

UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH→ **UPCPOSTGRAU**

# Cadenas de Markov →

## MÉTODOS CUANTITATIVOS PARA LA TOMA DE DECISIONES III

Joan B. Fonollosa  
Albert Suñé  
Josep M. Sallán  
Vicenç Fernández





UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH



iniciativa  
digital politècnica  
Publicacions Acadèmiques UPC



UPCPOSTGRAU

# Cadenas de Markov → MÉTODOS CUANTITATIVOS PARA LA TOMA DE DECISIONES III

Joan B. Fonollosa  
Albert Suñé  
Josep M. Sallán  
Vicenç Fernández

Primera edición: Noviembre de 2016

- © Los autores, 2016
- © Iniciativa Digital Politècnica, 2016  
Oficina de Publicacions Acadèmiques Digitals de la UPC  
Jordi Girona 31,  
Edifici Torre Girona, Planta 1, 08034 Barcelona  
Tel.: 934 015 885  
[www.upc.edu/idp](http://www.upc.edu/idp)  
E-mail: [info.idp@upc.edu](mailto:info.idp@upc.edu)

Fotografía de la cubierta: © iStock.com/

Producción: SERVICE POINT  
Pau Casals, 161-163  
08820 El Prat de Llobregat (Barcelona)

Depósito legal: B 22943-2016  
ISBN: 978-84-9880-611-3

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra sólo puede realizarse con la autorización de sus titulares, salvo excepción prevista en la ley.



# Índice

|                                                           |    |
|-----------------------------------------------------------|----|
| <b>Índice</b>                                             | 5  |
| <b>Presentación de la obra</b>                            | 9  |
| <b>1. Introducción a los métodos cuantitativos</b>        | 13 |
| 1.1 Concepto de modelo                                    | 14 |
| 1.2 Tipos de problemas                                    | 15 |
| 1.2.1 Problemas de inventarios ( <i>stocks</i> )          | 15 |
| 1.2.2 Problemas de reparto                                | 16 |
| 1.2.3 Problemas de secuencias                             | 17 |
| 1.2.4 Problemas de colas                                  | 17 |
| 1.2.5 Problemas de renovación                             | 18 |
| 1.2.6 Problemas de caminos                                | 18 |
| 1.2.7 Problemas de competencia                            | 19 |
| 1.2.8 Problemas de búsqueda                               | 19 |
| 1.3 Métodos cuantitativos de tratamiento de los problemas | 20 |
| 1.3.1 Métodos exactos                                     | 20 |
| 1.3.2 Métodos no exactos                                  | 22 |
| 1.4 Relación entre problemas y técnicas                   | 23 |
| 1.5 Notación y conocimientos previos                      | 24 |
| <b>2. Procesos polietápicos</b>                           | 27 |
| 2.1 Definición de proceso polietápico                     | 27 |
| 2.2 Comparación con la programación lineal                | 30 |
| 2.3 Modelos polietápicos existentes                       | 31 |
| 2.4 Introducción al cálculo recurrente                    | 32 |
| <b>3. Conceptos básicos de las cadenas de Markov</b>      | 37 |
| 3.1 Procesos markovianos                                  | 37 |
| 3.2 Cadenas de Markov                                     | 37 |
| 3.3 Modelización mediante las cadenas de Markov           | 38 |
| 3.4 Cadenas de orden 1 y de orden superior                | 40 |



|                                                                               |     |
|-------------------------------------------------------------------------------|-----|
| <b>4. Cadenas de Markov simples</b>                                           | 43  |
| 4.1 Matriz de una cadena de Markov                                            | 43  |
| 4.1.1 Matrices de probabilidad de transición                                  | 43  |
| 4.1.2 Matrices estocásticas                                                   | 43  |
| 4.1.3 Cadenas de orden superior                                               | 44  |
| 4.2 Probabilidades estacionarias y tipos de cadenas de Markov                 | 47  |
| 4.2.1 Probabilidades de transición de k pasos. Teorema de Chapman-Kolmogorov  | 47  |
| 4.2.2 Probabilidades estacionarias                                            | 48  |
| 4.2.3 Cadenas ergódicas                                                       | 49  |
| 4.2.4 Cadenas semiergódicas                                                   | 50  |
| 4.2.5 Cadenas no ergódicas                                                    | 50  |
| 4.2.6 Cadenas cíclicas                                                        | 51  |
| 4.2.7 Clasificación de las cadenas de Markov                                  | 52  |
| 4.3 Análisis topológico de las cadenas de Markov                              | 52  |
| 4.3.1 Propiedades de estado                                                   | 53  |
| 4.3.2 Propiedades de clase                                                    | 54  |
| 4.3.3 Periodicidad de una clase final                                         | 55  |
| 4.3.4 Clasificación de las cadenas a partir del análisis topológico           | 56  |
| 4.4 Análisis espectral                                                        | 58  |
| 4.4.1 Propiedades del espectro de una matriz estocástica                      | 58  |
| 4.4.2 Clasificación según el análisis espectral                               | 58  |
| 4.4.3 Comparación entre los dos tipos de análisis                             | 60  |
| 4.5 Significado de la clasificación                                           | 60  |
| 4.5.1 Significado de la ergodicidad                                           | 60  |
| 4.5.2 Significado de la periodicidad                                          | 61  |
| 4.6 Cálculo de las probabilidades estacionarias                               | 62  |
| 4.6.1 Probabilidades estacionarias y comportamiento a largo plazo             | 62  |
| 4.6.2 Probabilidades estacionarias para las cadenas ergódicas y semiergódicas | 63  |
| 4.6.3 Probabilidades estacionarias para las cadenas no ergódicas              | 67  |
| 4.6.4 Modelización con cadenas de Markov simples                              | 76  |
| <b>5. Cadenas de Markov con remuneración</b>                                  | 79  |
| 5.1 Introducción                                                              | 79  |
| 5.2 Remuneración esperada a corto plazo                                       | 80  |
| 5.2.1 Iteración en el espacio de los estados                                  | 80  |
| 5.2.2 Normalización                                                           | 81  |
| 5.3 Remuneración esperada a largo plazo                                       | 84  |
| 5.3.1 Paso del corto al largo plazo                                           | 84  |
| 5.3.2 Remuneración esperada a largo plazo en las cadenas regulares            | 86  |
| 5.3.3 Remuneración esperada en las cadenas cíclicas                           | 89  |
| 5.3.4 Remuneración esperada en las cadenas no ergódicas                       | 91  |
| 5.4 Duración del régimen transitorio                                          | 94  |
| <b>6. Cadenas de Markov con remuneración y decisión</b>                       | 101 |
| 6.1 Introducción                                                              | 101 |

|                                                    |            |
|----------------------------------------------------|------------|
| 6.2 Procesos markovianos de decisión .....         | 102        |
| 6.3 Iteración en el espacio de los estados .....   | 104        |
| 6.4 Iteración en el espacio de las políticas ..... | 109        |
| <b>7. Problemas resueltos .....</b>                | <b>119</b> |
| 7.1 Revisión de los conceptos teóricos .....       | 119        |
| 7.2 Un poco más de revisión de conceptos .....     | 121        |
| 7.3 Saratres Airlines .....                        | 122        |
| 7.4 El servicio en el tenis .....                  | 125        |
| 7.5 Tres elementos delicados .....                 | 129        |
| 7.6 Embarcaciones de lujo .....                    | 131        |
| 7.7 Mantenimiento .....                            | 133        |
| 7.8 Cultivos rotativos .....                       | 136        |
| <b>Bibliografía .....</b>                          | <b>141</b> |
| <b>Glosario .....</b>                              | <b>143</b> |





