

3 Cadenas de Markov

3.1 Cadenas de Markov: Introducción

3.1.1 Concepto de cadena de Markov

Las cadenas de Markov son una herramienta para analizar el comportamiento y el gobierno de determinados tipos de procesos estocásticos, esto es, procesos que evolucionan de forma no determinista a lo largo del tiempo en torno a un conjunto de estados.

Una cadena de Markov, por tanto, representa un sistema que varía su estado a lo largo del tiempo, siendo cada cambio una transición del sistema. Dichos cambios no están predeterminados, aunque sí lo está la probabilidad del próximo estado en función de los estados anteriores, probabilidad que es constante a lo largo del tiempo (sistema homogéneo en el tiempo). Eventualmente, en una transición, el nuevo estado puede ser el mismo que el anterior y es posible que exista la posibilidad de influir en las probabilidades de transición actuando adecuadamente sobre el sistema (decisión).

En este módulo nos ocuparemos de las llamadas cadenas de Markov finitas, caracterizadas porque el número de estados del sistema es finito.

Formalmente, para definir una cadena de Markov finita hace falta determinar por lo tanto los siguientes elementos:

- a) Un conjunto de *estados* del sistema.
- b) La definición de *transición*.
- c) Una *ley de probabilidad condicional*, que defina la probabilidad del nuevo estado en función de los anteriores.

Los *estados* son una caracterización de la situación en que se halla el sistema en un instante dado, dicha caracterización puede ser tanto cuantitativa como cualitativa. Desde un punto de vista práctico probablemente, la mejor definición de qué debe entenderse por estado es la respuesta que se daría a la pregunta "¿cómo están las cosas?".

Formalmente, el estado de un sistema en un instante t es una variable cuyos valores sólo pueden pertenecer al conjunto de estados del sistema. El sistema modelizado por la cadena, por lo tanto, es una variable que cambia de valor en el tiempo, cambio al que llamamos *transición*.

Dicho de otro modo, se trata de una colección indexada de variables E_t , donde t denota intervalos temporales significativos para el fenómeno estudiado. Los posibles valores de E_t se toman de un

conjunto de categorías mutuamente excluyentes, denominadas *estados* del sistema. Por ser el sistema estocástico, no se conocerá con certeza el estado del sistema en un determinado instante, sino tan solo la probabilidad asociada a cada uno de los estados. Este hecho puede expresarse en términos de probabilidad condicional:

$$p\{E_t = j \mid E_{t-1} = i, E_{t-2} = e_{t-2}, E_{t-3} = e_{t-3}, \dots, E_0 = e_0\} = p\{E_t = j \mid E_{t-1} = i, E_{t-2} = e_{t-2}, E_{t-3} = e_{t-3}, \dots, E_{t-k} = e_{t-k}\}$$

donde i, j, e_i pertenecen al conjunto de estados posibles del sistema.

Para el caso particular de una cadena de Markov de orden 1, tenemos:

$$p\{E_t = j \mid E_{t-1} = i, E_{t-2} = e_{t-2}, E_{t-3} = e_{t-3}, \dots, E_0 = e_0\} = p\{E_t = j \mid E_{t-1} = i\} = p_{ij}$$

Donde p_{ij} recibe el nombre de probabilidad de transición del estado i al estado j . En una cadena de Markov de orden 1, el estado del sistema en el futuro j sólo depende del estado presente i .

En este módulo sólo estudiaremos cadenas de Markov que tengan las propiedades siguientes:

- Con un número n finito de estados (en oposición a cadenas de Markov de infinitos estados posibles).
- De orden 1 (si bien veremos cómo transformar las de orden superior en cadenas de orden 1).

Para este tipo de cadenas, tendremos que la ley de probabilidad condicional es de la forma:

$$p\{E_{t+s} = j \mid E_{t+s-1} = i\} = p\{E_t = j \mid E_{t-1} = i\}$$

3.1.2 Matrices de probabilidades de transición

En las cadenas finitas de orden 1 (es decir las que estudiamos aquí), la forma más cómoda de expresar la ley de probabilidad condicional de la misma es mediante la llamada matriz de probabilidades de transición \mathbf{P} , o más sencillamente, matriz de la cadena.

Dicha matriz es cuadrada con tantas filas y columnas como estados tiene el sistema, y los elementos de la matriz representan la probabilidad de que el estado próximo sea el correspondiente a la columna si el estado actual es el correspondiente a la fila.

Como el sistema debe evolucionar a t a alguno de los n estados posibles, las probabilidades de transición cumplirán la propiedad siguiente:

$$\sum_{j=1}^n p_{ij} = 1$$

Además, por definición de probabilidad, cada una de ellas ha de ser no negativa:

$$p_{ij} \geq 0$$

Cuando las p_{ij} cumplen las propiedades arriba indicadas, la matriz \mathbf{P} es una matriz estocástica: la suma de valores de las filas de la matriz será siempre igual a 1 (la suma de valores de las columnas no tiene ninguna propiedad especial).

Para más adelante, es interesante recordar dos propiedades de este tipo de matrices:

- Si \mathbf{P} , \mathbf{Q} son matrices estocásticas, su producto $\mathbf{P} \cdot \mathbf{Q}$ también lo es.
- Los valores propios de una matriz estocástica tienen siempre módulo inferior o igual a 1.

3.1.3 Matrices de orden superior

La condición antes expresada de que únicamente se tratarán en este módulo las matrices de orden 1 puede parecer una limitación a las posibilidades de modelización de procesos estocásticos, pero en realidad no lo es, ya que es posible estudiar las cadenas de orden superior ($k > 1$) mediante las de orden 1, transformándolas adecuadamente.

Dada una cadena de orden k con n estados, se puede definir una cadena equivalente de orden 1 de la siguiente forma:

Los estados de la cadena de orden 1 se definen como los diferentes conjuntos (ordenados) en los que puede estar la cadena superior en las k últimas transiciones. Así los estados de la cadena de orden 1 se definirán por:

$$X_t = \{E_t, E_{t-1}, E_{t-2}, \dots, E_{t-k}\}$$

De esta forma, la nueva cadena tiene $x = n^k$ estados.

Naturalmente, no todas las transiciones serán posibles en la nueva cadena: solamente aquellas en que los estados correspondientes para la cadena superior sean coherentes con la definición anterior. Dicho de otro modo:

Si $X_t = \{E_t, E_{t-1}, E_{t-2}, \dots, E_{t-k}\}$ y $X'_{t-1} = \{E'_{t-1}, E'_{t-2}, \dots, E'_{t-k-1}\}$ la probabilidad de transición $P'\{X_t | X'_{t-1}\}$ podrá ser distinta de cero si y sólo si $E_{t-1} = E'_{t-1}, E_{t-2} = E'_{t-2}$, etc.

y en este caso:

$$P'\{X_t | X'_{t-1}\} = P\{E_t | E_{t-1}, E_{t-2}, \dots, E_{t-k}, E_{t-k-1}\}$$

Ejemplo 3.1.3.a: cadena de Markov de orden 2

Después de analizar en profundidad la evolución de la cotización de un valor bursátil de renta variable, un inversor ha llegado a la conclusión de que la probabilidad de que dicho valor suba su cotización (estado A), la mantenga (estado B) o baje (estado C) durante una semana de cotización depende de su comportamiento en las dos semanas anteriores. Las probabilidades se indican en la siguiente tabla:

E_{t-1}	E_t	$p(A)$	$p(B)$	$p(C)$
A	A	0,7	0,1	0,2
A	B	0,3	0,5	0,2
A	C	0,5	0,3	0,2
B	A	0,4	0,1	0,5
B	B	0,2	0,7	0,1
B	C	0,1	0,3	0,6
C	A	0,3	0,3	0,4
C	B	0,3	0,1	0,6
C	C	0,2	0,1	0,7

Atendiendo a las definiciones que hemos indicado antes, nos encontramos ante una cadena de Markov de $n = 3$ estados y de orden $k = 2$. Pero, definiendo los estados como (E_{t-1}, E_{t-2}) podemos convertir la cadena de orden 2 en cadena de orden 1. Contaremos con $3^2 = 9$ estados posibles. Como cada estado puede evolucionar sólo a otros 3 en el instante siguiente, tendremos 27 transiciones posibles. En definitiva, tendremos la matriz de transición que se muestra a continuación, en la que los orígenes son (E_{t-1}, E_{t-2}) y los destinos (E_t, E_{t-1}) . En cada celda, se ha indicado el valor de E_t , y la probabilidad correspondiente.

	AA	AB	AC	BA	BB	BC	CA	CB	CC
AA	A/0,7			B/0,1			C/0,2		
AB	A/0,3			B/0,5			C/0,2		
AC	A/0,5			B/0,3			C/0,2		
BA		A/0,4			B/0,1			C/0,5	
BB		A/0,2			B/0,7			C/0,1	
BC		A/0,1			B/0,3			C/0,6	
CA			A/0,3			B/0,3			C/0,4
CB			A/0,3			B/0,1			C/0,6
CC			A/0,2			B/0,1			C/0,7

3.1.4 Probabilidades de transición de k pasos. Teorema de Chapman – Kolmogorov

Puesto que las probabilidades de transición son estables en el tiempo, podemos interesarnos en conocer las propiedades de transición después de k pasos, definidas formalmente como:

$$p\{E_{t+k} = j \mid E_t = i\} = p\{E_k = j \mid E_0 = i\} = p_{ij}^{(k)}$$

Esto es, la probabilidad de que el proceso se encuentre en el estado j si k etapas antes se encontraba en el estado i .

Si conocemos las p_{ij} , podemos calcular las $p_{ij}^{(k)}$ haciendo el siguiente razonamiento: si al cabo de $m < k$ pasos, nos encontramos en el estado e , la probabilidad de alcanzar el estado j después de $k - m$ pasos será:

$$p_{ie}^{(m)} \cdot p_{ej}^{(k-m)}$$

Como el estado intermedio e puede ser cualquiera, podemos determinar una expresión para la probabilidad de transición de k pasos:

$$p_{ij}^{(k)} = \sum_{e=1}^n p_{ie}^{(m)} \cdot p_{ej}^{(k-m)}$$

Haciendo $m = 1$, y $m = k-1$ obtenemos las ecuaciones de Chapman – Kolmogorov, que permiten obtener las expresiones de las propiedades de transición en el estado k a partir de las de $k-1$.

$$p_{ij}^{(k)} = \sum_{e=1}^n p_{ie} \cdot p_{ej}^{(k-1)}$$

$$p_{ij}^{(k)} = \sum_{e=1}^n p_{ie}^{(k-1)} \cdot p_{ej}$$

Lo que indican las ecuaciones es que pueden obtenerse las matrices $\mathbf{P}^{(k)}$ de transición de k pasos a partir de las potencias de la matriz \mathbf{P} .

$$\mathbf{P}^{(2)} = \mathbf{P} \cdot \mathbf{P} = \mathbf{P}^2$$

$$\mathbf{P}^{(3)} = \mathbf{P}^{(2)} \cdot \mathbf{P} = \mathbf{P} \cdot \mathbf{P}^2 = \mathbf{P}^2 \cdot \mathbf{P} = \mathbf{P}^3$$

$$\mathbf{P}^{(k)} = \mathbf{P}^{(k-1)} \cdot \mathbf{P} = \mathbf{P} \cdot \mathbf{P}^{k-1} = \mathbf{P}^{k-1} \cdot \mathbf{P} = \mathbf{P}^k$$

Es decir, que las sucesivas potencias de la matriz \mathbf{P} indican las probabilidades de transición en tantas transiciones como se indica en el índice de la potencia. Esto puede generalizarse aún más observando que la \mathbf{P}^1 representa la probabilidad de una transición y que $\mathbf{P}^0 = \mathbf{I}$ es la probabilidad en cero transiciones: si no ha habido transición, el estado es el mismo y por lo tanto la matriz que representa la no-transición es la matriz identidad.

3.1.5 Probabilidades estacionarias y tipos de cadenas de Markov

Nos interesamos ahora por conocer la probabilidad de hallar el sistema en un estado determinado cuando lleva funcionando un tiempo indefinidamente largo. Tales probabilidades se denominan *probabilidades estacionarias*. El estudio de las probabilidades estacionarias puede entenderse, por lo tanto, como el estudio del comportamiento a largo plazo de las cadenas de Markov.

Dichas probabilidades se denotan como π_{ij} , y la matriz de probabilidades de estado estable como \mathbf{P}^* .

Puesto que, como hemos visto, las potencias de \mathbf{P} definen las probabilidades en un número cualquiera de transiciones, podemos aproximarnos al estudio de las probabilidades estacionarias viendo qué ocurre si calculamos potencias elevadas de algunas matrices \mathbf{P} . Veremos que el comportamiento del sistema puede tender a estabilizarse, en el sentido de que el sistema pasará por cada uno de los estados con una frecuencia regular, esto es, independiente del número de transiciones transcurridas.

3.1.5.1 Cadenas ergódicas

La cadena de Markov C1, de dos estados, tiene la matriz de probabilidades de transición:

$$\mathbf{P}_1 = \begin{vmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{vmatrix}$$

Calculemos la potencia decimosexta de esa matriz para aproximar la matriz de probabilidades estacionarias:

$$\mathbf{P}_1^{16} = \begin{vmatrix} 0,429 & 0,571 \\ 0,429 & 0,571 \end{vmatrix}$$

Se observa que las probabilidades de estado estable de los diferentes estados son independientes del estado de origen, razón por la que la matriz de probabilidades estacionarias tiene todas las filas

iguales. Tenemos entonces una cadena de Markov *regular*, en la que las probabilidades estacionarias no dependen del estado inicial. Además, ninguna de las probabilidades vale cero. Tenemos entonces una cadena de Markov *ergódica*.

3.1.5.2 Cadenas semi-ergódicas

Tenemos ahora una cadena C2 de cuatro estados, de matriz de probabilidades de transición.

$$\mathbf{P}_2 = \begin{vmatrix} 0,5 & 0,4 & 0,1 & 0 \\ 0,3 & 0,3 & 0,4 & 0 \\ 0 & 0 & 0,2 & 0,8 \\ 0 & 0 & 0,6 & 0,4 \end{vmatrix}$$

Si se observa la matriz de la transición decimosesta, se observa cómo todas las filas tienden a ser iguales (aunque no completamente, especialmente las dos primeras), con una diferencia respecto de las cadenas ergódicas: existen estados cuya probabilidad de estado estable tiende a ser cero (esto es, que no aparecerán en el comportamiento a largo plazo). Por lo tanto, no se trata de una cadena ergódica. Sin embargo, sigue siendo cierto que todas las filas tienden hacia un mismo valor, por lo que sigue siendo regular. Las cadenas de Markov regulares (y también otras que veremos más adelante) con algunas de las columnas de la matriz de probabilidades estacionarias igual a cero se llaman *semi-ergódicas*. Las cadenas ergódicas pueden considerarse como un caso particular de las cadenas semi-ergódicas, en las que no existen probabilidades de estado estable iguales a cero.

$$\mathbf{P}_2^{16} = \begin{vmatrix} 0,005 & 0,007 & 0,475 & 0,563 \\ 0,002 & 0,005 & 0,426 & 0,567 \\ 0 & 0 & 0,429 & 0,571 \\ 0 & 0 & 0,429 & 0,571 \end{vmatrix}$$

3.1.5.3 Cadenas no ergódicas

La cadena C3, de cuatro estados, tiene la siguiente matriz de transición:

$$\mathbf{P}_3 = \begin{vmatrix} 0,5 & 0,4 & 0,1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0,2 & 0,8 \\ 0 & 0 & 0,6 & 0,4 \end{vmatrix}$$

Si observamos la matriz de la transición 16, podemos ver que, mientras algunas filas tienen el mismo comportamiento que las de los casos anteriores, vemos que otras tienden a ciertos valores, diferentes de los de las otras filas. Ello quiere decir que, al contrario de lo que sucede con el caso regular, las probabilidades de estado estable sí dependen de cuál ha sido el estado inicial de la cadena. Se trata de una cadena *semi-regular*.

$$\mathbf{P}_3^{16} = \begin{vmatrix} 0,000 & 0,800 & 0,086 & 0,114 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0,429 & 0,571 \\ 0 & 0 & 0,429 & 0,571 \end{vmatrix}$$

3.1.5.4 Cadenas cíclicas

La cadena C4, cuya matriz de probabilidades de transición se muestra a continuación, después de un número elevado de transiciones presenta un comportamiento diferente del de las cadenas anteriores.

$$\mathbf{P}_4 = \begin{vmatrix} 0,5 & 0,4 & 0,1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0,2 & 0 & 0,8 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

Al ir obteniendo matrices de transición, se observa que éstas no convergen a un valor concreto, sino que muestran un comportamiento cíclico. En este caso, las transiciones impares tienden a un valor y las pares a otro:

$$\mathbf{P}_4^{2k} = \begin{vmatrix} 0 & 0,08 & 0,60 & 0,32 \\ 0 & 0,2 & 0 & 0,8 \\ 0 & 0 & 1 & 0 \\ 0 & 0,2 & 0 & 0,8 \end{vmatrix}$$

$$\mathbf{P}_4^{2k+1} = \begin{vmatrix} 0 & 0,2 & 0,4 & 0,48 \\ 0 & 0 & 1 & 0 \\ 0 & 0,2 & 0 & 0,8 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

Este tipo de cadenas son cadenas cíclicas. En este caso particular, nos encontramos ante una cadena de periodo $p=2$.

