 0, \forall i, j \in S$. *Además, por ser* $\mathbf{x}$ *vector propio de* $\mathbf{P}$ *con valor propio asociado* λ*, tenemos que:*

$$(\mathbf{P}^{k_0})\mathbf{x}^t = (\mathbf{P}^{k_0-1})\mathbf{P}\mathbf{x}^t = (\mathbf{P}^{k_0-1})\lambda\mathbf{x}^t = \lambda(\mathbf{P}^{k_0-1})\mathbf{x}^t = \lambda(\mathbf{P}^{k_0-2})\mathbf{P}\mathbf{x}^t = \lambda(\mathbf{P}^{k_0-2})\lambda\mathbf{x}^t =$$

$$\lambda^2(\mathbf{P}^{k_0-2})\mathbf{x}^t = \cdots = \lambda^{k_0}\mathbf{x}^t.$$

  Con un argumento similar al que hemos hecho para el caso particular anterior, obtenemos, de nuevo, que $x^1 = \cdots = x^N$ *aplicando la Proposición A.12.1. Además,* $\lambda^{k_0} = 1$.

  Pero, únicamente con este resultado, no podemos asegurar que $\lambda = 1$. *Debido a que si por ejemplo* $k_0 = 4$ *y* $\lambda = i$,[10] *entonces* $\lambda^{k_0} = 1$ *y* $\lambda \neq 1$.

  Por lo que nos hace falta algún resultado más para asegurar que $\lambda = 1$.

  Sabemos que $\mathbf{P}$ *es regular, esto quiere decir que no tiene ninguna columna de* 0*'s. Veamos por qué:*

  Supongamos que la columna $j_0 \in S$ *es t.q.* $p_{i\,j_0} = 0, \forall i \in S$. *Sea* $i_1 \in S$ *arbitrario pero fijado. Entonces:*

$$(\mathbf{P}^{k_0})^{i_0}_{j_0} = \sum_{k=1}^{N}(\mathbf{P}^{k_0-1})^{i_0}_k\mathbf{P}^k_{j_0} = \sum_{k=1}^{N}(\mathbf{P}^{k_0-1})^{i_0}_k \cdot 0 = 0.$$

  Por tanto, $(\mathbf{P}^{k_0})^{i_0}_{j_0} = 0$ *que es una contradicción. Debido a que por ser* $\mathbf{P}$ *regular, sabíamos que* $(\mathbf{P}^{k_0})$ *tiene todos sus elementos positivos.*

  Entonces $\mathbf{P}$ *no tiene ninguna columna toda de* 0*'s, y, por ello,* $(\mathbf{P}^{k_0+1})$ *también tiene todos sus elementos positivos. Como también se cumple que* $(\mathbf{P}^{k_0+1})\mathbf{x}^t = \lambda^{k_0+1}\mathbf{x}^t$, *con un argumento similar al que hemos hecho en el caso particular, acabaríamos obteniendo que* $\lambda^{k_0+1} = 1$.

  Ahora sí, como $\lambda^{k_0} = 1$, $\lambda^{k_0+1} = 1$ *y* $|\lambda| = 1 \Rightarrow |\lambda^{k_0}| = 1$ *(por hipótesis), tenemos que:*

$$\lambda^{k_0} = \lambda^{k_0+1} \Rightarrow \frac{\lambda^{k_0}}{\lambda^{k_0}} = \lambda \Rightarrow 1 = \lambda,$$

  tal y como queríamos ver.

□

[10]Véase la Definición A.6.1 del Apéndice A para más información sobre la unidad imaginaria.

Demostración B.2.7. *Sea $(\mathbf{J}_k)$ uno de los bloques de Jordan de la matriz $\mathbf{J}$ con $k \in \{1,\dots,p\}$ arbitrario pero fijado. Entonces*

$$(\mathbf{J}_k) = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda_k & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_k \end{pmatrix}_{n_k \times n_k},$$

donde n_k es la multiplicidad algebraica del valor propio λ_k. Sabemos que $|\lambda_k| < 1$ por el Teorema 3.3.1.