## Presentación de la obra

Las cadenas de Markov (también conocidas como *procesos markovianos*) deben su nombre al matemático ruso Andréi Andréyevich Markov (1856-1922), que definió teóricamente el concepto en 1906. En 1913, lo aplicó al estudio de un problema muy concreto, que permite comprender bien la esencia del modelo: el estudio de las primeras 20.000 letras (en este caso, del alfabeto cirílico) de la novela *Eugenio Onegin* de Pushkin.

En efecto, un texto cualquiera se puede considerar una secuencia aleatoria de símbolos. Pero, si estudiamos la frecuencia con que cada símbolo aparece en el texto completo (denominada *probabilidad a priori* o *estática*) y generamos una secuencia aleatoria con cada una de estas frecuencias no solo obtendremos un texto absurdo, sino que ni siquiera “parecerá” que pueda pertenecer al idioma original. Y es que la secuencia de las letras del alfabeto en un texto se rige por unas reglas mucho más complejas que la simple aparición aleatoria según su frecuencia a priori. Así, por ejemplo, aunque la letra *e* sea la más frecuente en un texto en castellano, difícilmente aparecerá después de otra vocal, salvo que la vocal anterior sea una *i* o una *u*, con las cuales forma un diptongo en palabras como *tiene*, *fiera*, *puede* o *hueso*. En cambio, la probabilidad de que aparezca una *u* después de una *q* es prácticamente uno, pese a que se trata de la vocal menos frecuente.

Ello nos lleva de manera natural a entender qué clase de sistemas se modelizan mediante la técnica de las cadenas de Markov: sistemas aleatorios cuya evolución a lo largo del tiempo depende del estado o los estados anteriores por los cuales haya pasado. Siguiendo con el ejemplo del castellano, a diferencia de lo que ocurre con el alemán, por ejemplo, nunca podremos hallar más de cuatro consonantes consecutivas, y aun este caso solo se da en unos grupos de consonantes muy concretas: *-bstr-* (*abstracción*) *-nstr-* (*instrucción*, *construcción*). Ello nos dice que, para modelizar este sistema, hemos de considerar por lo menos sus cuatro últimos estados, es decir, lo que



se conoce como una *cadena de orden 4*. Sin embargo, los ejemplos anteriores nos dan otra pauta: si indicamos con \_ el final de palabra e ignoramos los acentos por simplicidad, la secuencia *cción\_* es mucho más frecuente de lo que debería si solo atendemos a su frecuencia a priori; pero esta secuencia tiene 6 caracteres, de modo que el orden de la cadena en castellano se eleva ya a 6, por lo menos.

El concepto se ha generalizado mucho a partir de esta consideración inicial y se puede aplicar a un gran número de situaciones, con la posibilidad de influir en el comportamiento del sistema, es decir, en la capacidad de decisión, a fin de obtener un rendimiento óptimo.

El modelo de las *cadenas de Markov* forma parte de un conjunto muy general de modelos, denominados *modelos polietápicos*, que comparten el hecho de que toman en consideración la evolución del sistema objeto de estudio a lo largo del tiempo. Dentro de la presente colección, se incluyen otros dos volúmenes dedicados a sendos modelos de esta clase:

- La teoría de colas
- La programación dinámica

Los tres ítems comparten determinadas características y, al mismo tiempo, tienen algunas diferencias notables. Este volumen incluye un capítulo dedicado precisamente a señalar dichos paralelismos y diferencias.



→ 1



# Introducción a los métodos cuantitativos

Este volumen forma parte de una serie en que se introducen las técnicas y las aplicaciones de los métodos cuantitativos para la toma de decisiones en el contexto de la organización industrial.

Dichos volúmenes se corresponden con las diversas asignaturas incluidas en el plan de estudios de las titulaciones impartidas en la Escuela Superior de Ingenierías Industrial, Aeroespacial y Audiovisual de Terrassa (ESEIAAT) sobre métodos cuantitativos para la toma de decisiones.

En este capítulo, se introducen la naturaleza y las especificidades de los métodos cuantitativos en la organización industrial. Esta introducción sigue una lógica de modelos, problemas y técnicas:

- a) En muchas ocasiones, en el contexto organizativo nos encontramos (generalmente, en la gestión del sistema productivo, pero también en otros contextos) con situaciones de una cierta complejidad, que plantean problemas que pueden resolverse elaborando un modelo cuantitativo. Conviene, pues, tener una noción clara de qué es un *modelo* y, en concreto, un *modelo matemático o cuantitativo*.
- b) También conviene que, al enfrentarnos a una situación determinada, no elaboremos un modelo cuantitativo cualquiera. Diferentes modelos pueden presentar distintos grados de dificultad (y, en consecuencia, un coste de resolución muy variable). Por este motivo, es importante saber que existen diferentes *problemas* tipo, que representan situaciones que pueden darse en contextos muy diferentes. En ocasiones, puede ser útil tener en mente un cierto abanico de problemas tipo a la hora de elaborar un modelo.



- c) Finalmente, necesitamos un conjunto de *técnicas* que nos permitan resolver el modelo, es decir, obtener, a un coste asequible y con una precisión aceptable, la información que buscábamos al elaborar dicho modelo. Con frecuencia, estas técnicas están relacionadas con problemas tipo, de ahí su importancia.

En definitiva, en los cursos de métodos cuantitativos para la toma de decisiones, se pretende que los estudiantes sean capaces de formular modelos matemáticos a partir de situaciones poco estructuradas. Dichas situaciones pueden tener alguna relación con problemas tipo a los cuales pueda aplicarse un conjunto bien definido de técnicas para hallar la solución del modelo.

## 1.1 Concepto de modelo

Una definición clásica de modelo es:

Un objeto  $M$  es un modelo de una realidad  $R$  para un observador  $O$ , si  $O$  puede obtener, estudiando  $M$ , las respuestas a las preguntas que se hace sobre  $R$ .

Así, una misma realidad puede ser representada por diferentes modelos, en función de las preguntas que nos hagamos sobre ella. Una red eléctrica, por ejemplo, puede ser representada por:

- Un grafo cuyos arcos tienen asociada la distancia entre diferentes puntos, si queremos determinar cuál es el conjunto de conexiones que minimiza la longitud del cable que hay que emplear.
- Un mapa del terreno a escala, para llevar a cabo la obra civil del tendido eléctrico. El mapa tiene más información que el grafo: además de estar a escala, incluye los accidentes orográficos (ríos, etc.), las curvas de nivel, etc.
- Una red de Kirchhoff, a fin de determinar las intensidades de la línea, las pérdidas de energía por transporte, etc.