Obsérvese que la primera columna es siempre cero, por lo que el estado 1 no aparecerá en las probabilidades a largo plazo; quiere ello decir que la cadena considerada no es ergódica, aunque es claro que pueden existir cadenas cíclicas ergódicas, como veremos en ejemplos posteriores.

También debemos preguntarnos qué ocurre con las probabilidades estacionarias en las cadena cíclicas, ya que si las sucesivas potencias de \mathbf{P} no tienden hacia unos valores determinados. Más adelante, cuando estudiemos el cálculo sistemático de \mathbf{P}^* , daremos una respuesta a esta pregunta.

3.1.5.5 Clasificación de cadenas de Markov

Los distintos casos examinados son sólo algunos (los más importantes) de las diferentes situaciones que se pueden dar en cuanto a la clasificación de las cadenas de Markov. En el apartado 3.2.4 se darán de forma completa y sistematizada todas las definiciones.

De lo expuesto hasta ahora, si queremos analizar el comportamiento a largo plazo de un proceso estocástico que cumpla la propiedad markoviana, necesitamos:

- Una metodología para poder clasificar la cadena como ergódica o no ergódica por una parte, y como regular, semirregular o cíclica por otra, examinando la matriz de probabilidades de transición.
- Una metodología que permita el cálculo de la matriz de probabilidades estacionarias.

La clasificación de las cadenas de Markov puede realizarse mediante dos metodologías:

- El *análisis topológico*, examinando las propiedades de los estados de la cadena y estableciendo clases de equivalencia entre los estados.
- El *análisis espectral*, examinando los valores propios de la matriz de probabilidades de transición de un paso.

Una vez clasificada la cadena, puede obtenerse información acerca de la forma que presente la matriz de probabilidades estacionarias, lo cual facilita su obtención.

3.2 Análisis topológico de las cadenas de Markov

El análisis de topológico de las cadenas de Markov permite la clasificación de las cadenas a partir de la información suministrada por la matriz \mathbf{P} utilizando propiedades relativas a la relación entre estados (propiedades de estado). Estas propiedades permiten, a su vez, definir subconjuntos de estados denominados *clases*. También podremos definir, entonces, las *propiedades de clase*.

3.2.1 Propiedades de estado

Dados dos estados de una cadena, pueden establecerse dos tipos de relaciones entre ellos:

- El estado i es descendiente de j si cuando iniciamos el proceso en i existe una probabilidad no nula de que el proceso llegue a j . En este caso, diremos que existe un camino entre los estados i y j .
- Los estados i y j se comunican si i es descendiente de j y j es descendiente de i .
- Existirá un *ciclo* dentro de una cadena de Markov si existe un camino en la cadena que comunique al estado i consigo mismo. Dicho circuito se caracterizará por el número mínimo de transiciones que necesitará el sistema para volver al estado i , si se inició el proceso en ese estado. Dicho número constituirá la longitud del ciclo.

Obsérvese que, con las definiciones dadas, la existencia de un circuito implica que todos los estados que lo forman están comunicados. Se conviene que todo estado está comunicado consigo mismo, ya que se al menos puede acceder a él en cero transiciones (circuito de longitud cero), con independencia de que además existan otros circuitos de longitud mayor.

Para analizar estas relaciones entre estados, es útil recordar que, según la teoría de grafos, toda matriz cuadrada tiene asociado un grafo, cuya representación gráfica se puede elaborar a partir de la matriz de probabilidades de transición, el *diagrama de transiciones* de estados.

Cada estado de la cadena se representa por un vértice del grafo y cada transición con probabilidad no nula se representa por una relación entre los vértices que representan los estados anterior y posterior de la misma. De esta manera en el diagrama se representan todas las situaciones en las que un estado i es descendiente respecto de j . En la figura 3.2.1.a se muestra el diagrama de transición de una cadena de Markov, junto con su matriz de transición.

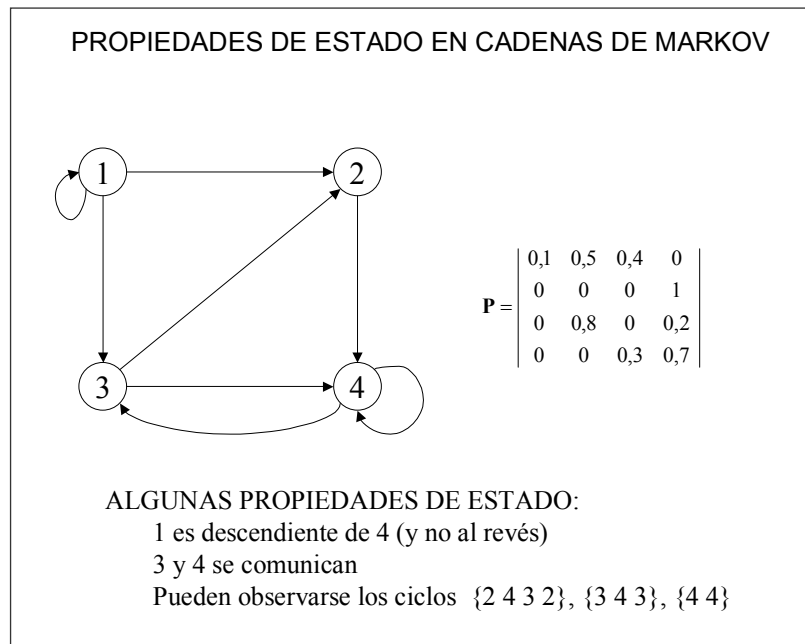


Fig. 3.2.1.a Propiedades de estado en cadenas de Markov

3.2.2 Propiedades de clase

Dado que, por convenio, hemos establecido que un estado está siempre comunicado consigo mismo, la relación entre estados *estar comunicado* es reflexiva, simétrica y transitiva, por lo que se trata de una relación de equivalencia. Por este motivo, podemos decir que un conjunto de estados comunicados entre sí constituye una *clase de equivalencia*. De esta manera, podemos clasificar en diversas clases los estados de una cadena de Markov.

A su vez, podemos definir la propiedad de clase siguiente para las clases de equivalencia que se hayan establecido:

- Una clase de equivalencia será una *clase final* si cuando el proceso llega a uno de los estados de la clase, en las transiciones siguientes el proceso evoluciona siempre dentro de los estados de la clase.
- Aquellas clases de equivalencia que no sean clases finales serán *clases de paso*. Las clases de paso tienen un interés muy limitado en el estudio de las cadenas de Markov.

Es claro que, puesto que el sistema debe ser capaz de evolucionar indefinidamente entre un número finito de estados, toda cadena debe tener al menos una clase final. Si en su evolución a lo largo de infinitas transiciones el sistema puede pasar por todos los estados, entonces habrá una única clase final que los englobará a todos ellos. Este caso es el que hemos definido anteriormente como *cadena ergódica*.

La figura 3.2.2.a muestra, sobre una cadena de Markov, una clase final y una clase de paso.

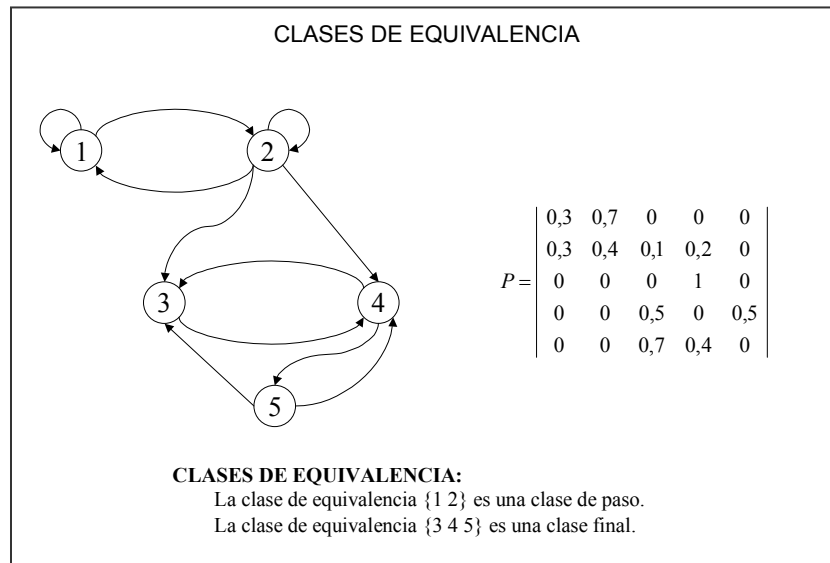


Fig. 3.2.2.a Propiedades de clase en una cadena de Markov

3.2.3 PERIODICIDAD DE UNA CLASE FINAL

Para una clase final, se puede obtener el *periodo* de una clase como el máximo común divisor de las longitudes de los ciclos que pueden encontrarse en esa clase. Tendremos entonces dos tipos de clases finales:

- *Clases cíclicas (o periódicas)*: aquellas que tengan un periodo de valor $p > 1$.
- *Clases acíclicas (o aperiódicas)*: aquellas que tengan un periodo de valor $p = 1$.

La figura 3.2.3.a muestra varias cadenas de Markov. La primera tiene ciclos de longitud 2 y 3, por lo que es aperiódica, ya que $\text{MCD}(2,3) = 1$. En la segunda se pueden observar ciclos de longitud 2 y 4, por lo que su periodo es $p = 2$. Obsérvese que la existencia de estos ciclos no impide la de otros mayores (por ejemplo 1-2-4-3-4-3-4-2-1, que tiene longitud ocho. Sin embargo, todos estos ciclos serán múltiplos de dos). La última, similar a la anterior, tiene sin embargo ciclos de longitud 1, 2 y 4, por lo que es también aperiódica. La mera existencia de un ciclo de longitud uno (en este caso el 1-1) implica que la clase es aperiódica.

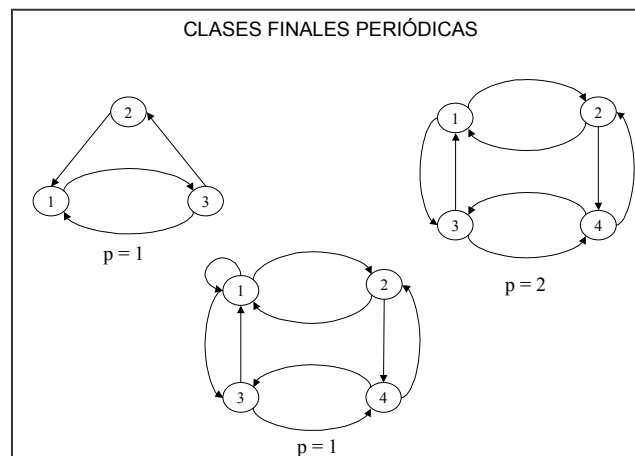


Fig. 3.2.3.a Ciclicidad en clases finales de las cadenas de Markov

3.2.4 Clasificación de las cadenas a partir del análisis topológico

Con las propiedades que se han definido, podemos establecer dos clasificaciones distintas de las cadenas de Markov según su número de clases finales y la periodicidad o no de sus clases finales:

Clasificación según el número de clases finales

Podemos tener las siguientes situaciones en relación con las clases finales de una cadena de Markov:

- La cadena tiene una única clase final y no tiene clases de paso. Se trata de una *cadena ergódica*.
- La cadena tiene una única clase final, pero tiene clases de paso. Tenemos entonces una cadena *semiergódica*.
- La cadena tiene más de una clase final y una o varias clases de paso. La cadena es entonces *no ergódica*.

Otra situación que podemos encontrar es la de varias clases de finales, sin que existan clases de paso. Entonces, más que una única cadena, tenemos varias cadenas ergódicas sin relación entre ellas. Por esta razón, no se ha incluido en la clasificación.

Clasificación según la periodicidad de las clases finales

Según el número y la periodicidad de las clases finales, tenemos diversas situaciones:

- Si tenemos una sola clase final aperiódica, tendremos una cadena de Markov *regular*. Una cadena regular puede ser ergódica o semiergódica, según tenga clases de paso o no.
- Si tenemos varias clases finales, todas ellas aperiódicas, tendremos una cadena de Markov *semirregular*.
- Si tenemos varias clases finales, y todas ellas son periódicas, se trata entonces de una cadena *policíclica*.
- Finalmente, si tenemos varias clases finales, y algunas son cíclicas y otras no, hablaremos de una *cadena mixta*.

Las cadenas semirregulares, policíclicas y mixtas son por tanto los tres casos posibles de cadenas no ergódicas. En la tabla 3.2.4.a se muestran las diversas posibilidades derivadas de esta doble clasificación:

Tabla 3.2.4.a. Clasificación cadenas de Markov

	Ergódica	Semiergódica	No ergódica
Regular	Regular positiva	Regular	(No posible)
Cíclica	Cíclica	Cíclica	(No posible)
Semirregular	(No posible)	(No posible)	Semirregular
Policíclica	(No posible)	(No posible)	Policíclica
Mixta	(No posible)	(No posible)	Mixta

En las figuras 3.2.4.a y 3.2.4.b se procede a clasificar las cadenas C1, C2, C3 y C4 a partir del análisis topológico, mostrando las clases finales y las clases de paso.

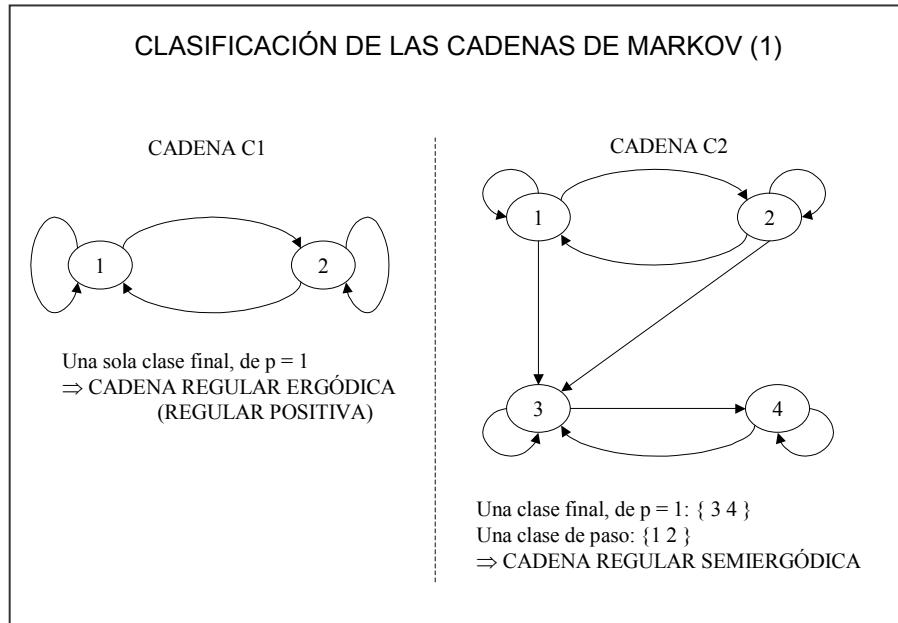


Fig. 3.2.4.a Ejemplos de clasificación de las cadenas de Markov

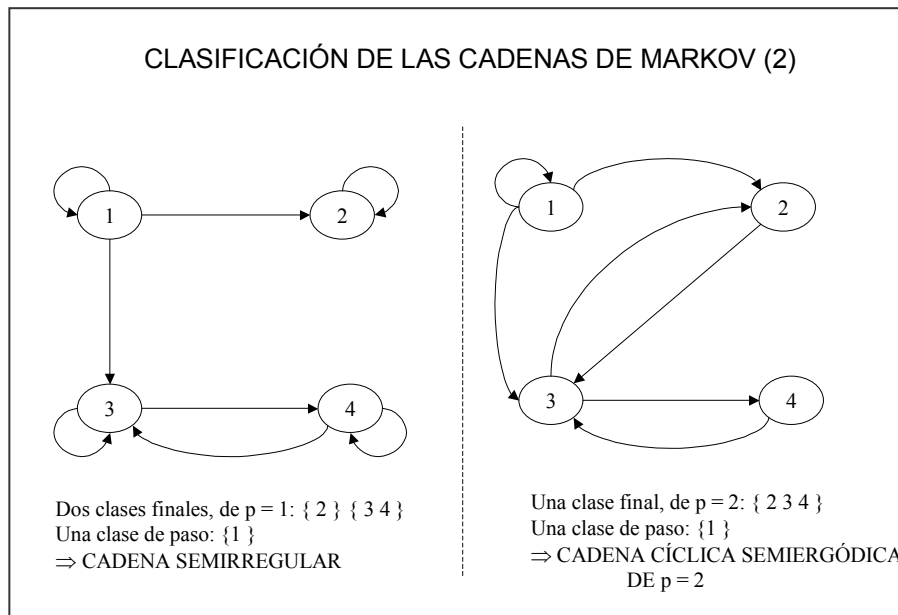


Fig. 3.2.4.b Ejemplos de clasificación de las cadenas de Markov

3.2.5 Significado de la ergodicidad

El concepto de cadena ergódica tiene relación con el comportamiento a largo plazo del sistema. En una cadena ergódica, todos los estados se pueden dar a largo plazo, es decir, que en régimen permanente todos los estados tienen una probabilidad no nula de aparecer: se trata de cadenas cuyo comportamiento no varía de forma cualitativa a lo largo del tiempo. En una cadena ergódica es lo mismo el largo plazo que el corto plazo, salvo que se conserva la memoria del estado inicial.