Calculemos $((\mathbf{J}_k)^2)$ y $((\mathbf{J}_k)^3)$ para ver qué podría valer $((\mathbf{J}_k)^n)$:

$$((\mathbf{J}_k)^2) = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda_k & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_k \end{pmatrix} \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda_k & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_k \end{pmatrix} = \begin{pmatrix} (\lambda_k)^2 & 2\lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 1 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 2\lambda_k \\ 0 & \cdots & \cdots & \cdots & 0 & (\lambda_k)^2 \end{pmatrix}$$

$$((\mathbf{J}_k)^3) = \begin{pmatrix} (\lambda_k)^2 & 2\lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 1 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 2\lambda_k \\ 0 & \cdots & \cdots & \cdots & 0 & (\lambda_k)^2 \end{pmatrix} \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \lambda_k & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_k \end{pmatrix} =$$

$$\begin{pmatrix} (\lambda_k)^3 & 3(\lambda_k)^2 & 3\lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 1 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 3\lambda_k \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 3(\lambda_k)^2 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & (\lambda_k)^3 \end{pmatrix}.$$

Si nos fijamos, los elementos que van surgiendo en el triángulo superior de $((\mathbf{J}_k)^n)$ a medida que n aumenta, se corresponden con las diferentes filas del triángulo de Pascal.[11] *De hecho, en general, tenemos que $\forall\, j \in \{1,\dots,n_k\}$ e $i \in \{1,\dots,n_k\}$:*

[11] Véase la Definición A.8.1 del Apéndice A para más información.

$$((\mathbf{J}_k)^n)^i_j = \begin{cases} \begin{pmatrix} n \\ j-i \end{pmatrix} (\lambda_k)^{n-(j-i)} & \textit{si } i \leq j \leq n+1 \\ 0 & \textit{en otro caso} \end{cases}$$

Ahora bien, como $|\lambda_k| < 1$ *tenemos que:*

$$\lim_{n\to+\infty} (\lambda_k)^n = 0 \Rightarrow \lim_{n\to+\infty} a \cdot (\lambda_k)^n = 0, \ \textit{con } a = \begin{pmatrix} n \\ j-i \end{pmatrix}.$$

Notemos que cuando $n \to +\infty$, $\begin{pmatrix} n \\ j-i \end{pmatrix}$ *puede hacerse muy grande, pero como* $(\lambda_k)^n$ *converge mucho más rápido a* 0*, que* a *a* $+\infty$*, tenemos que* $a \cdot (\lambda_k)^n$ *converge a* 0*.*

Como consecuencia de esto, cuando $n \to +\infty$*, obtenemos que* $((\mathbf{J}_k)^n)$ *converge a la matriz de* 0*'s de dimensión* $n_k \times n_k$*.*

Como esto lo hemos visto para un $k \in \{1, \ldots, p\}$ *arbitrario pero fijado, se cumple para cualquier* k *y, por tanto:*

$$\lim_{n\to+\infty} (\mathbf{J}^n) = \mathbf{K}.$$

□

Demostración B.2.8. *Vamos a ver si podemos acotar* $\|(\mathbf{J}^n) - \mathbf{K}\|$*:*

$$\|\mathbf{J}^n - \mathbf{K}\| = \max_{i\in S}\{\|(\mathbf{J}^n)^i - \mathbf{K}^i\|_1\} = \max_{i\in S}\{\sum_{j\in S} |(\mathbf{J}^n)^i_j - \mathbf{K}^i_j|\},$$

ahora, como la primera fila de $(\mathbf{J}^n)$ *y de* $\mathbf{K}$ *es la misma, podemos considerar desde* $i = 2$*. Como a partir de la fila* 2*, la matriz* $\mathbf{K}$ *es toda de* 0*'s, la igualdad anterior nos queda:*

$$\max_{i\in S}\{\sum_{j\in S} |(\mathbf{J}^n)^i_j - \mathbf{K}^i_j|\} = \max_{i=2,\ldots,N}\{\sum_{j\in S} |(\mathbf{J}^n)^i_j|\}.$$