En definitiva, el *modelo* viene determinado por el *problema* que hay que resolver y por los requerimientos de la *técnica* escogida para resolverlo.

El proceso de construcción y de resolución de un modelo puede dividirse en tres partes:

- a) Modelización: construcción y elaboración del modelo. Generalmente, se trata de un proceso que es difícil de sistematizar, puesto que podemos tener situaciones muy diversas que admitan modelos muy similares (es el caso, por ejemplo, de la programación lineal).



- b) Resolución. Diremos que hemos resuelto el modelo cuando hayamos podido responder a las preguntas que nos movieron a elaborarlo. Es decir, cuando hayamos obtenido la información que necesitábamos.
- c) Explotación. Una vez obtenidos los resultados, estos han de ser interpretados y se deben analizar las implicaciones para la gestión del sistema afectado. Otra cuestión importante es el mantenimiento del modelo: ver cómo evoluciona la solución cuando los parámetros del sistema evolucionan.

## 1.2 Tipos de problemas

Los métodos cuantitativos para la toma de decisiones centran su interés en un conjunto de problemas concretos, asociados generalmente a la organización de las interrelaciones entre los elementos del sistema que permiten optimizar su comportamiento. En muchas ocasiones, dicha optimización supone la minimización de los recursos empleados o del valor (coste) de esos recursos. En otras, buscaremos maximizar la utilidad asociada al uso del sistema. Dicha utilidad suele ir asociada, en el contexto empresarial, a la maximización del beneficio que se deriva de la actividad. A continuación, mostramos algunos problemas prototipo. Para cada uno de ellos, definimos tres aspectos:

- a) Las *características* del problema.
- b) El *objetivo* del problema, esto es, la magnitud a optimizar.
- c) Las posibles *variables de decisión*, es decir, las características del sistema que podemos controlar. En general, el valor de la magnitud a optimizar es función de dichas variables de decisión, además de unos *parámetros* sobre los cuales no tenemos control.

### 1.2.1 Problemas de inventarios (stocks)

Los problemas de inventarios aparecen cuando, por determinadas razones, los flujos de entrada de un recurso determinado en un sistema son diferentes de los flujos de salida. Este hecho exige que dispongamos de una cantidad determinada de dicho recurso, sin más finalidad que ajustar ambos flujos. Esta cantidad son las existencias o *stocks* del sistema.

Un sistema de este tipo tiene un conjunto de costes asociados a la gestión del sistema:

- *Coste de adquisición*. Es posible que el recurso tenga diferentes costes a lo largo del tiempo y que sea interesante adquirir una gran cantidad de recursos en un momento determinado, para su utilización cuando el coste sea elevado. Esta adquisición aumentará el valor de las existencias. En ocasiones, existe un componente de coste



independiente de la cantidad adquirida, que suele denominarse *coste de lanzamiento de un pedido*.

- *Coste de mantenimiento o posesión*. Mantener el recurso en *stock* supone unos costes de almacenamiento, financiación, etc., que tienden a limitar el valor del inventario.
- *Coste de rotura*. Puede darse el caso de que, con nuestra programación, no podamos servir el flujo de demanda y provoquemos una falta de suministro. Entonces, tendremos unos costes de rotura de *stocks*. Si estos costes son elevados, podrán hacer aumentar el valor del inventario.

El objetivo del problema será minimizar la suma total de costes de adquisición, mantenimiento y rotura asociados a un determinado flujo de salida. Las variables de decisión serán las cantidades que habrá que adquirir en cada instante de tiempo, así como el nivel de inventario entre cada entrada y cada salida.

La gestión de materiales en una empresa industrial es el ejemplo más evidente de un problema de *stocks*, aunque pueden encontrarse formulaciones similares en otros ámbitos:

- *Gestión de tesorería*, en que el inventario es el saldo de la cuenta bancaria en cada momento.
- *Gestión de un recurso*, por ejemplo, un embalse. En este caso, el inventario es el nivel de agua entre la entrada y la salida del embalse.

### 1.2.2 Problemas de reparto

Las características generales de un problema de reparto son que disponemos de un conjunto de recursos, con los cuales realizamos un conjunto de actividades. Los parámetros del sistema son:

- a) La cantidad que se dispone de cada recurso.
- b) La cantidad de cada recurso necesaria para producir una unidad de cada una de las actividades.
- c) El rendimiento o la utilidad que se deriva de producir una unidad de cada una de las actividades.

El objetivo es determinar cómo se asignan los recursos a las actividades, de manera que la utilidad o el rendimiento total sean máximos.

Una variante de este problema consiste en plantear una situación en que, además de la cantidad máxima de recursos, se deba realizar una cantidad mínima de cada una de las actividades. En este caso, podemos plantearnos



determinar si es posible servir esta cantidad mínima o no, averiguar qué actividades deben realizarse primero, etc.

Las variables de decisión serán las cantidades de actividades que habrá que realizar, o bien los recursos que se consumirán en cada actividad. También podemos plantear como variable de decisión qué actividades pueden realizarse y cuáles no.

Algunos problemas asimilables a un problema de reparto son:

- Determinar desde qué almacén debe servirse cada punto de consumo y en qué cantidad, para minimizar los costes de transporte.
- Los problemas de horarios: por ejemplo, qué profesores deben asignarse a cada asignatura.

### 1.2.3 Problemas de secuencias

En un problema de secuencias, tenemos un conjunto de tareas para realizar, relacionadas entre ellas por algunas restricciones. Por ejemplo, una tarea determinada no puede realizarse hasta que se hayan finalizado otras dos tareas previas. En ocasiones, dichas tareas consumen algunos recursos, de los cuales se dispone en una cantidad limitada.

El objetivo es determinar el momento en que debe iniciarse y acabarse cada tarea, así como el tiempo total de realización de dichas tareas. También puede que sea deseable conocer qué actividades pueden retrasarse sin retrasar la fecha de entrega del conjunto y cuáles no, así como la asignación de los recursos limitados a las tareas.

Los problemas de secuencias suelen ir asociados a actividades no repetitivas, como el lanzamiento de un producto al mercado o la construcción de un edificio.

### 1.2.4 Problemas de colas

Un problema de colas puede formularse del siguiente modo: determinadas unidades llegan a un sistema de forma aleatoria, según una *ley de llegada* conocida. Dichas unidades reciben un determinado servicio en uno o varios servidores. El tiempo de servicio sigue también una determinada *ley de servicio*. Además, el sistema puede tener varios componentes de coste: por ejemplo, el *coste de servicio* (coste asociado al hecho de disponer de un servidor) y el *coste de espera* (coste asociado a esperar a ser servido). Los problemas de colas están condicionados por el carácter aleatorio de las llegadas y del tiempo de servicio.