En las cadenas semiergódicas, el comportamiento a largo y a corto plazo es cualitativamente distinto. Existen unos estados transitorios, que a largo plazo desaparecerán y por tanto sólo pueden aparecer mientras el sistema no haya alcanzado su régimen permanente. Una vez alcanzado éste, el comportamiento de la cadena no difiere del caso ergódico, si se hace abstracción de los estados que ya no pueden aparecer.

Finalmente, en las cadenas no ergódicas el comportamiento a largo plazo depende de la situación inicial. Al revés de lo que ocurre en los dos casos anteriores, el comportamiento a largo plazo no está definido cualitativamente, sino que será distinto según la clase final a la que vaya a parar el sistema. Ello significa que la probabilidad estacionaria de cada estado depende de la situación inicial, que puede llegar a determinar que ciertos estados no se pueden alcanzar jamás si se parte de un cierto conjunto de estados.

3.2.6 Significado de la periodicidad o ciclicidad

Una clase final acíclica representa una situación en la que la cadena evoluciona entre los distintos estados de la misma de forma completamente aleatoria, descrita según las probabilidades indicadas por la ley condicional (o matriz \mathbf{P}). Podríamos decir que se trata de un sistema estocástico “puro”, en el sentido que no se puede hacer una descripción de su comportamiento más precisa que las meras probabilidades. El significado de las probabilidades estacionarias representa en estas clases finales la probabilidad real de que en una transición cualquiera el sistema se halle en uno u otro estado.

Las cadenas regulares y semiregulares, por tanto, representan sistemas que acaban yendo a parar a una situación de este tipo, sea ésta única o no.

Un caso particular interesante dentro de las clases finales acíclicas es el de los llamados *estados absorbentes*: se trata de estados que constituyen por sí mismos una sola clase final, puesto que la única transición posible es ir otra vez al mismo. Matemáticamente significa que la fila correspondiente de \mathbf{P} estará toda a ceros excepto un 1 en la diagonal principal; el significado de este tipo de situaciones suele ser el de un sistema que ha llegado a una situación de degradación, que ya no puede evolucionar más, etc.

En las clases finales cíclicas se puede establecer una partición de sus estados en p subclases, de manera que la evolución de la cadena dentro de ésta clase consistirá en que pasará sucesivamente por un estado de la subclase 1, luego uno de la subclase 2, etc. hasta llegar a la subclase p , después del cual volverá a un estado de la subclase 1; en el caso de C_4 ($p = 2$) estas dos subclases son $\{2,4\}$ y $\{3\}$. Las clases finales cíclicas, pues, son un caso intermedio entre un sistema estocástico puro en el que no se puede predecir nada de su evolución más allá de la mera descripción probabilística y el caso determinista en que la evolución es conocida y por lo tanto predecible y calculable: en este caso no se conoce cuál será el estado siguiente con exactitud, pero sí que forma parte de un subconjunto bien determinado.

Ello quiere decir que la interpretación de las probabilidades estacionarias es ligeramente distinta del caso acíclico. Si se observa el sistema en un momento cualquiera, la probabilidad de hallarlo en uno u otro estado es efectivamente la probabilidad estacionaria, pero ello es sólo fruto de nuestra ignorancia del número de transiciones realizadas. Si conociéramos este número (o más precisamente su módulo p), las probabilidades serían distintas, puesto que las potencias de \mathbf{P} evolucionan cíclicamente.

3.3 Análisis espectral de las cadenas de Markov

Si conocemos los valores propios de una matriz de probabilidades de transición, podemos conocer ciertas propiedades de su cadena de Markov asociada mediante el *análisis espectral*.

Se parte de la siguiente propiedad general de las matrices estocásticas:

Los valores propios complejos de una matriz estocástica tienen módulo inferior o igual a uno.

Para detectar el número de clases finales y su periodicidad, el análisis espectral parte de este hecho y de la siguiente propiedad:

Cada clase final de periodo p genera p valores propios de valor igual a las raíces p -ésimas de 1. El resto de valores propios serán de módulo inferior a 1.

En consecuencia, el número de clases finales será igual a la multiplicidad de 1 como valor propio.

De este modo, a partir del examen de los valores propios de la matriz de probabilidades de transición podemos llegar a las conclusiones siguientes, según el análisis espectral:

Una cadena de Markov, *regular*, tanto *ergódica* como *semiergódica*, tendrá un único valor propio de valor 1. El resto de valores propios serán de módulo inferior a 1.

Una cadena de Markov *semirregular*, tendrá un valor propio 1 múltiple, y el resto de valores propios serán de módulo inferior a 1. La multiplicidad del valor propio 1 será igual al número de clases finales de la cadena.

Una cadena de Markov *cíclica* (o *polícíclica*) se caracterizará por tener, entre sus valores propios, uno o varios conjuntos de raíces p -ésimas de 1. Cada uno de esos conjuntos revelará la presencia en la cadena de una clase cíclica de periodo p .

El análisis espectral no dice nada, sin embargo, sobre las clases de paso y por ello tampoco informa sobre si se trata de una cadena ergódica o no. Únicamente se puede deducir que si el valor propio +1 es múltiple, se tratará de una cadena no ergódica, y en caso contrario será ergódica o semiergódica.

A modo de ejemplo, podemos tener situaciones como las siguientes:

Ejemplo 3.3.a:

Una cadena de Markov con el conjunto de valores propios siguiente:

$$\{-0,8 \quad 0,4 \quad +1 \quad -1 \quad +1 \quad -0,3 \quad 0,01 \quad 0 \}$$

Es una cadena de Markov cíclica: cuenta con una clase final de periodo 1 y con una clase final de periodo 2.

Dado que existen ocho valores propios, la cadena debe tener precisamente ocho estados, pero no sabemos cuáles pertenecen a cada una de las dos clases finales, ni si hay alguno en clases de paso o no. Sí sabemos que por tratarse de una cadena mixta se trata de un caso no ergódico.

Ejemplo 3.3.b

Una cadena de Markov con el conjunto de valores propios siguiente:

$$\{-0,4 \quad 0,3 \quad 0,2 \quad -i \quad +i \quad -1 \quad +1 \quad -1 \quad +1 \}$$

Ahora tenemos una cadena de Markov con dos clases finales, puesto que, como en la cadena anterior, la multiplicidad de 1 es doble: una de las clases es de periodo 4, mientras que la segunda es de periodo 2. En consecuencia, se trata de una cadena policíclica. Obsérvese que las filas correspondientes a la clase de período 2 se repetirán alternativamente en las potencia pares e impares de \mathbf{P} , mientras que en la clase de período 4 se repetirán cada 4 potencias; por lo tanto, la matriz \mathbf{P}^k se repetirá cada 4 potencias.

3.4 Cálculo de las probabilidades estacionarias

La clasificación de la cadena de Markov permite identificar la forma de la matriz de probabilidades estacionarias \mathbf{P}^* , y así facilitar su cálculo exacto, sin necesidad de obtener una potencia elevada de la matriz \mathbf{P} (procedimiento desaconsejable, por otra parte, porque puede dar lugar a confusiones, como veremos).

Dado que la matriz de probabilidades de estado estable identifica las probabilidades de transición después de un número de pasos suficientemente grande, si la matriz \mathbf{P} es de una cadena ergódica o semi-ergódica, \mathbf{P}^* es una matriz con todas las filas iguales, hecho que facilita notablemente su cálculo. Para las cadenas no ergódicas, al ser algunas de las filas de \mathbf{P}^* diferente del resto, su cálculo será más complejo.

Para todas las clases de cadenas que no tengan clases finales cíclicas, podemos escribir la siguiente identidad:

$$\mathbf{P}^* = \lim_{n \rightarrow \infty} \mathbf{P}^n$$

con lo que basta hallar este límite para conocer \mathbf{P}^* si la cadena no tiene clases finales cíclicas.

En consecuencia, en las cadenas cíclicas, la relación de \mathbf{P}^* con su comportamiento a largo plazo requerirá alguna aclaración adicional. Para aclarar estas cuestiones se procederá, mediante ejemplos, a mostrar cómo se calcula \mathbf{P}^* para diferentes tipos de cadenas de Markov.

3.4.1. Probabilidades estacionarias para cadenas regulares ergódicas

Dado que se trata de una cadena regular y ergódica, al cabo de un cierto número de transiciones suficientemente elevado se habrá perdido la información del estado inicial, por lo que la probabilidad de que el sistema se halle en uno u otro estado no dependerá de dicho estado. Todas las filas de \mathbf{P}^* en una cadena regular son iguales, por lo que la matriz adopta la forma:

$$\mathbf{P}^* = \begin{pmatrix} \pi_1 & \pi_2 & \dots & \pi_n \\ \pi_1 & \pi_2 & \dots & \pi_n \\ \vdots & \vdots & \ddots & \vdots \\ \pi_1 & \pi_2 & \dots & \pi_n \end{pmatrix}$$

Sabiendo que la matriz \mathbf{P}^* cumple la ecuación:

$$\mathbf{P}^* \cdot \mathbf{P} = \mathbf{P}^* \cdot \mathbf{P} = \mathbf{P}^*$$

Desarrollando esta ecuación matricial para una cualquiera de las filas de las matrices, obtenemos el siguiente conjunto de ecuaciones:

$$\sum_{k=1}^n \pi_k \cdot p_{ik} = \pi_i \quad \text{para } i=1, 2, \dots, n$$

Al que hemos de añadir la ecuación, propia de toda matriz estocástica:

$$\sum_{i=1}^n \pi_i = 1$$

Tenemos entonces $n+1$ ecuaciones para n incógnitas. Sin embargo, las ecuaciones procedentes del producto de matrices están indeterminadas, puesto que podemos tener otras soluciones a partir de un conjunto de soluciones inicial, multiplicando todos los valores de ésta por un escalar. Es la última ecuación la que hace a \mathbf{P}^* estocástica. En consecuencia, deberá eliminarse alguna de las procedentes del producto de matrices, y nunca la que hace la suma de las π_i igual a 1.

Ejemplo 3.4.1.a Probabilidades de transición para una cadena regular ergódica

La cadena C1, de matriz de probabilidades de transición:

$$\mathbf{P}_1 = \begin{vmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{vmatrix}$$

Tiene una matriz de probabilidades estacionarias de la forma:

$$\mathbf{P}^* = \begin{vmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{vmatrix}$$

Dichas probabilidades pueden obtenerse con las ecuaciones siguientes, derivadas de la ecuación matricial $\mathbf{P}^* \cdot \mathbf{P} = \mathbf{P}^*$:

$$\begin{aligned} 0,2 \cdot \pi_1 + 0,6 \cdot \pi_2 &= \pi_1 \\ 0,8 \cdot \pi_1 + 0,4 \cdot \pi_2 &= \pi_2 \\ \pi_1 + \pi_2 &= 1 \end{aligned}$$

Eliminando una de las dos primeras ecuaciones, encontramos los valores:

$$\begin{aligned} \pi_1 &= 3/7 = 0,429 \\ \pi_2 &= 4/7 = 0,571 \end{aligned}$$

3.4.2 Probabilidades estacionarias para cadenas regulares no ergódicas

Las matrices de probabilidades estacionarias de este tipo de cadenas regulares tendrán la misma forma que las matrices de las cadenas regulares ergódicas, con el rasgo distintivo de que las $\pi_{\text{clases de paso}} = 0$, dado que después de infinitas transiciones el sistema se encontrará en una clase final con toda seguridad.

Ejemplo 3.4.2.a Probabilidades de transición para una cadena regular no ergódica

La matriz de probabilidades estacionarias de la cadena C2, es:

$$\mathbf{P}_2 = \begin{vmatrix} 0,5 & 0,4 & 0,1 & 0 \\ 0,3 & 0,3 & 0,4 & 0 \\ 0 & 0 & 0,2 & 0,8 \\ 0 & 0 & 0,6 & 0,4 \end{vmatrix}$$

Y la matriz de probabilidades estacionarias será de la forma:

$$\mathbf{P}^* = \begin{vmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 \\ \pi_1 & \pi_2 & \pi_3 & \pi_4 \end{vmatrix}$$

Haciendo de nuevo $\mathbf{P}^* \cdot \mathbf{P} = \mathbf{P}^*$ tenemos el conjunto de ecuaciones:

$$\begin{aligned} 0,5 \cdot \pi_1 + 0,3 \cdot \pi_2 &= \pi_1 \\ 0,4 \cdot \pi_1 + 0,3 \cdot \pi_2 &= \pi_2 \\ 0,1 \cdot \pi_1 + 0,4 \cdot \pi_2 + 0,2 \cdot \pi_3 + 0,6 \cdot \pi_4 &= \pi_3 \\ 0,8 \cdot \pi_3 + 0,4 \cdot \pi_4 &= \pi_4 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 &= 1 \end{aligned}$$

Eliminando una de las cuatro primeras ecuaciones, obtenemos los siguientes valores:

$$\begin{aligned} \pi_1 &= 0 \\ \pi_2 &= 0 \\ \pi_3 &= 3/7 = 0,429 \\ \pi_4 &= 4/7 = 0,571 \end{aligned}$$

Los estados 1 y 2, cada uno de ellos clase de paso, tienen probabilidades estacionarias iguales a cero: en el largo plazo, no encontraremos el proceso en ninguno de estos dos estados.

Aunque los resultados $\pi_1 = \pi_2 = 0$ se obtienen del sistema anterior, es obvio que si sabemos ya que estos estados son de paso, podemos eliminarlos directamente dos ecuaciones y dos incógnitas del sistema, con lo que el cálculo resultará más sencillo.

3.4.3 Probabilidades estacionarias para cadenas semirregulares

Aunque las cadenas semirregulares presentan mayor complejidad, la naturaleza de las distintas clases facilita también el cálculo de la matriz \mathbf{P}^* :

- Las columnas de \mathbf{P}^* pertenecientes a estados de paso serán todas cero.
- Las filas de los estados de *una misma clase final* serán todas iguales, con los valores de las columnas correspondientes a estados no pertenecientes a la clase final considerada iguales a cero. Por supuesto, cada clase final dará lugar a un "formato" distinto de fila.
- Las filas de los estados de las *clases de paso* serán de la forma siguiente: los valores de las π_i de las transiciones con destino en estados pertenecientes a clases de paso serán cero, y diferentes de cero los π_j de las transiciones con destino en las clases finales. Cada estado perteneciente a una clase de paso tendrá una fila diferente en \mathbf{P}^* .

Se trata entonces de proceder en dos pasos:

Se hallan las matrices \mathbf{P}^* correspondientes a las cadenas regulares positivas definidas por cada una de las clases finales (según el procedimiento descrito en 3.4.1) y se componen con ellas las correspondientes filas de sus estados.

Para hallar las filas de los estados de paso, se resuelven tantos sistemas de ecuaciones

$$\sum_{k=0}^n \pi_k \cdot p_{ki} = \pi_i \quad \text{para } i=1, 2, \dots, n$$

$$\sum_{i=1}^n \pi_i = 1$$

como clases finales y estados de las clases de paso tenga la cadena.

Ejemplo 3.4.3.a Matriz \mathbf{P}^* para una cadena semirregular

Sea la cadena de Markov definida por la siguiente matriz de probabilidades de transición:

$$\mathbf{P} = \begin{array}{c|cccccccc|} & 0.7 & 0.1 & 0.1 & 0 & 0 & 0 & 0 & 0.1 \\ & 0.4 & 0.3 & 0.1 & 0 & 0.2 & 0 & 0 & 0 \\ & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 & 0 \\ \mathbf{P} = & 0 & 0 & 0.6 & 0.4 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0.1 & 0.3 & 0.3 & 0.1 & 0.2 & 0 \\ & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 \\ & 0 & 0 & 0 & 0 & 0 & 0.4 & 0.6 & 0 \\ & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}$$

El análisis topológico nos dice que hay tres clases finales acíclicas: {3,4}, {6,7} y {8}, así como dos clases de paso: {1,2} y {5}. Se trata por tanto de una cadena semirregular.