Sea λ_k*, con* $k \in \{1, \ldots, p\}$*, el valor propio de* $\mathbf{P}$*, de entre todos los valores propios, con mayor módulo. Entonces:*

$$\max_{i=2,\ldots,N}\{\sum_{j\in S} |(\mathbf{J}^n)^i_j|\} = \max_{i=1,\ldots,n_k}\{\sum_{j=1}^{n_k} |((\mathbf{J}_k)^n)^i_j|\} = \max_{i=1,\ldots,n_k}\{\sum_{j=1}^{n_k} |((\lambda_k Id + \mathbf{Q})^n)^i_j|\},$$

donde $\mathbf{Q}$ *es la matriz nilpotente:*

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

En la matriz $\lambda_k Id + \mathbf{Q}$*, la primera fila es la que tendrá, de entre todas las filas, suma mayor. Esto es debido a que esta matriz es triangular superior con el mismo valor en cada* *una de las posiciones de la diagonal, y la secuencia de valores, para cada fila, se repite desde el valor de la diagonal, hasta la última columna. Además, todos los valores son mayores o iguales a* 0*, por ello, como acabamos de explicar:*

$$\sum_{j=1}^{n_k}((\mathbf{J}_k)^n)_j^{n_k} \leq \sum_{j=1}^{n_k}((\mathbf{J}_k)^n)_j^{n_k-1} \leq \cdots \leq \sum_{j=1}^{n_k}((\mathbf{J}_k)^n)_j^1.$$

Por tanto:

$$\max_{i=1,\ldots,n_k}\{\sum_{j=1}^{n_k}|((\lambda_k Id+\mathbf{Q})^n)_j^i|\} = \sum_{j=1}^{n_k}((\lambda_k Id+\mathbf{Q})^n)_j^1,$$

y, como ya hemos visto en la Demostración B.2.7 anterior, tenemos que:

$$\sum_{j=1}^{n_k}((\lambda_k Id+\mathbf{Q})^n)_j^1 = \begin{cases} \displaystyle\sum_{j=1}^{n_k}\binom{n}{j-1}(\lambda_k)^{n-(j-1)} & si\ n_k \leq n \\ \displaystyle\sum_{j=0}^{n}\binom{n}{j}(\lambda_k)^{n-j} & si\ n_k > n \end{cases} =$$

$$\begin{cases} C_\epsilon(\lambda_k+\epsilon)^n & si\ n_k \leq n \\ (\lambda_k+1)^n & si\ n_k > n \end{cases}$$

donde $\epsilon > 0$ *y* C_ϵ *(constante que depende de* ϵ*) surgen de que en el caso* $n_k \leq n$*, no tenemos todos los elementos del binomio* $(\lambda_k+1)^n$*, solo tenemos hasta el término* n_k*. Como* $|\lambda_k| < 1$*, podemos acotar el resultado de este binomio con* ϵ*, que dependerá del valor de* n_k *obviamente, y la constante* C_ϵ*.*

Este cálculo lo estamos haciendo para averiguar la velocidad de convergencia de $(\mathbf{J}^n)$ *hacia* $\mathbf{K}$*, por lo que, a priori, el* n *será lo suficientemente grande como para únicamente considerar el caso* $n_k \leq n$*. En definitiva:*

$$\|(\mathbf{J}^n)-\mathbf{K}\| \leq C_\epsilon(\lambda_k+\epsilon)^n.$$

De aquí deducimos que la velocidad de convergencia es exponencial. Al haber visto esto, también tendríamos que:

$$\|(\mathbf{P}^n)-\mathbf{W}\| \leq C_\epsilon(\lambda_k+\epsilon)^n.$$

□

APÉNDICE C

CÓDIGOS DE LOS PROGRAMAS UTILIZADOS

Como herramienta de cálculo se ha utilizado el programa *Octave*. Este programa es muy práctico si trabajamos con vectores y matrices.

Lo hemos usado principalmente para calcular potencias de matrices, inversas, determinantes... y para averiguar los resultados de aplicar el método de *Gauss* en el Capítulo 4 del trabajo.