Lo que se pretende obtener son los parámetros asociados al sistema:

- a) El tiempo medio de espera
- b) El tiempo medio en el sistema (el tiempo de espera, más el tiempo de servicio)
- c) El número medio de unidades en espera
- d) El número medio de unidades en el sistema

También pueden plantearse cuestiones relativas al diseño del sistema, como el número de servidores que minimiza la suma de costes de servicio más los costes de espera.

Podemos encontrar problemas de colas en cualquier situación en que se produzcan esperas (las colas en el supermercado o en el aeropuerto, el proceso de matriculación en una universidad, etc.). De forma menos evidente, se pueden encontrar problemas de colas en el diseño de sistemas de mantenimiento, en la asignación de máquinas a operarios, etc.

### **1.2.5 Problemas de renovación**

En un problema de renovación, existe un conjunto de elementos que envejece. Ello puede incrementar los costes de funcionamiento o de reparación, y aumentar las probabilidades de avería. Por otra parte, sustituir el elemento supone un coste, que puede ser superior al de seguir utilizando el elemento en un momento determinado.

El objetivo no es otro que conocer cuándo debe reemplazarse un equipo. En una situación de largo plazo, puede interesar conocer la política de decisión en cada momento.

Los problemas de renovación se plantean en cualquier sistema cuyos elementos envejezcan y en que el problema de la fiabilidad del conjunto del sistema sea relevante.

### **1.2.6 Problemas de caminos**

Los problemas de caminos son propios de situaciones en que se puede ir de un estado inicial a otro final de varias maneras. Se trata de encontrar cuál de estas vías es mejor para minimizar el tiempo, los costes o los recursos.

El objetivo es encontrar la política más indicada para hallar el camino más corto entre el inicio y el final del recorrido. Las variables de decisión deben considerar las acciones que hay que realizar, y han de formar parte del camino más corto.



Los problemas de caminos pueden darse cuando hay que encontrar el camino más corto entre dos puntos, o la forma más económica de realizar una actividad con un inicio y un final definidos. Algunos problemas de secuencias pueden formularse en términos de problema de caminos.

### 1.2.7 Problemas de competencia

Los problemas de competencia estudian situaciones en que diversos actores toman decisiones, de manera que la decisión de un actor afecta los resultados obtenidos por los demás, además de él. Partiendo de una hipótesis determinada, relativa al comportamiento de los actores (usualmente, que se comporten de manera racional), se modeliza el sistema para determinar los resultados que obtendrá cada actor.

El propósito de estos modelos es determinar la *estrategia* óptima para dichos actores, definida como las características de las acciones que cada uno de ellos ha de llevar a cabo para optimizar su utilidad. Estas estrategias son las variables de decisión del modelo.

Cualquier situación en que se dé la interdependencia descrita entre los actores es susceptible de ser analizada como un problema de competencia. Entre los ejemplos más conocidos, podemos citar las decisiones de las empresas de un sector, una subasta o una situación con votaciones sucesivas.

### 1.2.8 Problemas de búsqueda

Los problemas de búsqueda consisten en situaciones en que, para acceder a una información determinada, se ha de incurrir en un cierto coste de búsqueda. Pueden darse, incluso, situaciones en que no haya la seguridad de que la información exista o esté disponible.

El objetivo de los problemas de búsqueda es determinar una forma de buscar dicha información minimizando los costes de búsqueda y, si hay que detener la búsqueda, considerando los costes en que se haya incurrido y los beneficios obtenidos con el hallazgo. Una variante de este problema es determinar cómo clasificar la información para facilitar su recuperación en una búsqueda posterior, incurriendo en los mínimos costes de búsqueda.

La variable de decisión es la política de búsqueda, que ha de incluir las reglas para detener la búsqueda. En la segunda versión del problema, las variables definirán la política de clasificación relacionada con la política de búsqueda.

Son ejemplos de situaciones en que se presentan problemas de búsqueda el control de calidad (la búsqueda de defectos), la búsqueda de yacimientos de petróleo, la prospección arqueológica o la búsqueda de información en internet, así como la determinación de las reglas de clasificación y archivo.



### 1.3 Métodos cuantitativos de tratamiento de los problemas

Los métodos cuantitativos disponibles para resolver los problemas descritos en el apartado anterior pueden clasificarse en dos grandes categorías:

- a) *Métodos exactos*. Se caracterizan por el hecho de que aseguran hallar la solución óptima, si existe. En ocasiones, encontrar esta solución óptima puede comportar unos costes elevados, en términos de tiempo de cálculo.
- b) *Métodos no exactos*. Tienen en común que no aseguran la solución óptima, sino *una solución razonablemente buena, en un tiempo razonable*. Existe una cierta probabilidad de encontrar efectivamente el óptimo, pero no podremos asegurar si la solución obtenida es óptima; solo sabremos que es una buena solución.

#### 1.3.1 Métodos exactos

##### Optimización clásica (análisis matemático)

Engloba diversos procedimientos analíticos que tienen en común la idea de encontrar la función derivada e igualarla a cero para obtener los puntos extremos. Suele requerirse que la función que se quiere optimizar sea continua y derivable. Algunos desarrollos matemáticos, como los multiplicadores de Lagrange, permiten resolver de esta manera problemas en que las variables de decisión están sometidas a determinadas restricciones.

##### Programación matemática

Los programas matemáticos son formulaciones del tipo:

$$[\text{OPT}] z = f(x),$$

$$\text{donde } x \in E \subset \mathbb{R}^n$$

en que  $f(x)$  es la *función objetivo* y  $E$  es la *región factible*, subconjunto de  $\mathbb{R}^n$  de soluciones posibles del problema.

El programa matemático será más o menos difícil de resolver según las características de la función objetivo y de la región factible. En este sentido, existen algunos casos especialmente relevantes:

- a) *Programación lineal*. La función  $f(x)$  es lineal y  $E$  viene determinada por un conjunto de restricciones lineales.
- b) *Programación lineal entera, binaria o mixta*. Se trata de programación lineal en que los componentes de  $x$  son variables enteras



(programación lineal entera), binarias (programación lineal binaria), o bien algunas variables son enteras o binarias y otras reales (programación lineal mixta).

- c) *Programación no lineal*. Programa matemático que no cumple las condiciones de la programación lineal, tal como se acaba de definir. Los modelos de programación lineal pueden ser muy difíciles de resolver, aunque los dos casos que se indican a continuación son especialmente interesantes para la organización industrial:
- Programación cuadrática:  $f(x)$  es un polinomio de segundo grado y las restricciones son lineales o de segundo grado.
  - Programación semilineal:  $f(x)$  es un polinomio de grado  $n$  y las restricciones son lineales.

La programación lineal es una de las técnicas más utilizadas para resolver los problemas propios de los métodos cuantitativos, en el contexto de la organización industrial, y se desarrolla en profundidad en este volumen de la colección.

### Teoría de grafos

Esta teoría trata de unos objetos matemáticos denominados *grafos*, que consisten en un conjunto de elementos (vértices del grafo) y las relaciones entre ellos (que pueden ser aristas o arcos del grafo).