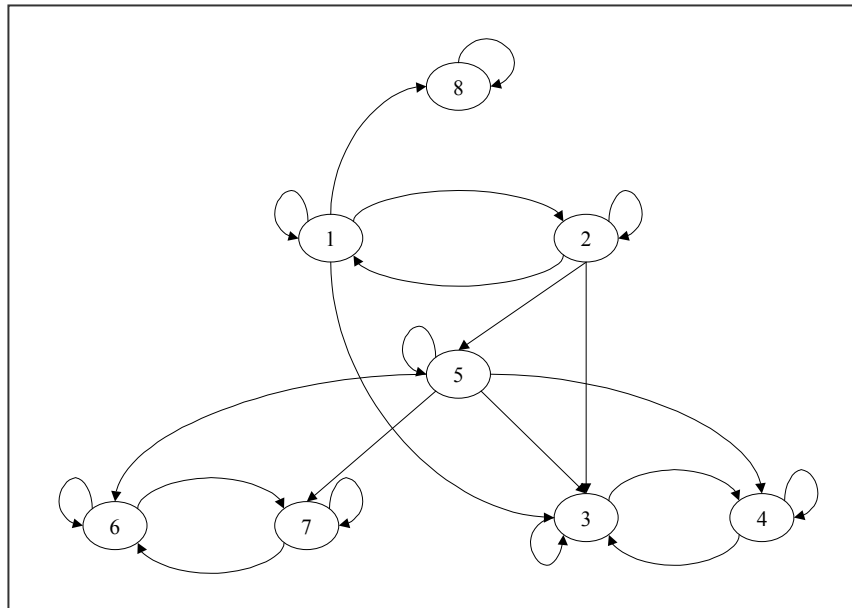


Fig. 3.4.3.a Cadena de Markov semirregular

Analizando las matrices ergódicas regulares de las tres clases finales, podemos hallar buena parte de los valores de \mathbf{P}^* :

$$\mathbf{P}^* = \begin{array}{c|cccccccc}
 & 0 & 0 & a & b & 0 & c & d & e \\
 & 0 & 0 & i & j & 0 & k & l & m \\
 & 0 & 0 & \boxed{2/3} & \boxed{1/3} & 0 & 0 & 0 & 0 \\
 & 0 & 0 & \boxed{2/3} & \boxed{1/3} & 0 & 0 & 0 & 0 \\
 & 0 & 0 & p & q & 0 & r & s & t \\
 & 0 & 0 & 0 & 0 & 0 & \boxed{4/9} & \boxed{5/9} & 0 \\
 & 0 & 0 & 0 & 0 & 0 & \boxed{4/9} & \boxed{5/9} & 0 \\
 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1}
 \end{array}$$

donde se han representado con letras los valores aún no conocidos y se han recuadrado las submatrices correspondientes a las clases finales.

Sin embargo, del diagrama de transiciones se puede deducir fácilmente que desde el estado 5 no se puede ir a la clase $\{8\}$ por lo que $t = 0$.

Obsérvese que aunque la transición 2-8 tiene probabilidad cero, lo mismo que las 1-6 y 1-7, *no* se puede deducir de ello que, por ejemplo, $c = 0$, ya que el estado 6 sí es descendente del 1, a través del 2 y del 5.

Nos quedan por lo tanto 14 incógnitas, por lo que debemos ahora establecer 14 ecuaciones para hallar estos valores. Escribimos en primer lugar las siempre obligadas:

$$\begin{aligned}
 a + b + c + d + e &= 1 \\
 i + j + k + l + m &= 1 \\
 p + q + r + s &= 1
 \end{aligned}$$

El resto de ecuaciones se deben establecer a partir de las expresiones ya conocidas:

$$\mathbf{P P}^* = \mathbf{P}^* \mathbf{P} = \mathbf{P}^*$$

Esta doble igualdad matricial nos permite, en teoría, escribir $(2^n)^2 = 128$ igualdades, de entre las cuales deberemos elegir las 11 que faltan. Muchas de estas igualdades serán simples identidades que no sirven para nuestro propósito; también existen combinaciones lineales entre ellas que no podemos utilizar.

La tarea, por tanto, parece complicada; sin embargo, de las propiedades de estas matrices se pueden deducir algunas relaciones que permiten establecer fácilmente algunas de las ecuaciones necesarias. Hay que decir que todas ellas se pueden hallar a partir de las expresiones matriciales indicadas, pero de esta manera es mucho más sencillo hallarlas sin tener que buscar en la maraña de 128 igualdades posibles.

Consideremos en primer lugar los valores a y b : representan las probabilidades de hallar el sistema en los estados 3 y 4 sabiendo que ha empezado su evolución en el 1. Pero ello quiere decir que el sistema ha ido a parar a la clase $\{3,4\}$ lo cual tiene una probabilidad que llamaremos $P_{1,34}$. Una vez el sistema ha ido a parar a dicha clase, la probabilidad que se halle en uno u otro estado es justamente la probabilidad estacionaria de la clase final, en nuestro caso, $2/3$ y $1/3$. De ello se deduce que $a = P_{1,34} * 2/3$ y $b = P_{1,34} * 1/3$, de lo que se concluye que a y b son proporcionales a $2/3$ y $1/3$. Por supuesto la probabilidad $P_{1,34}$ es desconocida, aunque sí sabemos que vale precisamente $a + b$.

Esta propiedad, que se cumple siempre para todas las clases finales, nos permite establecer inmediatamente seis ecuaciones más:

$$\begin{aligned} a / b &= i / j = p / q = (2/3) / (1/3) = 2 \\ c / d &= k / l = r / s = 4 / 5 \end{aligned}$$

En este caso existe además otra propiedad (que no se da siempre) que también nos facilita escribir algunas de las ecuaciones. En el estado 5, la cadena sólo puede hacer dos cosas: o bien se mantiene en él, o bien va directamente a una clase final. Si seguimos la notación anteriormente establecida, es evidente que $P_{5,34} + P_{5,67} = 1 - 0,3 = 0,7$ y como $p + q = P_{5,34} = (0,1 + 0,3) / 0,7$ y análogamente para $P_{5,67} = r + s$, esto supone que podemos escribir:

$$\begin{aligned} p + q &= 4/7 \\ r + s &= 3/7 \end{aligned}$$

Desgraciadamente, estas igualdades no proporcionan dos ecuaciones más, sino sólo una, ya que hay combinación lineal con las anteriores; no obstante, basta con ella para hallar los valores que permiten completar la fila 5:

$$p = 8/21; q = 4/21; r = 4/21; s = 5/21; t = 0$$

Téngase en cuenta que esto *no* se puede escribir para los estados 1 y 2, ya que de ellos no se va directamente sólo a clases finales. Tampoco vale el hecho de que $\{1,2\}$ sea una misma clase de paso: los estados de la misma clase de paso no tienen por qué tener las mismas probabilidades estacionarias. Por supuesto, un análisis probabilístico parecido nos permitiría establecer expresiones semejantes para los estados 1 y 2, pero ello generalmente conduce a situaciones más complejas que no suele valer la pena acometer.

Así pues, en este momento tenemos ya 5 valores calculados y otras 6 ecuaciones para hallar los 15 valores desconocidos de la matriz escrita más arriba. Pero no hay que preocuparse demasiado por el sistema de ecuaciones resultante: usualmente se puede reducir con cierta facilidad.

Para ello lo más productivo suele ser desarrollar las expresiones para el cálculo de los valores desconocidos mediante la expresión $\mathbf{P}^* = \mathbf{P} \cdot \mathbf{P}^*$ en grupos de la misma columna:

Empecemos por la pareja a, i :

$$\begin{aligned} a &= 0,7a + 0,1i + 0,1 \cdot (2/3) & \Rightarrow & \quad 3a - i = 2/3 \\ i &= 0,4a + 0,3i + 0,1 \cdot (2/3) + 0,2p & \Rightarrow & \quad 7i - 4a = 1/7 \end{aligned}$$

Como p es conocido, queda tan solo un sistema de dos ecuaciones y dos incógnitas, que se resuelve fácilmente y permite además obtener inmediatamente los valores de b y j :

$$a = 128/357; b = 64/357; i = 146/357; j = 73/357$$

Repitamos la técnica con c y k :

$$\begin{aligned} c &= 0,7c + 0,1k & \Rightarrow & \quad 3c - k = 0 \\ k &= 0,4c + 0,3k + 0,2r & \Rightarrow & \quad 7k - 4c = 8/21 \end{aligned}$$

con lo que $c = 8/357; d = 10/357; k = 8/119; l = 10/119$

Y finalmente, para e y m :

$$\begin{aligned} e &= 0,7e + 0,1m + 0,1 \cdot 1 & \Rightarrow & \quad 3e - m = 1 \\ m &= 0,4e + 0,3m & \Rightarrow & \quad 7m - 4e = 0 \end{aligned}$$

con lo cual $e = 7/17; m = 4/17$

Se puede comprobar que, efectivamente, la suma de los valores de cada fila es 1, y la matriz de probabilidades estacionarias resulta ser:

$$\mathbf{P}^* = \begin{pmatrix} 0 & 0 & 128/357 & 64/357 & 0 & 8/357 & 10/357 & 7/17 \\ 0 & 0 & 146/357 & 73/357 & 0 & 24/357 & 30/357 & 4/17 \\ 0 & 0 & 2/3 & 1/3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2/3 & 1/3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 8/21 & 4/21 & 0 & 5/21 & 4/21 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4/9 & 5/9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4/9 & 5/9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

De esta matriz es inmediato deducir las probabilidades de que el sistema vaya a una u otra clase final según el estado inicial:

$$\begin{aligned} P_{1,34} &= 64/119; & P_{1,67} &= 6/119; & P_{1,8} &= 7/17 \\ P_{2,34} &= 73/119; & P_{2,67} &= 18/119; & P_{2,8} &= 4/17 \\ P_{5,34} &= 4/7; & P_{5,67} &= 3/7; & P_{5,8} &= 0 \end{aligned}$$

Tres observaciones finales:

- Primera: en este caso la existencia de un estado de paso del cual sólo se puede salir para ir a alguna clase final permite el cálculo directo de la fila correspondiente; por supuesto eso no tiene por qué ocurrir en general.
- Segunda: la utilización en cada paso de los valores ya obtenidos permite la reducción de los sistemas de ecuaciones resultantes y por tanto también facilita el cálculo incluso manual.

- Tercera: la técnica de escribir las ecuaciones para los elementos de la misma columna permite en general trocear el sistema en otros más pequeños y con ello reducir notablemente la dificultad del cálculo.

Ejemplo 4.3.b Matriz P^* para una cadena de Markov semirregular

Sea la cadena de Markov cuya matriz de transición de un paso es:

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

El lector puede comprobar fácilmente, mediante análisis topológico, que se trata de una cadena semirregular, en la que existe una clase de paso $\{1,2\}$, y dos clases finales $\{3,4\}$ y $\{5,6\}$. Según lo expuesto anteriormente, la matriz de probabilidades estacionarias tiene el aspecto siguiente:

$$P^* = \begin{pmatrix} 0 & 0 & a & b & c & d \\ 0 & 0 & e & f & g & h \\ 0 & 0 & \pi_1 & \pi_2 & 0 & 0 \\ 0 & 0 & \pi_1 & \pi_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \pi_3 & \pi_4 \\ 0 & 0 & 0 & 0 & \pi_3 & \pi_4 \end{pmatrix}$$

Los valores no nulos de los estados pertenecientes a clases finales pueden hallarse resolviendo la clase final concreta. En el caso que nos ocupa, tenemos:

$$\begin{aligned} \pi_1 &= \pi_3 = 1/3 \\ \pi_2 &= \pi_4 = 2/3 \end{aligned}$$

Para las dos primeras filas, se trata de plantear, para las dos primeras filas de P , la ecuación matricial:

$$P \cdot P^* = P^*$$

El resultado es la matriz:

$$P^* = \begin{pmatrix} 0 & 0 & 0,3162 & 0,6324 & 0,0171 & 0,0342 \\ 0 & 0 & 0,2949 & 0,5897 & 0,0385 & 0,0769 \\ 0 & 0 & 0,3333 & 0,6667 & 0 & 0 \\ 0 & 0 & 0,3333 & 0,6667 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,3333 & 0,6667 \\ 0 & 0 & 0 & 0 & 0,3333 & 0,6667 \end{pmatrix}$$

3.4.4 Probabilidades estacionarias para cadenas con clases finales cíclicas

En las cadenas regulares y semirregulares se cumple que:

$$\mathbf{P}^* = \lim_{n \rightarrow \infty} \mathbf{P}^n$$

En las cadenas cíclicas, por su naturaleza, no existe este límite. La cadena evoluciona hacia tantas matrices de transición a largo plazo como el periodo p de la cadena, por lo que tendremos un conjunto de matrices:

$$\mathbf{P}^{*(1)}, \mathbf{P}^{*(2)}, \dots, \mathbf{P}^{*(p)}$$

Sin embargo, el hecho de que no exista un límite de las potencias de \mathbf{P} no significa que no exista \mathbf{P}^* entendida como la matriz de probabilidades estacionarias. Afortunadamente, el proceso indicado para las cadenas regulares y semirregulares sirve también aquí, por lo que no hace falta repetirlo. La diferencia es que el significado es ahora diferente: en lugar del límite de las potencias de \mathbf{P} es el promedio de las matrices de evolución a largo plazo de cada uno de los periodos.

$$\mathbf{P}^* = \frac{1}{p} \cdot [\mathbf{P}^{*(1)} + \mathbf{P}^{*(2)} + \dots + \mathbf{P}^{*(p)}]$$

Ejemplo 3.4.4.a Matriz \mathbf{P}^* para una clase final cíclica

Como se ha visto anteriormente, la cadena C_4 , cuya matriz de probabilidades de transición de un paso se muestra a continuación, cuenta con una clase cíclica de $p=2$.

$$\mathbf{P}_4 = \begin{vmatrix} 0,5 & 0,4 & 0,1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0,2 & 0 & 0,8 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

Para hallar la matriz de probabilidades estacionarias de \mathbf{P}_4 procedemos formalmente como si la clase final no fuese cíclica; entonces tendría las filas iguales: se trata de una cadena regular semiergódica. La diferencia es que ahora no se trata del límite de las potencias de \mathbf{P}_4 , sino del promedio de las dos matrices de probabilidades a las que tiende el comportamiento del sistema:

$$\mathbf{P}^* = \frac{1}{2} \cdot \mathbf{P}^{*(2k)} + \frac{1}{2} \cdot \mathbf{P}^{*(2k+1)}$$

$$\begin{vmatrix} 0 & 0,1 & 0,5 & 0,4 \\ 0 & 0,1 & 0,5 & 0,4 \\ 0 & 0,1 & 0,5 & 0,4 \\ 0 & 0,1 & 0,5 & 0,4 \end{vmatrix} = \frac{1}{2} \cdot \begin{vmatrix} 0 & 0,08 & 0,6 & 0,32 \\ 0 & 0,2 & 0 & 0,8 \\ 0 & 0 & 1 & 0 \\ 0 & 0,2 & 0 & 0,8 \end{vmatrix} + \frac{1}{2} \cdot \begin{vmatrix} 0 & 0,12 & 0,4 & 0,48 \\ 0 & 0 & 1 & 0 \\ 0 & 0,2 & 0 & 0,8 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

Para determinar las matrices \mathbf{P}^{nk} basta tener en cuenta las igualdades:

$$\mathbf{P}^* = 1/p (\mathbf{P}^k + \mathbf{P}^{k+1} + \dots + \mathbf{P}^{k-1})$$

$$\mathbf{P}^{k+1} = \mathbf{P}^k \cdot \mathbf{P}; \mathbf{P}^{k+2} = \mathbf{P}^{k+1} \cdot \mathbf{P}; \dots; \mathbf{P}^k = \mathbf{P}^{k-1} \cdot \mathbf{P}$$

Hay que darse cuenta aquí de la posibilidad de error que existe si se clasifica la cadena simplemente calculando potencias de \mathbf{P} , tentación en la que es fácil de caer disponiendo de un ordenador con una hoja de cálculo. Lo más fácil es elevar \mathbf{P} al cuadrado, luego hacer lo mismo con \mathbf{P}^2 y así sucesivamente. Entonces no se verá la aparición de ciclos de longitud 2, como en este caso. Compruébese que en la potencia \mathbf{P}^{16} ya se obtiene \mathbf{P}^{2k} , por lo que esta matriz se puede confundir con \mathbf{P}^* , cosa evidentemente errónea.