A continuación, mostramos los códigos para encontrar el vector solución de un sistema compatible determinado mediante el método de *Gauss*:

```
function y = Gauss(A,b)
  p = 0;
  q = 0;
  n = length(b);
  x = zeros(n,1);
  for i=1:n-1
    if A(i,i)!=0
      for j=i+1:n
        p = A(j,i)/A(i,i);
        A(j,i:n) = A(j,i:n)-p*A(i,i:n);
        b(j) = b(j)-p*b(i);
      endfor
    else
      printf("Error, ");
      A(i,i);
      printf(" vale 0.");
      return
    endif
  endfor
```

```
  for i=n:-1:1
    if i!=n
      for j=i+1:n
        q = q + x(j)*A(i,j);
      endfor
      x(i) = (b(i)-q)/A(i,i);
      q = 0;
    else
      x(i)=b(i)/A(i,i);
    endif
  endfor
  y = x;
endfunction
```

Y para encontrar la matriz del sistema compatible indeterminado aplicando el mismo método:

```
function y = Gauss2(A,b)
  p = 0;
  n = length(b);
  for i=1:n-1
    if A(i,i)!=0
      for j=i+1:n
        p = A(j,i)/A(i,i);
        A(j,i:n) = A(j,i:n)-p*A(i,i:n);
        b(j) = b(j)-p*b(i);
      endfor
    else
      printf("Error, ");
      A(i,i);
      printf(" vale 0.");
      return
    endif
  endfor
  b
  y = A;
endfunction
```

BIBLIOGRAFÍA

[1] G. Modica and L. Poggiolini, *A First Course in Probability and Markov Chains*. Chichester: John Wiley & Sons Ltd, 2012. 1

[2] M. Olinick, *Mathematical Modeling in the Social and Life Sciences*. John Wiley & Sons Ltd, 2014. 1

[3] J. C. Arya and L. Robin W., *Matemáticas Aplicadas a la Administración y a la Economía*. México: Pearson Educación, 2009. 1

[4] E. Behrends, *Introduction to Markov Chains*, ser. Advanced Lectures in Mathematics. Wiesbaden: Vieweg+Teubner Verlag, 2000. [Online]. Available: http://link.springer.com/10.1007/978-3-322-90157-6 1

[5] J. I. Gallardo, "Periódico Deportivo Online: http://www.marca.com/estadisticas/futbol/primera/2016_17/," 1938. [Online]. Available: http://www.marca.com/estadisticas/futbol/primera/2016{_}17/ 2, 9, 10

[6] E. Miller, *Getting Started in Hold'em*. Two Plus Two Publishing LLC, 2005. 4.2.4

[7] G. Mayor Forteza, "Apuntes de la asignatura (20311) - Àlgebra Lineal II," Palma de Mallorca, 2011. 5

[8] A. E. Teruel Aguilar, "Apuntes de la asignatura (20316) - Mètodes Numèrics I," Palma de Mallorca, 2011. 5

[9] C. Sbert Juan and A. Buades Capó, "Apuntes de la asignatura (20332) - Mètodes Numèrics II," Palma de Mallorca, 2015. 5

[10] M. González Hidalgo and J. Torrens Sastre, "Apuntes de la asignatura (20308) - Introducció a l'Anàlisi Matemàtica," Palma de Mallorca, 2010. 5

[11] M. d. l. M. Llabrés Segura and G. Cardona Juanals, "Apuntes de la asignatura (20300) - Matemàtica Discreta," Palma de Mallorca, 2010. 5

[12] J. Suñer Llabrés, "Apuntes de la asignatura (20320) - Probabilitats II," Palma de Mallorca, 2015. 5

[13] M. Monserrat, "Apuntes de la asignatura (20313) - Topologia," Palma de Mallorca, 2011. 5

[14] J. Torrens Sastre, "Apuntes de la asignatura (20307) - Matemàtiques I (Àlgebra Lineal)," Palma de Mallorca, 2010. 5

[15] J. Wales, "Enciclopedia Online: https://www.wikipedia.org/," 2001. [Online]. Available: https://www.wikipedia.org/ 5