### Programación dinámica

La programación dinámica es una metodología para encontrar las políticas óptimas en los procesos polietápicos de decisión. Se trata de sistemas que evolucionan en varias etapas, en cada una de las cuales hay que tomar una decisión, que hará que el sistema evolucione hacia un estado determinado, de manera determinista o siguiendo una distribución de probabilidad.

El objetivo es determinar qué política debemos seguir (decisión a tomar si nos encontramos en un estado determinado, es decir, en una etapa determinada) para optimizar la función objetivo. Dicha función se caracteriza por ser recursiva (eso es, depende de la decisión actual y del estado al cual el sistema puede evolucionar en la etapa siguiente).

### Cadenas de Markov

Una cadena de Markov es un sistema que evoluciona a lo largo del tiempo dentro de un conjunto de estados. Se caracteriza por que la probabilidad de que el sistema evolucione a un estado determinado depende exclusivamente del estado en que se encuentra en el momento presente.



### Teoría de colas

Se trata de modelos desarrollados específicamente para analizar problemas de colas. Consiste, fundamentalmente, en un conjunto de modelos descriptivos de diversas situaciones, relativas a las leyes de llegada y servicio, y otras características propias de estos sistemas.

Son de especial interés, por su (relativa) sencillez conceptual, los modelos de cola en que las tasas de llegada y de servicio siguen una ley de Poisson.

### Teoría de la decisión

Se trata de procedimientos que permiten valorar las diferentes alternativas y los posibles modelos de experimentos que hay que realizar para la toma de decisiones en situaciones que pueden describirse mediante un modelo matemático.

### Teoría de juegos

Se trata de modelos que estudian, de forma específica, los problemas de competencia. Se diferencia de la teoría de la decisión por la existencia de varios actores y por la interdependencia entre las acciones de cada uno de ellos.

### Procedimiento de separación y acotación

Es un procedimiento para obtener la solución de problemas en que el conjunto de soluciones es finito. La táctica adoptada consiste, en primer lugar, en dividir el conjunto en dos partes, siguiendo una *regla de separación*. A continuación, se acota el valor de la función objetivo en cada una de las partes, mediante una *regla de acotamiento*. Así, se explora un conjunto de soluciones cada vez menor hasta encontrar el óptimo.

## 1.3.2 Métodos no exactos

### Métodos heurísticos

Un método heurístico genera soluciones a un problema determinado mediante un método del cual se sabe, o bien por experiencia empírica o bien por razonamiento teórico, que genera *buenas soluciones* con una alta probabilidad. Suele clasificarse en tres grupos:

- a) *Algoritmos de un solo paso*. Son métodos que generan una única solución en cada etapa, tomando decisiones sucesivas que ya no son reconsideradas. Cada vez hay menos alternativas y estas están más condicionadas, de forma que las últimas decisiones pueden ser muy malas.



- b) *Métodos iterativos*. Son métodos que generan una solución en cada etapa, que permiten reconsiderar las soluciones obtenidas en etapas anteriores.
- c) *Métodos de mejora*. Son métodos que, partiendo de una solución determinada, van mejorándola en las etapas sucesivas.

## Simulación

Proceso en que se representa el estado del sistema mediante variables relacionadas por unas reglas determinadas. Una vez establecidas, se observa su evolución a lo largo del tiempo según unas hipótesis y unas reglas de gestión predeterminadas. Permite conocer la solución de problemas difícilmente resolubles mediante métodos exactos.

## Modelos descriptivos

Modelos que describen o reproducen, de una manera más o menos simplificada, la realidad que hay que estudiar, con lo cual permiten experimentar y estudiar sus reacciones ante determinadas decisiones, incidencias o políticas.

## Algoritmos genéticos

Se trata de procedimientos basados en la selección natural, en que una generación de individuos (conjunto de soluciones) da lugar, mediante el entrecruzamiento entre ellos, a una nueva generación de soluciones que, por término medio, es mejor, desde el punto de vista de la calidad del resultado obtenido.

## Redes neuronales

Sistemas basados en la simulación de las conexiones neuronales de los seres vivos, capaces de aprender de la experiencia y obtener soluciones cada vez mejores a problemas similares.

Existen varias relaciones entre las diversas técnicas. Por ejemplo, la programación lineal entera utiliza procedimientos de separación y acotación; determinados problemas de programación dinámica utilizan propiedades de las cadenas de Markov, y la teoría de la decisión utiliza técnicas de programación lineal, de la teoría de grafos y de simulación, etc.

## 1.4 Relación entre problemas y técnicas

No existe una relación directa entre los tipos de problemas y las técnicas utilizadas para resolverlos. Además, la mayoría de los problemas no pueden clasificarse de manera unívoca y las técnicas se pueden combinar. La tabla 1



es una aproximación para establecer relaciones entre los problemas (columnas) y las técnicas (filas). Se ha utilizado la notación siguiente:

P: Técnica principal para resolver el problema

S: Técnicas empleadas de forma secundaria

A: Técnicas auxiliares o utilizadas en casos especiales

Tabla 1.  
Problemas y técnicas  
de métodos  
cuantitativos

|                        | Inventarios | Reparto | Colas | Secuencias | Renovación | Caminos | Competencia | Búsqueda |
|------------------------|-------------|---------|-------|------------|------------|---------|-------------|----------|
| <b>Opt. clásica</b>    | P           |         | A     |            | P          |         | S           | S        |
| <b>Prog. lineal</b>    | S           | P       |       | A          | A          | S       | P           | A        |
| <b>Prog. no lin.</b>   | S           | S       |       |            |            |         |             |          |
| <b>P. dinámica</b>     | P           | S       |       | S          | S          | S       |             | P        |
| <b>T. grafos</b>       |             |         | A     | P          | S          | P       | A           | A        |
| <b>T. colas</b>        |             |         | P     | S          |            |         |             |          |
| <b>P. sep. y acot.</b> |             | A       |       | P          | P          | P       |             | A        |
| <b>T. decisión</b>     | S           |         | A     |            | S          |         | P           | P        |
| <b>T. juegos</b>       |             |         |       |            |            |         | P           |          |
| <b>Cad. Markov</b>     | S           | A       | S     |            |            |         |             | S        |
| <b>M. heurísticos</b>  | S           | S       |       | S          | S          | S       | A           | A        |
| <b>Simulación</b>      | A           | S       | P     | S          | P          |         | S           | S        |
| <b>M. descript.</b>    | A           |         |       | S          |            |         | P           | P        |
| <b>Alg. genéticos</b>  |             | A       |       | S          |            | P       |             | P        |
| <b>R. neuronales</b>   |             |         |       | S          | S          | P       | P           | S        |

## 1.5 Notación y conocimientos previos

En las expresiones matemáticas, se ha seguido la notación siguiente:

- Las variables escalares se representan en cursiva, en mayúscula o en minúscula, según convenga. Especialmente en la teoría de colas, se han utilizado letras griegas para representar los escalares. Por ejemplo:  $n$ ,  $k$ ,  $p$ ,  $N$ ,  $\rho$ ,  $\lambda$ ,  $\mu$ .
- Los vectores se representan en minúscula negrita. Si no se indica lo contrario, se trata de vectores columna. Por ejemplo: **v**, **g**.
- Las matrices se representan en mayúscula negrita. Por ejemplo: **P**, **A**.
- En cuanto a los conocimientos previos necesarios, se recomienda revisar las siguientes materias:



- Distribuciones de probabilidad: especialmente las propiedades de las distribuciones exponencial y de Poisson.
- Cálculo matricial: el producto de matrices y el significado de los valores y los vectores propios.
- Obtención de raíces complejas de un número real.