Ejemplo 3.4.4.b Matriz \mathbf{P}^* para una cadena mixta

Sea la cadena de Markov definida por la siguiente matriz de probabilidades de transición:

$$\mathbf{P} = \begin{array}{c} \begin{array}{cccccccccccccccc} 0.7 & 0.2 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.9 & 1/20 & 0 & 0 & 1/20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.8 & 0.1 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3/4 & 0 & 0.1 & 0 & 0 & 0 & 0 & 0.1 & 0 & 1/20 & 0 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{0} & \boxed{2/3} & \boxed{1/6} & \boxed{1/6} & \boxed{0} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} & \boxed{0} & \boxed{0} & \boxed{0} & \boxed{0} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{0} & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{0.7} & \boxed{0.3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{0} & \boxed{0.4} & \boxed{0.6} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & \boxed{0} & \boxed{0} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & \boxed{0} & \boxed{0} \end{array} \end{array}$$

Se observa que se trata de una cadena mixta (se han recuadrado las submatrices ergódicas correspondientes a las cuatro clases finales).

Obsérvese que el estado 5 es un estado absorbente, y la clase {11,12} es aperiódica, mientras que las otras dos tienen ciclos de longitud 3 y 2 respectivamente.

El lector debe hacer por su cuenta los cálculos para determinar los valores de \mathbf{P}^* y comprobará que obtiene los siguientes valores (se han suprimido las columnas de clases de paso):

$$\mathbf{P} = \begin{array}{c} \begin{array}{cccccccccccc} 23/42 & 19/315 & 38/945 & 19/1990 & 19/1990 & 19/315 & 19/255 & 38/357 & 19/420 & 19/1050 & 19/700 \\ 4/7 & 2/35 & 4/105 & 1/105 & 1/105 & 2/35 & 6/85 & 12/119 & 3/70 & 3/175 & 3/350 \\ 1/2 & 1/15 & 2/45 & 1/90 & 1/90 & 1/15 & 7/85 & 2/17 & 1/20 & 1/50 & 3/100 \\ 0 & 2/15 & 4/45 & 1/45 & 1/45 & 2/15 & 14/85 & 4/17 & 1/10 & 1/25 & 3/50 \\ \boxed{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \boxed{1/3} & \boxed{2/9} & \boxed{1/18} & \boxed{1/18} & \boxed{1/3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \boxed{1/3} & \boxed{2/9} & \boxed{1/18} & \boxed{1/18} & \boxed{1/3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \boxed{1/3} & \boxed{2/9} & \boxed{1/18} & \boxed{1/18} & \boxed{1/3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \boxed{1/3} & \boxed{2/9} & \boxed{1/18} & \boxed{1/18} & \boxed{1/3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{7/17} & \boxed{10/17} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{7/17} & \boxed{10/17} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1/2} & \boxed{1/5} & \boxed{3/10} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1/2} & \boxed{1/5} & \boxed{3/10} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1/2} & \boxed{1/5} & \boxed{3/10} \end{array} \end{array}$$

Sin embargo, el significado de estas probabilidades es diferente según el estado (fila) de la que se trate.

Si se introduce la matriz \mathbf{P} en una hoja de cálculo y se eleva al cuadrado unas cuantas veces (con seis veces será suficiente) se observará como la fila 5 no varía, las filas 11 y 12 tienden hacia los valores finales más o menos rápidamente, mientras que en las filas 13-15 se repiten los mismos valores ya desde \mathbf{P}^2 . En cambio, en las filas 6-10 se observa como los mismos valores se van "moviendo" cíclicamente de unas filas a otras sin que sin embargo se lleguen a observar todos los valores en todas las filas.

Esto es un efecto de los ciclos de las clases finales, que no se observan fácilmente si nos limitamos a calcular potencias muy elevadas de \mathbf{P} simplemente elevando al cuadrado. Solamente si se multiplica por \mathbf{P} para obtener la potencia siguiente se puede ver como en efecto los valores de las clases {6-10} y {13-15} se repiten cada tres y cada dos potencias, como corresponde a su naturaleza. De ahí la advertencia hecha al principio de no intentar clasificar la cadena mediante la "fuerza bruta" del ordenador, ya que ello puede llevar fácilmente a engaño.

También se ve ahora como el significado de \mathbf{P}^* es distinto: en la clase {11-12} son efectivamente probabilidades una vez perdida la memoria del estado inicial, mientras que en las dos clases cíclicas es un promedio de tres y dos potencias sucesivas respectivamente. Para los estados de paso se trata de una combinación según las leyes del cálculo de probabilidades de las diferentes situaciones finales que se pueden dar. Véase que si se inicia la evolución del sistema en el estado 4, la probabilidad de caer en el estado absorbente 5 es nula.

3.5 Cadenas de Markov con remuneración

Decimos que a una cadena de Markov definida por una matriz \mathbf{P} se le ha asociado una remuneración cuando a cada transición posible (esto es, con $p_{ij} > 0$) se le ha asociado un valor numérico arbitrario, r_{ij} . De esta forma, junto a la matriz \mathbf{P} ya conocida, tenemos ahora una matriz \mathbf{R} de remuneraciones.

$$\mathbf{P} = \begin{vmatrix} p_{11} & p_{1n} \\ & \\ & \\ p_{n1} & p_{nn} \end{vmatrix} \quad \mathbf{R} = \begin{vmatrix} r_{11} & r_{1n} \\ & \\ & \\ r_{n1} & r_{nn} \end{vmatrix}$$

La palabra remuneración no debe hacer presuponer nada sobre la naturaleza de este valor, que puede ser cualquier cosa según convenga al modelo: puede tener, en efecto, un significado de ganancias, beneficios etc. pero también puede tenerlo de costes, pérdidas o cualquier otra cosa. De hecho, es un modelo bastante habitual definirla de manera que su valor sirva para contar el número de veces que han sucedido determinadas transiciones, por ejemplo.

El problema al que nos enfrentamos ahora es el de encontrar la remuneración esperada de un proceso estocástico representable por una cadena de Markov, tanto a largo plazo como a corto plazo.

3.5.1 Remuneración esperada a corto plazo

En esta sección, estudiaremos cómo obtener la remuneración total esperada a corto plazo de un proceso estocástico que pueda ser representado por una cadena de Markov. Se trata de determinar dicha remuneración cuando el número de transiciones N es fijo, mediante la iteración en el espacio de

los estados. También se introducirá la normalización, procedimiento que ayuda a simplificar el proceso de cálculo.

3.5.1.1 Iteración en el espacio de los estados

Si $v_i(N)$ representa la remuneración total esperada cuando quedan N transiciones para finalizar el proceso, y si partimos del estado i , podemos escribir:

$$v_i(N) = \sum_{j=1}^n p_{ij} \cdot [r_{ij} + v_j(N-1)] = q_i + \sum_{j=1}^n p_{ij} \cdot v_j(N-1)$$

$$q_i = \sum_{j=1}^n p_{ij} \cdot r_{ij}$$

En las dos expresiones, j representa los n estados posibles que podrá tomar el sistema en la siguiente transición, en la que faltarán $N-1$ transiciones para acabar. En cuanto a la q_i , nótese que se obtiene como producto de la fila i de las matrices de probabilidad de transición \mathbf{P} y de remuneraciones \mathbf{R} (no se trata, entonces del producto de matrices $\mathbf{P} \cdot \mathbf{R}$, sino que matricialmente se escribiría $\mathbf{P} \cdot \mathbf{R}^T$). Representa la remuneración esperada de la transición del estado N al estado $N-1$ y es función del estado actual i .

Si queremos obtener todos los $v_i(N)$, es decir, las remuneraciones esperadas para el sistema cuando queden N iteraciones para acabar partiendo desde cualquier estado i inicial, podemos utilizar la misma ecuación, pero esta vez en forma vectorial:

$$\mathbf{v}(N) = \mathbf{q} + \mathbf{P} \cdot \mathbf{v}(N-1)$$

Esta expresión recursiva (se obtiene el valor del estado N a partir del valor para $N-1$) nos permite obtener las remuneraciones esperadas del sistema para valores cada vez más alejados del instante final conociendo unas condiciones de contorno que definen la remuneración esperada *después* del final del proceso $\mathbf{v}(0)$.

La aplicación de esta ecuación constituye la *iteración en el espacio de los estados*.

3.5.1.2 Normalización

La técnica de la normalización consiste en restar sistemáticamente al valor obtenido en cada transición una cantidad determinada y operar en lo sucesivo con la diferencia. Por ello las ecuaciones de recurrencia quedan la forma siguiente:

$$\begin{aligned} \mathbf{v}(N) &= \mathbf{q} + \mathbf{P} \cdot \mathbf{v}'(N-1) \\ \mathbf{v}'(N) &= \mathbf{v}(N) - \mathbf{g}_N \end{aligned}$$

donde \mathbf{g}_N es un vector con todas las componentes iguales.

Obsérvese que una vez hechos los cálculos con normalización, obtener el valor verdadero de la remuneración esperada es muy sencillo:

$$\mathbf{v}^R(N) = \mathbf{v}'(n) + \mathbf{g}_N + \mathbf{g}_{N-1} + \dots + \mathbf{g}_1$$

Esta técnica tiene dos ventajas: en primer lugar evita que los valores de $\mathbf{v}(N)$ crezcan de forma desmesurada, lo cual tiene ventajas indiscutibles para el cálculo, y en segundo lugar permite observar con más facilidad la llegada del régimen permanente, puesto que en el momento que en dos transiciones (consecutivas o no) se repitan los valores tanto de la \mathbf{g}_N como de $\mathbf{v}(N)$ podemos afirmar que se ha superado el régimen transitorio y se ha alcanzado el régimen permanente.

Queda por determinar como se fija la \mathbf{g}_N , valor que tendrán todas las componentes de \mathbf{g}_N : en principio dicho valor es arbitrario, y lo importante es que se determine según una regla fija. Las dos más habituales son:

Regla 1:
$$g_N = \min_i \{v_i(N)\}$$

Regla 2:
$$g_N = v_k(N)$$
, donde k es un estado elegido previamente.

Ambas reglas conducen a largo plazo a los mismos resultados (salvo unas diferencias constantes), pues en régimen permanente el estado de mínimo valor será siempre el mismo. La Regla 1 tiene la ventaja de que se asegura que los valores de $\mathbf{v}'(N)$ son siempre mayores o iguales a cero, lo cual evita errores en los cálculos manuales, mientras que la Regla 2 facilita la comparación de los vectores \mathbf{v}'

Ejemplo 3.5.1.2.a Iteración en el espacio de los estados y normalización

La tabla siguiente muestra las sucesivas iteraciones realizadas para determinar el comportamiento de la cadena C1, con las matrices de probabilidades de transición \mathbf{P} y de remuneraciones \mathbf{R} siguientes:

$$\mathbf{P} = \begin{vmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{vmatrix} \quad \mathbf{R} = \begin{vmatrix} 10 & 20 \\ 40 & -20 \end{vmatrix}$$

Las filas segunda y tercera muestran el vector \mathbf{v} , obtenido a través de sucesivas normalizaciones, y la quinta y sexta \mathbf{v}' . La última fila es la correspondiente a \mathbf{g}_N , obtenida siempre como el valor más pequeño de \mathbf{v} .

	0	1	2	3	4	5	6	7	8	9	10	11
v_1	0.00	18.00	18.40	18.24	18.30	18.28	18.29	18.28	18.29	18.29	18.29	18.29
v_2	0.00	16.00	17.20	16.72	16.91	16.84	16.87	16.85	16.86	16.86	16.86	16.86
	0	1	2	3	4	5	6	7	8	9	10	11
v'_1	0	2.00	1.20	1.52	1.39	1.44	1.42	1.43	1.43	1.43	1.43	1.43
v'_2	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
g_N	0	16.00	17.20	16.72	16.91	16.84	16.87	16.85	16.86	16.86	16.86	16.86

Como se ve fácilmente en la tabla normalizada, a partir de la séptima transición aparece ya el régimen permanente. Dicha observación no es en absoluto evidente si sólo disponemos de la tabla sin normalizar.

3.5.2 Remuneración esperada a largo plazo

Mediante la iteración en el espacio de los estados se puede calcular la remuneración esperada en un número cualquiera de transiciones, de manera que teóricamente basta repetir dicha iteración hasta llegar al número deseado; sin embargo, las propiedades que se obtuvieron en la primera parte del

módulo hacen pensar que el sistema tenderá a tener una remuneración promedio (remuneración esperada dividida por el número de transiciones) estable a largo plazo.

Dicha remuneración dependerá fundamentalmente de la remuneración obtenida en la clase final a la que finalmente vaya a parar la evolución del sistema, ya que es en ella donde se producirán la mayoría de las transiciones (supuesto el número de estas suficientemente grande); sin embargo, es obvio que dicho valor promedio ha de venir afectado por un factor de corrección que dependa del estado i en el que se inicie el proceso. Podremos representar este hecho haciendo que, para un número de transiciones lo bastante elevado:

$$v_i^*(N) \rightarrow Ng_i + w_i$$

$$\mathbf{v}^*(N) \rightarrow N\mathbf{g} + \mathbf{w}$$

En estas expresiones, g representa la ganancia media por transición, y w_i un factor de corrección función del estado inicial i .

Si sustituimos las $v_i(N)$ por estas expresiones en las ecuaciones de iteración en el espacio de los estados, encontramos el siguiente resultado:

$$v_i(N) = q_i + \sum_{j=1}^n p_{ij} \cdot v_j(N-1)$$

$$Ng_i + w_i = q_i + \sum_{j=1}^n p_{ij} \cdot [(N-1)g_j + w_j] = q_i + (N-1) \cdot \sum_{j=1}^n p_{ij} \cdot g_j + \sum_{j=1}^n p_{ij} \cdot w_j$$

La ganancia media por transición para el estado inicial i puede obtenerse a partir de las ganancias medias del estado siguiente:

$$g_i = \sum_{j=1}^n p_{ij} \cdot g_j$$

Esto nos permite obtener la siguiente ecuación, para cada uno de los estados:

$$w_i + g_i = q_i + \sum_{j=1}^n p_{ij} \cdot w_j$$

La ecuación puede expresarse también en forma vectorial, para el conjunto de estados i :

$$\mathbf{w} + \mathbf{g} = \mathbf{q} + \mathbf{P}\mathbf{w}$$

En principio, tenemos un sistema altamente indeterminado, puesto que tenemos n ecuaciones para $2n$ incógnitas. Sin embargo, hemos de tener en cuenta dos cosas:

- La indeterminación afecta a las w_i , puesto que dichas variables no tienen sentido por ellas mismas, sino que lo tienen sus diferencias. Efectivamente, podemos escribir:

$$w_i - w_j = v_i^*(N) - v_j^*(N)$$

- La estructura del vector \mathbf{g} depende del tipo de la cadena de Markov. Puede deducirse la forma de \mathbf{g} a partir de la matriz de probabilidades estacionarias:

$$\mathbf{g} = \mathbf{P}^* \cdot \mathbf{q} = \mathbf{P} \cdot \mathbf{g}$$

No resulta ventajoso obtener \mathbf{g} de esta manera, puesto que puede obtenerse a partir de las ecuaciones obtenidas anteriormente con menos esfuerzo. Sin embargo, a través de esta ecuación podemos conocer cuántas componentes diferentes tendrá el vector \mathbf{g} .

3.5.2.1 Remuneración esperada a largo plazo en cadenas regulares

Consideraremos por el momento el caso sin actualización (o con $\alpha = 1$ si se prefiere). El caso con actualización se trata al final para todos los tipos de cadenas conjuntamente.

Si la cadena es regular, todas las filas de \mathbf{P}^* son iguales, por lo que todos los componentes del vector \mathbf{g} son iguales:

$$\text{Cadena regular} \Rightarrow g_1 = g_2 = \dots = g_n = g$$

Por lo tanto, para una cadena regular tendremos un sistema de n ecuaciones con $n+1$ incógnitas (la g y las w_i), es decir, con un grado de libertad que nos permite fijar arbitrariamente el valor de una de las w_i . Lo más conveniente es hacer una $w_i = 0$ para simplificar los cálculos, y así obtener el resto de ecuaciones.

El hecho de fijar a cero uno u otro de los valores no tiene una significación arbitraria, ya que equivale a la determinación del estado para el cual se hace cero el valor $v'(N)$ al normalizar. En efecto, aunque los valores de w_i no están determinados debido al grado de libertad del sistema, sí lo están sus diferencias:

$$w_i - w_j = \lim_{N \rightarrow \infty} \{v'_i(N) - v'_j(N)\}$$

Además, como que:

$$v_i^*(N) \rightarrow Ng_i + w_i$$

esto puede entenderse de la siguiente forma:

Como ya se ha dicho, g_i representa el valor promedio de la remuneración obtenida en las diferentes transiciones de la cadena dentro de la clase final.

Aceptemos ahora que se ha fijado un estado k de la cadena tal que $w_k = 0$, y que k pertenece a la clase final (si la cadena es ergódica tal cosa es obligada). Si la evolución del sistema parte de un estado i , habrá un cierto número de transiciones (v) hasta que pase por primera vez por el estado k durante las cuales se obtendrá una remuneración esperada $v_i(v)$. Pues bien, w_i representa la diferencia entre esta remuneración y la que correspondería según el valor medio:

$$w_i = v_i(v) - v \cdot g_i$$

Evidentemente, si tomamos el estado k fuera de la clase final o el valor fijado en uso del grado de libertad del sistema, ello no significa más que un simple corrimiento del "nivel de referencia".