→2



# Procesos polietápicos

## 2.1 Definición de proceso polietápico

Los denominados *procesos polietápicos* son una clase muy general de modelos matemáticos que presentan las características siguientes (v. figura 2.1):

1. El sistema representado por el modelo se halla siempre en uno y solo uno de un conjunto de estados:  $X$
2. El sistema cambia de estado en unos instantes del tiempo, denominados *transiciones* o *etapas*.
3. En cada transición o etapa, puede haber la posibilidad de influir en el comportamiento del sistema mediante una decisión ( $u_t$ ).
4. Igualmente, en cada transición o etapa, puede haber la posibilidad de obtener una remuneración o un rendimiento ( $r_t$ ). Este nombre no debe prejuzgar acerca de su naturaleza, ya que tanto puede significar beneficios como costes.

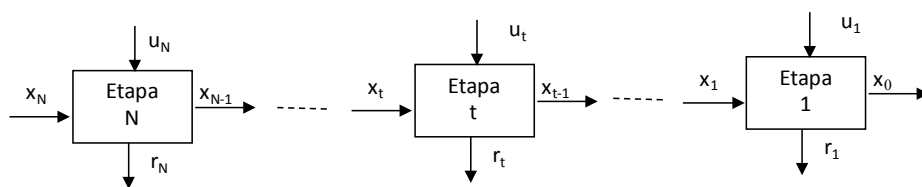


Figura 2.1.  
Procesos polietápicos



Algunas observaciones sobre la definición:

- La variable que cuenta el número de etapas ( $t$ ) recibe el nombre de *variable cronológica* porque usualmente tiene un significado temporal.
- Aunque no es estrictamente obligado, hemos optado por definir dicha variable como el número de etapas que faltan para acabar el proceso.
- El cambio de estado del sistema en una transición puede consistir en mantenerse en el mismo estado:

$$X_{t-1} = X_t$$

- La existencia de remuneración es opcional. Pueden existir sistemas sin remuneración.
- En los sistemas con remuneración, usualmente se define un rendimiento total que es función de los rendimientos de cada etapa:
- $R = R(r_1, r_2, \dots, r_N)$
- Esta función  $R$  es normalmente la suma, pero puede ser otra función cualquiera monótona creciente con todas las  $r_t$ .
- La existencia de decisión es también opcional. Dicha decisión implica que se desea optimizar el rendimiento total,  $R$ , que recibe entonces el nombre de *función objetivo* o *función económica*. El sentido de la optimización puede ser de máximo o de mínimo, según el significado del rendimiento.
- La decisión pretende optimizar el rendimiento total  $R$ , lo cual puede conllevar que, en determinados casos, la decisión óptima produzca un rendimiento muy pobre en la etapa donde se tome ( $r_t$ ), pero ello implicará dejar el sistema en condiciones mejores para los rendimientos futuros.
- Puesto que solo tiene sentido plantear una decisión si existe una función económica, puede haber sistemas con rendimiento y decisión, con rendimiento sin decisión o sin ninguno de ellos, pero no puede haber sistemas con decisión y sin rendimiento.

El número de etapas ( $N$ ) definido por la variable cronológica se denomina *horizonte* y puede ser:

- a) Continuo: la variable cronológica es continua y, por tanto, las transiciones pueden ocurrir en cualquier momento.
- b) Discreto: la variable cronológica es discreta, de modo que las transiciones ocurren en determinados instantes de tiempo (aunque no necesariamente han de estar distanciados uniformemente). En este caso, distinguiremos entre:



- b.1. Finito: el número de etapas (N) es conocido y finito.
- b.2. Indeterminado: el número de etapas (N) es finito pero desconocido.
- b.3. Infinito: el número de etapas (N) tiende a un infinito numerable.

El caso b3 es, evidentemente, una abstracción; sin embargo, dado que el comportamiento del sistema, en general, se estabiliza (en determinadas condiciones) al cabo de un cierto número de etapas, puede plantearse como un límite del horizonte finito. Del mismo modo, el caso b2 se tratará como el caso b3, ya que, si se desconoce el número de etapas que faltan, el caso debe tratarse como si el horizonte fuese infinito. De hecho, aunque nos hallemos en el caso finito, si N es lo bastante grande, se tratará como infinito. En realidad, el caso b1 solo tiene interés para estudiar el régimen transitorio o bien cuando el horizonte del problema es tan corto que no da tiempo a alcanzar el régimen permanente.

Tanto si el tiempo es continuo como si es discreto, se pueden distinguir dos situaciones diferentes y consecutivas:

- El *régimen transitorio* (también llamado *corto plazo*), durante el cual cada etapa es distinta en función de las condiciones iniciales; eso es, el sistema conserva la memoria de la situación de partida, la cual influye en la decisión.
- El *régimen permanente* (o *largo plazo*) en que el sistema ya ha alcanzado la estabilidad, aunque puede seguir conservado parte de la información inicial.

El comportamiento del sistema viene definido por dos distribuciones de probabilidad, que, de forma general, pueden definirse así:

- Probabilidad condicional de cambio de estado:

$$\Pr_t \{x_{t+1} \mid x_t, x_{t+1}, \dots, x_{t+k-1}; u_t\}$$

- Probabilidad del rendimiento:

$$\Pr_t \{r_t \mid x_t, x_{t+1}; u_t\}$$

La probabilidad condicional del cambio de estado indica que la probabilidad del nuevo estado depende de los k últimos estados. Si el tiempo es continuo, solamente será posible k = 1; en cambio, si el tiempo es discreto, puede darse cualquier valor de k. De la misma manera, el rendimiento obtenido en una etapa (si existe) dependerá de la decisión (si existe) y de los estados anterior y posterior.



Tal como se ha definido, el sistema es aleatorio, pero engloba como fenómeno particular el caso determinista, ya que basta con considerar que la distribución de probabilidad tiene un valor 1 para una alternativa, mientras que el resto son 0.

Las probabilidades se han indicado con el subíndice  $t$ , porque pueden ser diferentes para cada etapa, aunque ello no siempre es así:

- Decimos que el sistema es *homogéneo en el tiempo* (HT) si todas las probabilidades, tanto del nuevo estado como del rendimiento, son iguales para toda  $t$ ; entonces, prescindimos del subíndice y escribimos  $Pr\{\}$  en lugar de  $Pr_t\{\}$ .
- Por el contrario, si el sistema no cumple la condición anterior, decimos que es *no homogéneo en el tiempo* (NHT).