Todo lo expuesto en este apartado es válido para toda cadena regular, sea ergódica o semiergódica.

Ejemplo 3.5.2.1.a.:

Se desea obtener la ganancia media que se obtendrá con la cadena ergódica de matriz \mathbf{P} que a continuación se indica, si las remuneraciones asociadas a cada transición se indican en la matriz \mathbf{R} . En este caso, la matriz \mathbf{R} se ha obtenido del siguiente modo:

- Si el proceso permanece en el mismo estado i , se remunera con una cantidad igual a i .
- Si el proceso pasa de un estado i a otro j , se remunera o penaliza con una cantidad igual a $j - i$.

$$\mathbf{P} = \begin{pmatrix} 0 & 0,2 & 0,4 & 0,3 & 0,1 \\ 0,5 & 0,5 & 0 & 0 & 0 \\ 0 & 0,2 & 0,4 & 0 & 0,4 \\ 0 & 0,7 & 0 & 0 & 0,3 \\ 0 & 0 & 0,8 & 0,1 & 0,1 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 \\ -1 & 2 & 1 & 2 & 3 \\ -2 & -1 & 3 & 1 & 2 \\ -3 & -2 & -1 & 4 & 1 \\ -4 & -3 & -2 & -1 & 5 \end{pmatrix}$$

El vector \mathbf{q} se obtiene multiplicando las filas de \mathbf{P} y \mathbf{R} (una vez más debemos insistir en el hecho de que no se trata del producto de ambas matrices). Para el caso que nos ocupa, el valor de este vector es:

$$\mathbf{q} = \begin{pmatrix} 2,3 \\ 0,5 \\ 1,8 \\ -1,1 \\ -1,2 \end{pmatrix}$$

Finalmente, a partir de la ecuación matricial:

$$\mathbf{w} + \mathbf{g} = \mathbf{q} + \mathbf{P}\mathbf{w}$$

obtenemos el conjunto de ecuaciones siguiente, en el que todos los valores de las componentes de \mathbf{g} son iguales a un valor g :

$$\begin{aligned} w_1 + g &= 2,3 + 0,2w_2 + 0,4w_3 + 0,3w_4 + 0,1w_5 \\ w_2 + g &= 0,5 + 0,5w_1 + 0,5w_2 \\ w_3 + g &= 1,8 + 0,2w_2 + 0,4w_3 + 0,4w_5 \\ w_4 + g &= -1,1 + 0,7w_2 + 0,3w_5 \\ w_5 + g &= -1,2 + 0,8w_3 + 0,1w_4 + 0,1w_5 \end{aligned}$$

Tenemos cinco ecuaciones y seis incógnitas. Por lo tanto, podemos hacer $w_2 = 0$, y entonces tendremos los siguientes valores:

$$\begin{aligned}
 w_1 &= 0,5545 \\
 w_2 &= 0 \\
 w_3 &= 0,1178 \\
 w_4 &= -2,5913 \\
 w_5 &= -2,3801 \\
 g &= 0,7772
 \end{aligned}$$

Los resultados indican que la ganancia obtenida en el largo plazo, tras N transiciones, es aproximadamente de $0,7772N$ (téngase en cuenta que como N es muy grande, cualquier variación finita es irrelevante ante este valor: para un número lo suficientemente grande de transiciones, la ganancia *media* es igual a $g = 0,7772$). En cambio, si empezamos en el estado 4 el valor esperado de dicha ganancia es 2,5913 unidades menor que si empezamos en el estado 2.

3.5.2.3 Remuneración esperada en cadenas cíclicas

El cálculo para una cadena cíclica es formalmente idéntico al de una cadena regular, si bien el significado de los resultados es diferente. Los valores obtenidos son promedios de los valores que se presentan de forma cíclica con el número de transiciones.

Ejemplo 3.5.2.3.a Remuneración esperada para una cadena cíclica de $p=2$

Consideremos ahora una cadena cíclica, de $p = 2$, tal como:

$$\mathbf{P} = \begin{vmatrix} 0 & 0,5 & 0 & 0,5 \\ 0,5 & 0 & 0,5 & 0 \\ 0 & 0,5 & 0 & 0,5 \\ 0,5 & 0 & 0,5 & 0 \end{vmatrix}$$

Planteemos el problema de obtener \mathbf{g} y \mathbf{w} con la matriz de remuneraciones:

$$\mathbf{R} = \begin{vmatrix} - & 1 & - & 3 \\ 4 & - & 6 & - \\ - & 2 & - & 0 \\ 2 & - & 8 & - \end{vmatrix}$$

Para estas \mathbf{P} y \mathbf{R} , tenemos que \mathbf{q} vale:

$$\mathbf{q} = \begin{vmatrix} 2 \\ -1 \\ 1 \\ 3 \end{vmatrix}$$

De manera que, prescindiendo del hecho de que la cadena sea cíclica, y dado que tiene una única clase final, podemos plantear las ecuaciones:

$$\begin{aligned}
 w_1 + g &= 2 + (w_2 + w_4)/2 \\
 w_2 + g &= -1 + (w_1 + w_3)/2
 \end{aligned}$$

$$\begin{aligned}w_3 + g &= 1 + (w_2 + w_4)/2 \\w_4 + g &= 3 + (w_1 + w_3)/2\end{aligned}$$

De donde obtenemos:

$$\begin{aligned}g &= 5/4 \\w_1 &= 0 \\w_2 &= -11/4 \\w_3 &= -1 \\w_4 &= 5/4\end{aligned}$$

Estos valores describen el comportamiento *promedio* del sistema en el largo plazo, exactamente del mismo modo que \mathbf{P}^* describe las propiedades estacionarias para una cadena periódica. Podemos conocer con más detalle el comportamiento del sistema si realizamos iteraciones en el espacio de las políticas, a partir de la ecuación:

$$\mathbf{v}(N) = \mathbf{q} + \mathbf{P} \cdot \mathbf{v}(N-1)$$

El lector puede comprobar, por inducción, que la ganancia total del sistema cuando faltan N etapas para terminar tiene el comportamiento siguiente:

	$N = 2k$	$N = 2k+1$
v_1	$(5/4)N + 1/2$	$(5/4)N + 3/4$
v_2	$(5/4)N + 2$	$(5/4)N - 9/4$
v_3	$(5/4)N - 1/2$	$(5/4)N - 1/4$
v_4	$(5/4)N + 2$	$(5/4)N + 7/4$

De la tabla se deduce que, tal como habíamos hallado, $g = 5/4$. Las diferencias entre componentes de \mathbf{v} se comportan de modo diferente en periodos pares que en periodos impares, y las w_i obtenidas con las ecuaciones muestran el comportamiento promedio:

	$N = 2k$	$N = 2k + 1$	$w_i - w_1$
$v_1 - v_1$	0	0	0
$v_2 - v_1$	-5/2	-3	-11/4
$v_3 - v_1$	-1	-1	-1
$v_4 - v_1$	3/2	5/4	5/4

3.5.2.4 Remuneración esperada en cadenas no ergódicas

En general, una cadena no ergódica tendrá:

CF_1, CF_2, \dots, CF_m clases finales (sean cíclicas o acíclicas)
 $ecp_1, ecp_2, \dots, ecp_r$ estados pertenecientes a clases de paso

Entonces el vector g tendrá $m + r$ componentes diferentes. Todos los estados de una misma clase final tendrán el mismo valor de g , y los estados de las clases de pasos tendrán un valor de g diferente cada uno:

Cadena no ergódica $\Rightarrow g_{CF1}, g_{CF2}, \dots, g_{CFm}, g_{ecp1}, g_{ecp2}, \dots, g_{ecpr}$ valores distintos

Tenemos entonces un sistema con n ecuaciones y $n + m + r$ incógnitas. El sistema quedará determinado si se hacen $m + r$ valores de $w_i=0$.

Ejemplo 3.5.2.4.a Remuneración esperada para una cadena semirregular

Consideremos la matriz de transición \mathbf{P} de una cadena semirregular y las componentes activas de la matriz de remuneraciones \mathbf{R} siguientes:

$$\mathbf{P} = \begin{vmatrix} 0,1 & 0,2 & 0,3 & 0,4 & 0 \\ 0 & 0,4 & 0,6 & 0 & 0 \\ 0 & 0,7 & 0,3 & 0 & 0 \\ 0 & 0 & 0 & 0,3 & 0,7 \\ 0 & 0 & 0 & 0,8 & 0,2 \end{vmatrix}$$

$$\mathbf{R} = \begin{vmatrix} 4 & 7 & -8 & 9 & - \\ - & 8 & 3 & - & - \\ - & 2 & -13 & - & - \\ - & - & - & 6 & -4 \\ - & - & - & 6 & -4 \end{vmatrix}$$

En este caso, el vector de ganancias medias por transición \mathbf{g} tendrá tres componentes diferentes:

- Un componente g_1 asociado a la clase de paso 1.
- Un componente g_A asociado a la clase final $A = \{2,3\}$.
- Un componente g_B asociado a la clase final $B = \{4,5\}$.

En primer lugar, multiplicando las filas de \mathbf{P} y \mathbf{R} obtenemos el vector \mathbf{q} :

$$\mathbf{q} = \begin{vmatrix} 3 \\ 5 \\ -2,5 \\ -1 \\ 4 \end{vmatrix}$$

Podemos, por ejemplo, plantear las ecuaciones $\mathbf{g} + \mathbf{w} = \mathbf{q} + \mathbf{Pw}$ para la clase final A:

$$\begin{aligned} w_2 + g_A &= 5 + 0,4w_2 + 0,6w_3 \\ w_3 + g_A &= -2,5 + 0,7w_2 + 0,3w_3 \end{aligned}$$

De estas ecuaciones obtenemos:

$$\begin{aligned} g_A &= 20/13 \\ w_2 &= 0 \\ w_3 &= -75/13 \end{aligned}$$

Planteando las ecuaciones para la clase final B, tenemos:

$$\begin{aligned}w_4 + g_B &= -1 + 0,3w_4 + 0,7w_5 \\w_5 + g_B &= 4 + 0,8w_4 + 0,2w_5\end{aligned}$$

$$\begin{aligned}g_B &= 4/3 \\w_4 &= 0 \\w_5 &= 10/3\end{aligned}$$

Ahora quedan por encontrar los valores de la clase de paso, g_1 y w_1 . Para ello contamos con la ecuación:

$$w_1 + g_1 = 3 + 0,1w_1 + 0,2w_2 + 0,3w_3 + 0,4w_4$$

Ahora no podemos encontrar g_1 haciendo $w_1 = 0$. Ambos valores están determinados, porque hemos agotado los grados de libertad del sistema al hacer $w_2=0$ y $w_4=0$. Por lo tanto, necesitamos otra ecuación. Dicha ecuación puede encontrarse a partir de las ecuaciones matriciales que \mathbf{g} satisface:

$$\begin{aligned}\mathbf{P}^* \cdot \mathbf{q} &= \mathbf{g} \\ \mathbf{P} \cdot \mathbf{g} &= \mathbf{g}\end{aligned}$$

La opción más económica es la segunda, que nos exonera de obtener \mathbf{P}^* . Desarrollando para la primera fila, tenemos:

$$0,1g_1 + 0,2g_A + 0,3g_A + 0,4g_B = g_1$$

De donde obtenemos de manera directa que $g_1 = 508/351$. Volviendo a la ecuación $g_1 + w_1 = \dots$ obtenemos que $w_1 = 625/3159$.

El significado de los valores de \mathbf{g} y \mathbf{w} para los estados de paso se puede deducir de las expresiones utilizadas para hallarlos: en primer lugar g_1 en la ecuación anterior resulta ser un promedio de los valores g_A , g_B de las diferentes clases finales; ello proviene del hecho que, a largo plazo, lo que contará como remuneración de la cadena es precisamente el valor de g que corresponda a la clase final donde vaya a parar; por ello, la g de los estados de paso es la esperanza matemática de dichos valores, de modo que si en la ecuación anterior aislamos g_1 resulta:

$g_1 = 5/9 g_A + 4/9 g_B$, lo cual nos dice que las probabilidades de que el sistema acabe yendo a la clase final A o a la clase final B si inicia su evolución en el estado 1 son 5/9 y 4/9 respectivamente.

Por lo que respecta a las w_i , su significado es el mismo ya visto en las cadenas ergódicas y semi-ergódicas; la única diferencia es que en este caso, al depender de dos grados de libertad, es un valor referido a los niveles fijados en cada clase final; recuérdese que ello no significa que el valor real de la remuneración sea exactamente $v_i = N g + w_i$, sino solamente que $v_i - v_j = w_i - w_j$.

3.6 Cadenas de Markov con remuneración y decisión

Se dice que un proceso markoviano tiene decisión si en cada transición se puede fijar una variable (llamada de *decisión*) mediante la cual se puede elegir entre diferentes juegos de probabilidades de transición, y eventualmente entre diferentes valores de la remuneración asociada. De esta forma, tanto la matriz de probabilidades \mathbf{P} como la de remuneraciones \mathbf{R} , tienen tres índices o dimensiones: el estado anterior, el posterior y la decisión. Por ello notamos sus elementos como p_{iju} y r_{iju} , donde u representa la variable de decisión. Así, los valores posibles que podrán tomar los subíndices serán:

Subíndice i (estados de origen):	$i = 1, 2, \dots, n$
Subíndice j (estados de destino):	$j = 1, 2, \dots, n$
Subíndice u (decisiones):	$u = 1, 2, \dots, d_i$

Como sucede a lo largo del texto, n es el número de estados de la cadena. El valor d_i es el número de decisiones posibles, que depende del estado de partida i .

Se llama *política* a una regla que fija para cada estado de la cadena el valor de la decisión a adoptar. Podrá expresarse entonces de forma vectorial:

$$\mathbf{p}^T = \{u_1, u_2, \dots, u_i, \dots, u_n\}$$

donde las u_i es el valor de la variable de decisión fijado por la política para el estado i .

Una vez determinada una política, el sistema se transforma en una cadena de Markov con remuneración como las que se han visto en el apartado anterior. Es importante tener en cuenta que mientras no se ha fijado una política de hecho el sistema no responde propiamente al modelo de las cadenas tal como se ha visto, ya que ni las probabilidades de transición ni las remuneraciones están determinadas. Por ello, hasta que no se haya fijado una política, no tiene sentido hablar de su clasificación ni de probabilidades estacionarias ni de remuneraciones esperadas.

Si existe capacidad de decisión es porque se pretende que la remuneración obtenida sea lo mejor posible, lo cual puede significar máxima o mínima según la naturaleza de la remuneración: de esta manera se puede hablar de política óptima, que es aquella que permite obtener la mejor remuneración esperada. Téngase en cuenta que por la naturaleza aleatoria del sistema no se puede optimizar la remuneración *obtenida realmente* (que es una magnitud aleatoria), sino tan sólo el valor esperado de la misma.

La política óptima no tiene sentido si no es con relación a un horizonte determinado. Se habla así de *política óptima a corto plazo* (esto es con un horizonte finito y determinado) o *política óptima a largo plazo*, es decir, con un horizonte ilimitado o con un horizonte finito pero desconocido.

3.6.1 Procesos markovianos de decisión

Algunos procesos polietápicos de decisión pueden modelizarse como cadenas de Markov con remuneración y decisión. Para que esto sea posible, deben tener las características siguientes:

- El sistema puede evolucionar dentro de un número finito de estados n .
- La evolución del estado en el futuro depende exclusivamente de:
 - El estado i en el que se encuentra en el momento presente (esto es, el sistema cumple la propiedad markoviana).
 - Del valor de la variable de decisión escogido para el estado i , de entre los d_i valores posibles de dicha variable de decisión.

La evolución del sistema, pues, puede definirse como una secuencia de evolución (E) según un proceso markoviano y decisión (D) que condiciona la evolución a la etapa siguiente: este tipo se sistemas pueden caracterizarse como sistemas E / D.

Los datos de partida para resolver el sistema son:

- La matriz de probabilidades de transición, para cada una de las políticas $\mathbf{P}^{(p)}$. Sus componentes serán del tipo p_{iju} , que no es más que la probabilidad de transición del estado i al estado j , una vez se ha tomado la decisión u desde el estado i .
- La matriz de remuneraciones $\mathbf{R}^{(p)}$, cuyos componentes son r_{iju} , definidos como la remuneración (ganancia o pérdida) obtenida al pasar del estado i al estado j , una vez se ha tomado la decisión u desde el estado i .

Definido el problema en estos términos, es importante destacar que sólo cuando hayamos establecido una política determinada, tendremos un proceso markoviano con remuneración.

El objetivo es ahora el de establecer la política óptima, de manera que se encuentre la política que optimiza la ganancia esperada. Existen para ello dos métodos:

La *iteración en el espacio de los estados* consiste en encontrar la política que nos da la $\mathbf{v}(N)$ óptima, esto es, $\mathbf{v}^*(N)$, a partir de $\mathbf{v}^*(N-1)$. Este método puede emplearse para obtener la política óptima para un número finito de etapas (problema de *horizonte infinito*), y si se observa que el sistema converge a una determinada política después de un número determinado de etapas, la política óptima a largo plazo (problema de *horizonte finito*).

La *exploración del espacio de las políticas*, consistente en encontrar las \mathbf{g} y las \mathbf{w} para cada una de las posibles políticas del sistema. Por su propia naturaleza, se trata de una metodología para resolver el problema de *horizonte infinito*.

Seguidamente se describirán cada una de estas técnicas usando diversos ejemplos prototipo.

3.6.2 Iteración en el espacio de los estados

Como se ha expuesto anteriormente, se trata de encontrar $\mathbf{v}^*(N)$, una vez obtenida $\mathbf{v}^*(N-1)$. Para ello, se hará uso de la ecuación de recurrencia¹ definida para las cadenas de Markov con remuneración, teniendo en cuenta ahora que, para cada uno de los estados origen i , hemos de valorar las políticas posibles. En primer lugar, calcularemos la remuneración esperada si partimos del estado i y tomamos la decisión u :

$$v_{iu}(N) = q_{iu} + \sum_{j=1}^n p_{iju} \cdot v_j^*(N-1)$$

Una vez calculados todos estos valores, podremos encontrar el valor óptimo (máximo o mínimo, según convenga) para el estado i . Una vez hayamos obtenido las decisiones óptimas para todos los estados, quedará definida la política óptima para esa etapa:

$$v_i^*(N) = \underset{u}{OPT} \{v_{iu}(N)\}$$

Con esta ecuación de recurrencia, obtendremos, para la etapa N , la política óptima a seguir si en ese momento nos encontramos en el estado i .

¹ La iteración en el espacio de los estados es, de hecho, un caso particular de programación dinámica aleatoria. En el siguiente módulo, se desarrollará con detalle la problemática de la programación dinámica, en un contexto más general.

También podemos estudiar el comportamiento a largo plazo, si normalizamos los resultados en cada iteración, tal como se describe en la sección 5.2.2. Podremos obtener conclusiones acerca del comportamiento del sistema a largo plazo con este sistema si se cumplen las siguientes condiciones:

Se observa que, para un número suficientemente elevado de estados, la política óptima es siempre la misma.

Se observa que el sistema converge, esto es:

$$\begin{aligned} g_N^* &\rightarrow g_{N-1}^* \\ v^*(N) &\rightarrow v^*(N-1) \end{aligned}$$

Entonces podremos decir que hemos encontrado la política óptima p^* , teniendo además:

$$\begin{aligned} g_N^{(p)} &\rightarrow g^* \\ v^*(N) &\rightarrow w^* \end{aligned}$$

Si no observamos una convergencia clara, o bien el sistema escoge diversas políticas, podemos recurrir a la exploración del espacio de las políticas.

Ejemplo 3.6.2.a

Un taxista trabaja en una ciudad dividida en tres zonas: A, B y C. Situado en una zona determinada, puede seguir tres políticas:

Política 1: circular por la calle en busca de clientes.

Política 2: permanecer estacionado en una parada de taxis en espera del cliente.

Política 3: captar clientes por un servicio de radio taxi, al que el taxista está abonado. Dicho servicio no cubre la zona B.

La zona en que se encuentre el taxi en este momento y la política seguida, determinarán:

- Las probabilidades de que el destino de su próximo servicio sea la zona A, B o C.
- La remuneración (que depende de las características de los clientes captados con una política u otra, las comisiones a pagar al servicio de radio taxi, el combustible gastado circulando en busca de cliente, etc) obtenida al partir de una zona i y llegar a una zona j .

Dichas probabilidades, y remuneraciones, así como la ganancia esperada $q_i^{(p)}$ (obtenidas a partir de las $p_{ij}^{(p)}$ y las $r_{ij}^{(p)}$), se detallan en la tabla adjunta:

ZONA	política	A	B	C	A	B	C	$q_i^{(p)}$
A	1	0,5	0,2	0,3	10	3	8	8
	2	0,1	0,7	0,2	8	2	4	3
	3	0,3	0,1	0,6	3	7	4	4
B	1	0,5	0	0,5	14	0	18	16
	2	0,1	0,8	0,1	6	16	6	14
C	1	0,2	0,3	0,5	12	2	8	7
	2	0,1	0,8	0,1	6	4	2	4
	3	0,7	0,1	0,2	4	0	6	4

Se pide determinar la política óptima para el taxista a largo plazo.

Resolución:

Claramente nos encontramos dentro del marco de los modelos de Markov con remuneración y decisión, dado que la evolución del sistema en la siguiente transición depende de:

- La zona (que será el estado del sistema) en la que se encuentre el taxista en el momento actual: el sistema es de tipo markoviano.
- La política seguida por el taxista.

Lo que deseamos obtener es la política a seguir si nos encontramos en un estado determinado. Así, por ejemplo, la política $\mathbf{p}^7 = \{1, 2, 3\}$ significa que el taxista debe circular por la calle en busca de clientes si se encuentra en la zona A, estacionarse en la parada si se encuentra en la zona B y atender al radio taxi si se encuentra en la zona C.

Tenemos entonces $3 \times 2 \times 3 = 18$ políticas posibles, por lo que no resulta operativo explorarlas todas. Se trata, entonces, de llevar a cabo una exploración en el espacio de las políticas. La operativa de la exploración se detalla en la página siguiente.

N	0	1	2	3	4	5	6	7
$v_1^{(1)}$	0	8	10,3	10,5	11	11	11	11
$v_1^{(2)}$	0	3	9,4	10,9	11,8	11,8	11,8	11,8
$v_1^{(3)}$	0	4	5,2	5,7	6,4	6,4	6,4	6,4
$v_2^{(1)}$	0	16	16,5	16,5	17	17	17	17
$v_2^{(2)}$	0	14	21,3	22,9	23,8	23,8	23,8	23,8
$v_3^{(1)}$	0	7	9,9	10,8	11,6	11,6	11,6	11,6
$v_3^{(2)}$	0	4	11,3	12,9	13,8	13,8	13,8	13,8
$v_3^{(3)}$	0	4	5,6	5,3	5,6	5,6	5,6	5,6
v_1^*	0	8	10,3	10,9	11,8	11,8	11,8	11,8
v_2^*	0	16	21,3	22,9	23,8	23,8	23,8	23,8
v_3^*	0	7	11,3	12,9	13,8	13,8	13,8	13,8
v_1^{*1}	0	1	0	0	0	0	0	0
v_2^{*1}	0	9	11	12	12	12	12	12
v_3^{*1}	0	0	1	2	2	2	2	2
g	0	7	10,3	10,9	11,8	11,8	11,8	11,8
política								
A		1	1	2	2	2	2	2
B		1	2	2	2	2	2	2
C		2	2	2	2	2	2	2

Nótese que las políticas óptimas (marcadas en rojo en la tabla en cada caso) se han obtenido como:

$$\begin{aligned}
 v_1^* &= \max \{ v_1^{(1)}, v_1^{(2)}, v_1^{(3)} \} \\
 v_2^* &= \max \{ v_2^{(1)}, v_2^{(2)} \} \\
 v_3^* &= \max \{ v_3^{(1)}, v_3^{(2)}, v_3^{(3)} \}
 \end{aligned}$$

La normalización se ha llevado a cabo según las expresiones:

$$\begin{aligned}
 g &= \min \{ v_1^*, v_2^*, v_3^* \} \\
 v_i^{*1} &= v_i^* - g
 \end{aligned}$$

Nótese como el sistema evoluciones de las políticas $\{1, 1, 1\}$ a la $\{1, 2, 2\}$ y finalmente converge a la $\{2, 2, 2\}$. Por tanto, podemos decir que el taxista deberá esperar siempre en la parada de taxis, obteniendo así una ganancia media de 11,8.

3.6.3 Iteración del espacio de las políticas

La iteración del espacio de las políticas nos da la política óptima para el sistema a largo plazo para cada uno de los posibles i estados de partida. Dicha iteración no es más que determinar \mathbf{g} y \mathbf{w} para cada una de las políticas. Formalmente, consiste en resolver, para cada política p :

$$\mathbf{w}^{(p)} + \mathbf{g}^{(p)} = \mathbf{q}^{(p)} + \mathbf{P}^{(p)}\mathbf{w}^{(p)}$$

La iteración en el espacio de las políticas tiene la ventaja de considerar comportamientos *anómalos* (semirregulares o periódicos) de las diferentes cadenas de Markov asociadas a una política, pero tiene el serio inconveniente de que supone resolver un sistema de n ecuaciones con n incógnitas para cada una de las posibles políticas, y frecuentemente resulta más laborioso que la iteración en el espacio de los estados. Por ejemplo, para el ejemplo 3.6.2.a debemos resolver 18 sistemas de tres ecuaciones. Es frecuente utilizar esta exploración como confirmación de los resultados obtenidos mediante la iteración en los estados.

El ejemplo siguiente, que representa una situación con solamente dos estados, permite mostrar las posibilidades de la exploración del espacio de las políticas.

Ejemplo 3.6.3.a

Un profesor de Métodos Cuantitativos de la Universitat Politècnica de Catalunya (U.P.C.) debe decidir, dos veces cada trimestre, si pone un examen fácil o difícil. Lo ideal sería poner un examen difícil, y que hubieran muchos aprobados: esto mostraría que los alumnos han estudiado y han aprendido muchas cosas de la asignatura. Por el contrario, si hay pocos aprobados, a pesar de que el examen ha sido fácil, significa que no han estudiado lo suficiente. Las otras dos situaciones (examen difícil y pocos aprobados, y muchos aprobados con un examen fácil) el profesor las considera de valor intermedio.

Ahora bien, los alumnos estudian más o menos según si en el examen anterior ha habido pocos o muchos aprobados: si han aprobado muchos estudian poco, y si ha habido muchos suspensos estudian mucho. Esto afecta a la probabilidad de aprobar: si después de un examen con muchos aprobados el profesor pone un examen fácil, sólo tres de cada cinco veces habrá muchos aprobados. Si después de una situación con pocos aprobados pone un examen difícil, habrá muchos o pocos aprobados con igual probabilidad. Por descontado, si pone un examen difícil después de uno con muchos aprobados habrá pocos aprobados con toda seguridad, mientras que si después de haber habido pocos aprobados pone un examen fácil, estima que la probabilidad de que haya muchos aprobados es del 90%.

- a) *¿Cuál es la política más adecuada para asegurar a largo plazo el máximo nivel medio de conocimientos de los alumnos?*
- b) *La Universidad ha fijado unas normas para valorar la calidad de la docencia de los profesores, en las que el parámetro principal es directamente proporcional a la frecuencia con que hay muchos aprobados. ¿Modificará esto la política del profesor? ¿Cuál será la valoración de los conocimientos de los alumnos que hará el profesor si aplica esta política?*

RESOLUCIÓN:

a) ¿Cuál es la política más adecuada para asegurar a largo plazo el máximo nivel medio de conocimientos de los alumnos?

Este ejemplo está menos estructurado que el del caso anterior, aunque resulta relativamente sencillo establecer los estados y las decisiones:

- El sistema (es decir, el conjunto de los alumnos) puede encontrarse en dos estados: con muchos aprobados (M) o con pocos (P). El hecho de que el comportamiento de los alumnos se vea condicionado exclusivamente por el *último* examen hace que nos encontremos con una (o varias, puesto que variarán en función de la política) cadenas de Markov de orden 1.
- Antes de la siguiente evolución, el profesor toma la decisión de que el examen sea fácil (F) o difícil (D). Dado que el profesor debe establecer cuál es su comportamiento en función de que haya habido muchos o pocos aprobados, podemos definir en este caso la política como:

{comportamiento si M, comportamiento si P}

Siendo así posibles cuatro políticas:

{F, F}
 {F, D}
 {D, F}
 {D, D}

Cada política tendrá sus propias matrices **P** y **R**. Dichas matrices pueden obtenerse de la lectura atenta del enunciado, y se muestran en la siguiente tabla: la matriz de la izquierda muestra las probabilidades, y la de la derecha las remuneraciones. También se han obtenido los valores de $q_i^{(p)}$ para cada caso.

ESTADOS	política	M	P	M	P	$q_i^{(p)}$
M	F	0,6	0,4	0	-2	-0,8
	D	0	1	2	0	0
P	F	0,9	0,1	0	-2	-0,2
	D	0,5	0,5	2	0	1

En este caso, el enunciado da cierta libertad para la determinación de los valores de $\mathbf{R}^{(p)}$, dado que se trata de funciones de utilidad. En este caso se ha optado por asignar a la situación más deseada (muchos aprobados con examen difícil) una utilidad de +2, a la menos deseada (pocos aprobados con examen fácil) una utilidad de -2, y al resto de casos (situados en un término medio en el enunciado) un valor intermedio de cero. Pueden darse formulaciones alternativas a la matriz $\mathbf{R}^{(p)}$, que darán lugar a diferentes valores de $q^{(p)}$.

La siguiente tabla muestra los resultados de la iteración en el espacio de los estados:

N	0	1	2	3	4	5	6
$v_M^{(F)}$	0	-0,8	-0,4	-0,6	-0,5	-0,55	-0,525
$v_M^{(D)}$	0	0	1	0,5	0,75	0,625	0,6875
$v_P^{(F)}$	0	-0,2	-0,1	-0,15	-0,125	-0,1375	-0,13125
$v_P^{(D)}$	0	1	1,5	1,25	1,375	1,3125	1,34375
v_M^*	0	0	1	0,5	0,75	0,625	0,6875
v_P^*	0	1	1,5	1,25	1,375	1,3125	1,34375
v_M^{*1}	0	0	0	0	0	0	0
v_P^{*1}	0	1	0,5	0,75	0,625	0,6875	0,65625
g	0	0	1	0,5	0,75	0,625	0,6875

Del examen de los resultados, obtenemos:

Parece que el sistema converge a la política $\{D, D\}$ desde un primer momento.

Sin embargo, no podemos asegurar que el sistema converja, dado que tanto g como $v_M^{*1} \approx w_P$ no convergen con las iteraciones dadas en la tabla².

Por este motivo, podemos realizar una exploración en el espacio de las políticas, evaluando el comportamiento a largo plazo del sistema si establecemos la política $\{D, D\}$. De la primera tabla podemos ver que:

$$\mathbf{P} = \begin{vmatrix} 0 & 1 \\ 0,5 & 0,5 \end{vmatrix}$$

$$\mathbf{R} = \begin{vmatrix} 2 & 0 \\ 2 & 0 \end{vmatrix}$$

$$\mathbf{q} = \begin{vmatrix} 0 \\ 1 \end{vmatrix}$$

Un examen de \mathbf{P} muestra rápidamente que la cadena de Markov asociada a esta política es ergódica, por lo que el vector \mathbf{g} tiene todas sus componentes iguales. Realizada esta precisión, podemos plantear el sistema:

$$\begin{aligned} g + w_M &= w_P \\ g + w_P &= 1 + 0,5w_M + 0,5w_P \end{aligned}$$

Para comparar con facilidad los resultados obtenidos con lo obtenido en la iteración en el espacio de los estados, haremos $w_M = 0$. Entonces se obtiene:

$$g = w_M = 2/3 = 0,6667$$

Resultado que confirma lo obtenido con la iteración en el espacio de las políticas. Ahora podemos decir que la política óptima es $\{D, D\}$ ³.

- b) *La Universidad ha fijado unas normas para valorar la calidad de la docencia de los profesores, en las que el parámetro principal es directamente proporcional a la frecuencia con que hay muchos aprobados. ¿Modificará esto la política del profesor? ¿Cuál será la valoración de los conocimientos de los alumnos que hará el profesor si aplica esta política?*

Ahora la matriz de remuneraciones cambia (las preferencias del profesor son diferentes a las de la Universidad), y tenemos la siguiente situación:

ESTADOS	política	M	P	M	P	q
M	F	0,6	0,4	2	0	1,2
	D	0	1	2	0	0
P	F	0,9	0,1	2	0	1,8
	D	0,5	0,5	2	0	1

² De hecho sí acaba convergiendo, pero bastante despacio: podemos afirmar que $w_M = g = 0,666$ después de 11 iteraciones.

³ Para poder asegurarlo completamente, deberíamos calcular \mathbf{g} y \mathbf{w} para las cuatro políticas. La de mayor g sería la política óptima.