Dado que en los sistemas NHT debe conocerse el comportamiento de cada etapa, el problema con horizonte infinito o indeterminado solo tiene sentido en el caso HT.

Cuando existe decisión, una *política* es una regla de decisión que, para cada estado del sistema y para cada etapa, indica qué acción debe elegirse. Obsérvese que fijar una política equivale a convertir un sistema con decisión en uno sin ella. Dicho de otro modo, una vez determinada la política, puede dejarse el control del sistema a un dispositivo automático que la aplique.

Evidentemente, no todas las políticas ofrecen el mismo rendimiento. La política que da un resultado óptimo ante un horizonte determinado se denomina *política óptima*.

Normalmente, la optimalidad de una política depende del horizonte que se haya planteado. Cuando el horizonte es finito (supongamos a  $N$  transiciones), existen  $N$  políticas óptimas, una para cada etapa; dicho de otro modo, no es igual la política óptima cuando quedan 3 etapas para terminar que cuando quedan 2 o 1. La política óptima depende de la variable cronológica y se habla entonces de *política óptima a corto plazo*.

En cambio, si se trata de un problema con horizonte infinito o ilimitado, la política óptima no depende de la variable cronológica y se habla de *política óptima a largo plazo*.

## 2.2 Comparación con la programación lineal

La programación lineal, a la cual se dedica un volumen de esta colección, es probablemente el modelo cuantitativo más conocido en general; por ello, es interesante compararla con los procesos polietápicos, dado que sus características son muy diferentes y permiten, en conjunto, modelizar una gran cantidad de situaciones distintas.



|                                       | <b>Programación lineal</b> | <b>Procesos polietápicos</b> |
|---------------------------------------|----------------------------|------------------------------|
| <b>Comportamiento del sistema</b>     | Determinista               | Aleatorio                    |
| <b>Tipo de sistema</b>                | Estático                   | Dinámico                     |
| <b>Tipo de decisión</b>               | Única                      | Múltiple / No existe         |
| <b>Consideración del tiempo</b>       | Ninguna                    | Discreto / Continuo          |
| <b>Distinción a corto/largo plazo</b> | No                         | Sí                           |

Tabla 2.1.  
Comparación entre la programación lineal y los procesos polietápicos

## 2.3 Modelos polietápicos existentes

Existen tres tipos de modelos polietápicos a los cuales se dedican tres volúmenes de esta colección:

- La teoría de colas
- Las cadenas de Markov
- La programación dinámica

La *teoría de colas* es un modelo polietápico específico para una situación muy concreta, que presenta las características siguientes:

- El tiempo es continuo.
- No existe decisión en cada etapa; únicamente podemos observar el comportamiento del sistema y, por tanto, solo existe la decisión previa del diseño del sistema en función de los resultados que se pretenden obtener.
- El estado del sistema se caracteriza por el número de elementos que se hallan en el sistema en cada momento, ya sea esperando a ser servidos o recibiendo un servicio.
- Existe una remuneración, que es el tiempo que cada elemento pasa en el sistema.

Se trata, pues, de un modelo poco general, en el sentido de que solo puede aplicarse a un tipo de problema muy concreto. Sin embargo, su frecuencia de aparición en la realidad es tan grande que vale la pena tomarlo en consideración.

Las *cadenas de Markov* (CM) y la *programación dinámica* (PD) son dos modelos de investigación operativa, de orígenes muy distintos, que responden a conceptos teóricos muy diferentes. Sin embargo, ambos se pueden considerar casos particulares de los procesos polietápicos, tal como se han definido. Por ello, es útil tener en cuenta sus paralelismos y diferencias dentro de este marco general, a fin de facilitar su comprensión.



Un efecto de este planteamiento conjunto es que determinados conceptos tienen dos nombres, ya que proceden de las terminologías de la PD y las CM. Por ejemplo, puede indicarse que los términos *transición* y *remuneración* proceden de la literatura sobre CM, mientras que *etapa* y *rendimiento* proceden de la de PD.

Existen tres tipos de cadenas de Markov:

- Las cadenas de Markov simples (CM), sin remuneración ni decisión
- Las cadenas de Markov con remuneración (CMR) pero sin decisión
- Las cadenas de Markov con remuneración y decisión (CMRD)

La relación con el marco general descrito se puede establecer así:

Tabla 2.2.  
Relaciones entre la  
PD y las CM

|     | Con decisión | Sin decisión |
|-----|--------------|--------------|
| HT  | CMRD, PD     | CM, CMR      |
| NHT | PD           | (PD)         |

Así pues, las cadenas de Markov tratan con los sistemas homogéneos en el tiempo (HT), tengan o no decisión o rendimiento, mientras que la programación dinámica exige que exista decisión (y, por tanto, rendimiento), pero admite la posibilidad de sistemas NHT.

Así pues, los problemas HT con decisión pueden tratarse de las dos maneras ya que, como se verá, si se comparan las exposiciones en ambos volúmenes, se trata poco más que de un cambio de notación.

Los sistemas NHT sin decisión no entran en el esquema anterior, por lo cual no se abordan explícitamente en esta colección. Se trata de situaciones relativamente poco frecuentes y que no suelen tener gran interés; sin embargo, puede extenderse fácilmente la programación dinámica para tratarlos, si es necesario: basta con considerar que se trata de una decisión con una única alternativa.

## 2.4 Introducción al cálculo recurrente

En este apartado, hacemos una breve introducción al cálculo recurrente, que es una herramienta fundamental en el tratamiento de los modelos polietápicos, mediante un ejemplo en que no hay decisión.

El problema es el siguiente: Por una carretera muy estrecha donde no se puede avanzar, circulan  $n$  coches, cada uno de los cuales tiene una velocidad propia,  $v_i$ , a la cual circularía si no tuviera ningún otro coche delante que se lo



impidiera. Cuando se encuentra con un coche más lento, como no lo puede avanzar, tiene que adaptar su velocidad a la de este. Ello quiere decir que se forman "paquetes" de coches, en cada uno de los cuales el primero es el más lento del paquete. Dado que el orden de los coches y sus velocidades propias son magnitudes no correlacionadas (y, por tanto, estadísticamente independientes), el número de paquetes que se formen será una magnitud aleatoria y se tratará de saber el número esperado de paquetes que se formarán a largo plazo (es decir, en una carretera infinita).

Consideramos el más lento de los coches, es decir, el de menor  $v_i$ . Detrás de él, se forma un único paquete y, si se halla en la posición  $k$ , delante suyo habrá  $k-1$  coches, que formarán  $p(k-1)$  paquetes.

Por tanto:

$$p(n / k) = 1 + p(k - 1)$$

si el más lento va en posición  $k$ . Como el orden es aleatorio:

$$p(n) = \frac{1}{n} \sum_{k=0}^n [1 + p(k - 1)]$$

con  $p(0) = 0$ , evidentemente. Esta expresión permite el cálculo recurrente de  $p(n)$ . Por tanto, el problema está resuelto, pero con una expresión final no demasiado satisfactoria.