Iterando en el espacio de los estados, tenemos:

N	0	1	2	3	4	5	6
$v_M^{(F)}$	0	1,2	1,44	1,368	1,3896	1,38312	1,385064
$v_M^{(D)}$	0	0	1,2	0	0	0	0
$v_P^{(F)}$	0	1,8	1,86	1,842	1,8474	1,84578	1,846266
$v_P^{(D)}$	0	1	1,3	1,21	1,237	1,2289	1,23133
v_M^*	0	1,2	1,44	1,368	1,3896	1,38312	1,385064
v_P^*	0	1,8	1,86	1,842	1,8474	1,84578	1,846266
v_M^{*1}	0	0	0	0	0	0	0
v_P^{*1}	0	0,6	0,42	0,474	0,4578	0,46266	0,461202
g	0	1,2	1,44	1,368	1,3896	1,38312	1,385064

Efectivamente, vemos que la política escogida ahora es {F, F}. Si juzgamos la política escogida por la Universidad a partir de los criterios del profesor, debemos evaluar esta política {F, F} a largo plazo con los datos del apartado a).

Tenemos ahora que:

$$\mathbf{P} = \begin{vmatrix} 0,6 & 0,4 \\ 0,9 & 0,1 \end{vmatrix}$$

$$\mathbf{R} = \begin{vmatrix} 0 & -2 \\ 0 & -2 \end{vmatrix}$$

$$\mathbf{q} = \begin{vmatrix} -0,8 \\ -0,2 \end{vmatrix}$$

Podemos plantear el sistema:

$$\begin{aligned} g + w_M &= -0,8 + 0,6w_M + 0,4w_P \\ g + w_P &= -0,2 + 0,9w_M + 0,1w_P \end{aligned}$$

Si hacemos $w_M = 0$, obtenemos los siguientes resultados:

$$\begin{aligned} g &= -0,615 \\ w_P &= 0,4615 \end{aligned}$$

De donde cabe deducir que esta política de la Universidad da como resultado una obtención de conocimientos por parte de los alumnos subóptima: la política {D, D} es claramente mejor que la {F, F}.

3.7 Problemas resueltos

3.7.1 Servicio en tenis

Es bien conocido que en el deporte de tenis, el jugador que sirve tiene dos oportunidades para que la pelota entre en el cuadrado de servicio. Si falla las dos veces, pierde el punto.

En cada servicio, el tenista puede optar por sacar fuerte o flojo. Si saca fuerte, tiene una probabilidad de que la pelota entre en el cuadro de 0,5. Si saca flojo, la probabilidad de entrar la pelota en el cuadro es de 0,8. Ahora bien, si entra el servicio fuerte, su probabilidad de ganar el punto es de 0,7, y si entra el servicio flojo su probabilidad de ganar el punto es de 0,4. La ganancia por punto ganado es de +10, y la ganancia por punto perdido de -10.

Se desea conocer la estrategia a seguir para maximizar la ganancia esperada por servicio en el largo plazo.

Se pregunta:

- Formule el problema como un proceso markoviano con remuneración y decisión, definiendo los estados posibles del sistema, la variable de decisión y las matrices **P** y **R** para todas las posibilidades.
- Mediante iteración en el espacio de los estados, identifique una posible política de servicio óptima.
- Determine la ganancia esperada por servicio a largo plazo, si el tenista opta siempre por el servicio fuerte.

Solución problema del tenis

La forma más adecuada de resolver el problema es definir sus parámetros del siguiente modo:

- Etapas:* cada servicio es una etapa en la evolución del sistema. Si se trata de un primer servicio, puede ser que el servicio dé lugar a un punto (si entra en el cuadro) o a un segundo servicio (si no entra). El segundo servicio dará lugar, en todos los casos, a un punto, que podrá ser ganado o perdido.
- Estados:* por lo que se acaba de exponer, la información que necesitamos saber para conocer la evolución del sistema consiste en saber si estamos en el primer servicio (estado 1) o en el segundo servicio (estado 2). Por lo tanto, se trata de un sistema dinámico con dos estados.
- Variable de decisión:* el tenista puede, en parte, controlar con sus acciones la evolución del sistema. En este caso, puede hacerlo sirviendo fuerte (FU) o sirviendo flojo (FL). El tenista puede sacar fuerte o flojo tanto en el primer servicio como en el segundo.

Por lo tanto, nos encontramos ante un sistema de *Markov con remuneración y decisión* (caso particular de los modelos de programación dinámica aleatoria homogénea en el tiempo), del que nos interesa conocer el comportamiento a largo plazo (esto es, el problema de *horizonte infinito*).

Las reglas del juego del tenis, junto con las probabilidades que expresan las capacidades del jugador, nos permitirán determinar las matrices de probabilidades de transición y de remuneración, para todos los valores de estados del sistema y de variables de decisión.

Primer servicio

Si estamos en el primer servicio y lo entramos en el cuadro, se disputará el punto y el sistema evolucionará al estado 1 (volveremos al primer servicio). Veamos qué probabilidades y qué remuneraciones obtenemos para cada una de las variables de decisión.

Si saca fuerte, la probabilidad de entrar el servicio es de 0,5. Ésta será, entonces, la probabilidad de transición. Dado que el tenista tiene una probabilidad de ganar el punto de 0,7 (y de perderlo de 0,3), la remuneración será:

$$0,7 \cdot (+10) + 0,3 \cdot (-10) = +4$$

Si, en cambio, saca flojo, la probabilidad de entrar el servicio es de 0,8. Ésta será, una vez más, la probabilidad de transición al estado 1. Sin embargo, las probabilidades de ganar el punto en estas condiciones no son tan halagüeñas como en el caso anterior, y tenemos:

$$0,4 \cdot (+10) + 0,6 \cdot (-10) = -2$$

Si no entra el primer servicio, no se disputa el punto y pasamos al segundo servicio: en términos del problema, evolucionamos al estado 2. Al disputarse el punto, la remuneración es cero. Las probabilidades de que esto suceda dependen de la variable de decisión:

Si saca fuerte, la probabilidad de fallar el servicio es de 0,5.

Si saca flojo, la probabilidad de fallar el servicio es de 0,2.

Segundo servicio

En el segundo servicio, el punto está en juego, y si el tenista falla el servicio pierde el punto. Pase lo que pase, el próximo saque será un primer servicio. Por lo tanto, en el estado 2 la probabilidad de transición al estado 1 es de 1, y en consecuencia la probabilidad de transición al estado 2 es de 0.

¿Cómo puede ganar el punto el tenista en el segundo servicio? Entrando el servicio y ganando el punto. Por el contrario, el tenista pierde el punto si falla el servicio, o si entra el servicio y pierde el punto. En consecuencia, las remuneraciones valdrán:

$$\begin{aligned} \text{Si saca fuerte:} & \quad 0,5 \cdot 0,7 \cdot (+10) + \{0,5 \cdot 0,3 + 0,5\} \cdot (-10) = -3 \\ \text{Si saca flojo:} & \quad 0,8 \cdot 0,4 \cdot (+10) + \{0,8 \cdot 0,6 + 0,2\} \cdot (-10) = -3,6 \end{aligned}$$

Matrices de transición y remuneración

Ahora ya podemos formular el problema como un sistema markoviano de remuneración y decisión. Las matrices de remuneración y decisión quedan como:

		Estado 1	Estado 2	Estado 1	Estado 2	
Estado 1	FU	0,5	0,5	+04	0	+2
	FL	0,8	0,2	-2	0	-1,6
Estado 2	FU	1	0	-3		-3
	FL	1	0	-3,6		-3,6
		Matriz P		Matriz R		Vector q

En la tabla adjunta se muestra la iteración para el espacio de los estados, que muestra que la política óptima a largo plazo es la {FU, FU}, esto es, sacar siempre fuerte.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
f(1,FU)	0.000	2.000	4.500	3.250	3.875	3.563	3.719	3.641	3.680	3.660	3.670	3.665	3.667	3.666
f(1,FL)	0.000	-1.600	2.400	0.400	1.400	0.900	1.150	1.025	1.088	1.056	1.072	1.064	1.068	1.066
f(0,FU)	0.000	-3.000	2.000	-0.500	0.750	0.125	0.438	0.281	0.359	0.320	0.340	0.330	0.335	0.333
f(0,FL)	0.000	-3.600	-3.600	-3.600	-3.600	-3.600	-3.600	-3.600	-3.600	-3.600	-3.600	-3.600	-3.600	-3.600
f*(1)	0.000	2.000	4.500	3.250	3.875	3.563	3.719	3.641	3.680	3.660	3.670	3.665	3.667	3.666
f*(0)	0.000	-3.000	2.000	-0.500	0.750	0.125	0.438	0.281	0.359	0.320	0.340	0.330	0.335	0.333
f(1)	0.000	5.000	2.500	3.750	3.125	3.438	3.281	3.359	3.320	3.340	3.330	3.335	3.333	3.334
f(0)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
g	0.000	-3.000	2.000	-0.500	0.750	0.125	0.438	0.281	0.359	0.320	0.340	0.330	0.335	0.333

Podemos determinar de manera exacta el comportamiento del sistema a largo plazo planteando para la política {FU, FU}:

$$\begin{aligned} g + w_1 &= 2 + 0,5 \cdot w_1 + 0,5 \cdot w_2 \\ g + w_2 &= -3 + w_1 \end{aligned}$$

Haciendo $w_1 = 0$, obtenemos: $g = 0,33$ y $w_2 = 3,33$.

En definitiva, adoptando la política {FU, FU} el tenista puede ganar su servicio a largo plazo, puesto que la ganancia por servicio a largo plazo es positiva.

En la resolución manual, basta con iterar dos o tres etapas y resolver el sistema para la política obtenida (como puede verse, la convergencia es bastante lenta en este caso).

3.8 Glosario de términos

Cadena de Markov:

Una cadena de Markov de orden k es un proceso estocástico cuyas probabilidades de transición están determinadas por los valores de los estados alcanzados en las k transiciones anteriores. Dichas probabilidades de transición se mantienen inalterables a lo largo del tiempo (de otro modo, son independientes de la transición en que nos encontremos). Cualquier cadena de Markov puede reducirse a una cadena de orden 1.

Cadena de Markov cíclica:

Cadena de Markov con una única clase final cíclica. Puede tener clases de paso (cadena semiérgódica cíclica) o no tenerlas (cadena érgódica cíclica).

Cadena de Markov érgódica:

Cadena de Markov en la que todas las filas de la matriz de probabilidades estacionarias de \mathbf{P}^* son iguales, de manera que *ninguno* de los elementos de \mathbf{P}^* es igual a cero. Se trata de una cadena con una sola clase final y *sin* clases de paso.

Cadena de Markov mixta:

Cadena de Markov no érgódica con clases finales cíclicas y acíclicas.

Cadena de Markov no ergódica:

Cadena de Markov cuyas probabilidades estacionarias dependen del estado inicial, por lo que las filas de \mathbf{P}^* pueden ser diferentes entre si. Se trata de una cadena con dos o más clases finales y al menos una clase de paso.

Cadena de Markov regular:

Cadena de Markov semiergódica cuya única clase final es acíclica.

Cadena de Markov regular positiva:

Cadena de Markov ergódica cuya única clase final es acíclica.

Cadena de Markov semiergódica:

Cadena de Markov en la que todas las filas de la matriz de probabilidades estacionarias de \mathbf{P}^* son iguales, de manera que *alguno* de los elementos de \mathbf{P}^* es igual a cero. Se trata de una cadena con una sola clase final y *al menos una* clases de paso.

Cadena de Markov semirregular:

Cadena de Markov no ergódica con todas sus clases finales *acíclicas*.

Cadena de Markov policíclica:

Cadena de Markov no ergódica con todas sus clases finales *cíclicas*.

Camino:

Decimos que existe un camino entre los estados i y j si i es descendiente de j .

Ciclo:

Decimos que existe un ciclo en una cadena de Markov si existe un camino que comunique a un estado i consigo mismo.

Clase final:

Conjunto de estados comunicados entre sí tal que si en algún momento el proceso llega a alguno de los estados de la clase final, en las transiciones siguientes el estado evoluciona entre los estados de la clase final.

Clase final acíclica (o aperiódica):

Clase final de una cadena de Markov de periodo superior a uno.

Clase final cíclica (o periódica):

Clase final de una cadena de Markov de periodo igual a uno.

Clase de paso:

Conjunto de estados comunicados entre sí tal que si en algún momento el proceso llega a alguno de los estados de la clase de paso, en las transiciones siguientes puede evolucionar a algún estado diferente.

Decisión:

Variable de un proceso estocástico que puede ser gobernada por el observador. Cada estado tiene definido un número de decisiones d_i . Cada una de ellas da lugar a diferentes probabilidades de transición y remuneraciones para ese estado.

Estado:

Variable descriptiva de la situación en que se encuentra el proceso en un momento del tiempo. Puede ser una variable cuantitativa o cualitativa. También puede ser continua o discreta, aunque en este módulo sólo se trata el último caso.

Estado absorbente:

Estado cuya única transición posible es volver al mismo estado. Un estado absorbente constituye una clase final de un único estado.

Estado descendiente:

Un estado i es descendiente de j si cuando iniciamos el proceso en i existe una probabilidad no nula de que el proceso llegue a j .

Estados comunicados:

Decimos que dos estados i y j se comunican cuando i es descendiente de j y j es descendiente de i . La comunicación entre estados es una relación de equivalencia, puesto que es reflexiva, simétrica y transitiva.

Iteración en el espacio de las políticas:

Obtención de la remuneración esperada a largo plazo para una política determinada.

Iteración en el espacio de los estados:

Obtención de la remuneración esperada cuando faltan N etapas para finalizar el proceso, operando transición a transición. Usando la normalización, podemos llegar a obtener la remuneración esperada por transición a largo plazo después de operar para un número de transiciones lo bastante elevado.

Longitud de un ciclo:

Número mínimo de transiciones necesarias para volver al estado i en un ciclo, si se inició el proceso en ese estado.

Matriz de probabilidades de transición de un paso (\mathbf{P}):

Matriz cuadrada de orden n , donde n es igual al número de estados del proceso. Sus componentes son las probabilidades de transición de un paso p_{ij} .

Matriz de probabilidades estacionarias (\mathbf{P}^):*

Matriz cuadrada de orden n , donde n es igual al número de estados del proceso. Sus componentes son las probabilidades estacionarias π_{ij} .

Matriz de remuneraciones (\mathbf{R}):

Matriz cuadrada de orden n , donde n es igual al número de estados del proceso. Sus componentes son las remuneraciones r_{ij} .

Módulo z :

El módulo z de un número n es el resto de la división de n por z . Los valores posibles del módulo z van de 0 a $z - 1$. Por ejemplo, el módulo 2 de 7 es 1 , y el módulo 3 de 11 es 2 .

Normalización:

Técnica utilizada en la iteración en el espacio de los estados, para obtener la remuneración esperada por transición en el largo plazo. Véase la sección 5.1.2 para una definición en su contexto.

Periodo de una clase final:

Máximo común divisor de las longitudes de ciclo que pueden encontrarse en una clase final.

Política:

Se obtiene una política para una transición escogiendo una variable de decisión para cada uno de los estados. Una vez fijada una política, un proceso markoviano de decisión se convierte en una cadena de Markov.

Probabilidad de transición de k pasos ($p_{ij}^{(k)}$):

Probabilidad de que el proceso evolucione al estado j dentro de k transiciones, si en este momento se encuentra en el estado i .

Probabilidad de transición de un paso (p_{ij}):

Probabilidad de que el proceso evolucione al estado j en la siguiente transición, si en este momento se encuentra en el estado i . En una cadena de Markov de orden uno, las probabilidades de transición de un paso no varían con el tiempo y no dependen de la evolución seguida en etapas anteriores.

Probabilidades estacionarias (π_{ij}):

Probabilidad de que el proceso se encuentre en el estado j después de un número elevado de transiciones (una vez alcanzado el régimen permanente), si el sistema comenzó su evolución en el estado i .

Proceso estocástico:

Proceso que evoluciona de manera no determinista a lo largo del tiempo, tomando cada vez valores de entre un conjunto de estados. Suele conocerse la ley de probabilidad que rige la transición de un estado a otro.

Remuneración (r_{ij}):

Valor numérico asociado a la transición de i a j con probabilidad de transición p_{ij} diferente de cero. Puede ser un valor positivo, negativo o nulo. Su significado dependerá de la situación modelizada.

Transición:

Momento en que una cadena de Markov puede cambiar de estado. Es usual que una transición se produzca a intervalos regulares de tiempo.

Valor propio y vector propio:

Para una matriz cuadrada \mathbf{A} , podremos encontrar un conjunto de escalares complejos y otro de vectores tal que $\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}$. Diremos entonces que \mathbf{v} es un vector propio y λ un valor propio de la matriz \mathbf{A} . Los valores propios pueden encontrarse resolviendo el sistema $\det(\mathbf{A} - \lambda \cdot \mathbf{I}) = 0$. Una vez hallados los valores λ , pueden hallarse los \mathbf{v} correspondientes resolviendo el sistema (compatible indeterminado) $\mathbf{A} - \lambda \cdot \mathbf{I} = \mathbf{0}$.