Consideremos ahora el más rápido. Salvo que vaya el primero, nunca formará un paquete, sino que se añadirá al de delante.

Por tanto:

$$p(n / 1) = 1 + p(n - 1)$$

$$p(n / \neq 1) = p(n - 1)$$

Entonces:

$$p(n) = \frac{1}{n} [1 + p(n - 1)] + \frac{n-1}{n} p(n - 1) = \frac{1}{n} + p(n - 1)$$

Y, dado que  $p(0) = 0$ , puede deducirse:

$$p(n) = \sum_{k=1}^n \frac{1}{k}$$

lo cual es una fórmula directa, que permite también el cálculo recurrente.



Este problema posibilita ver dos aspectos interesantes:

- a) Los fundamentos del cálculo recurrente y la forma de plantear los problemas de este tipo.
- b) Cómo uno u otro planteamiento, aunque matemáticamente equivalentes, pueden dar soluciones que permitan calcular los resultados más o menos fácilmente.

Estos son dos detalles a tener muy en cuenta a la hora de utilizar los modelos, en lo referente tanto a los modelos basados en las cadenas de Markov como a los modelos basados en la programación dinámica.



→ 3



# Conceptos básicos de las cadenas de Markov

## 3.1 Procesos markovianos

Un proceso markoviano es aquel que cumple la denominada *condición de Markov*.

Un sistema cumple la condición de Markov si su estado es variable en el tiempo, pero en todo momento se halla en uno y solo uno de un conjunto de estados determinado, y la probabilidad del nuevo estado cuando se produzca el cambio depende de un cierto número de estados anteriores.

Observemos que no se fija cómo ni cuándo se produce el cambio de estado, ni cuál debe ser este conjunto de estados, ni tampoco cómo se determinan las probabilidades. Así pues, podemos tener sistemas markovianos cuyo conjunto de estados sea finito o infinito, y el cambio podrá producirse en instantes discretos de tiempo o no.

## 3.2 Cadenas de Markov

En este volumen, estudiamos las denominadas *cadenas de Markov finitas* (abreviadamente, cadenas de Markov o CM) que son sistemas markovianos cuyo conjunto de estados es finito y los cambios de estado se producen en momentos discretos del tiempo.

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.

Por tanto, una cadena de Markov representa un sistema cuyo estado varía a lo largo del tiempo, en que cada cambio es 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 cabe la posibilidad de influir en las probabilidades de transición actuando adecuadamente sobre el sistema (decisión).

### 3.3 Modelización mediante las cadenas de Markov

Para construir un modelo que nos permita estudiar un sistema mediante las cadenas de Markov, hemos de definir los tres elementos clave que caracterizan las cadenas de Markov: estados, transiciones y probabilidades.

Formalmente, para definir una cadena de Markov finita, es preciso determinar los elementos siguientes:

- a) Un conjunto finito 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  $k$  anteriores. Se dice entonces que se trata de una cadena de Markov de orden  $k$ .

Los *estados* son una caracterización de la situación en que se halla el sistema en un instante determinado. Dicha caracterización puede ser tanto cuantitativa como cualitativa. Desde el punto de vista práctico, 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 solo pueden pertenecer al conjunto de estados del sistema. El sistema modelizado por la cadena posee una variable que cambia de valor en instantes discretos de tiempo, cambio que denominamos *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. Al ser el sistema estocástico, no se conoce con certeza el estado del sistema en un instante determinado, sino tan solo la probabilidad asociada a cada uno de los estados. Este hecho puede expresarse en términos de probabilidad condicional:

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

donde las  $e_i$  pertenecen al conjunto de estados posibles del sistema.



Observemos que, junto con la probabilidad de transición, hemos de fijar el momento en que *puede* producirse el cambio, es decir, hemos de definir la transición. Ello se complica porque, a veces, es posible que el nuevo estado sea el mismo que el anterior, por lo cual deben elegirse muy cuidadosamente los conceptos de estado y de transición, que normalmente han de definirse mutuamente.

### ***Ejemplo 3.1. La modelización de un semáforo***

Consideremos un semáforo que presenta el comportamiento siguiente: la luz permanece dos minutos y medio en rojo, luego pasa otros tres minutos en verde y, finalmente, permanece 30 segundos en ámbar antes de pasar nuevamente a rojo.

Se trata de un sistema determinista, ya que la secuencia de colores es fija, pero de momento obviaremos este detalle puesto que, a fin de cuentas, un sistema determinista siempre se puede considerar un sistema aleatorio muy simple, ya que sus probabilidades solo pueden ser 0 o 1.

La primera idea será, probablemente, definir tres estados (R, V, A) y la transición será que el semáforo cambie de color. Es una modelización correcta si solo se desea estudiar la secuencia de colores, pero no nos dará información sobre la probabilidad de encontrarlo en un color u otro cuando lleguemos a él, que es probablemente lo que deseamos saber.

Aunque existen técnicas que permiten resolver este problema con el modelo indicado, podemos hacerlo también con una definición adecuada de qué es un estado y qué es un sistema. Definamos como transición el hecho de que hayan transcurrido 30 segundos desde el último cambio de estado; entonces, el estado del sistema no será solo el color de la luz sino también el tiempo que lleva en el mismo. Dado que el ciclo completo dura 6 minutos, tendremos 12 estados, cuya secuencia será:

$$R1 \rightarrow R2 \rightarrow R3 \rightarrow R4 \rightarrow R5 \rightarrow V1 \rightarrow V2 \rightarrow V3 \rightarrow V4 \rightarrow V5 \rightarrow V6 \rightarrow A \rightarrow R1$$

con lo cual es inmediato obtener la probabilidad de encontrar el semáforo en rojo, verde o ámbar al llegar a él.

Este simple ejemplo ha de servir para observar la importancia de definir bien qué debe entenderse por estado y por transición, ya que de una definición correcta dependerá, en buena parte, la eficacia del modelo y su utilidad.

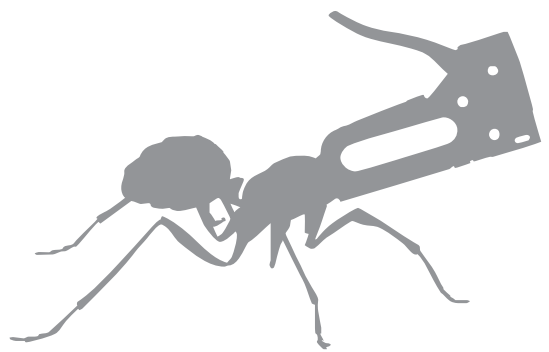


### 3.4 Cadenas de orden 1 y de orden superior

Las cadenas cuyo orden es 2 o más se denominan de *orden superior*.

En este volumen, tratamos solamente de las cadenas finitas de orden 1, es decir, en que la probabilidad del nuevo estado depende solo del estado anterior. Ello puede parecer una limitación importante, pero no lo es. Como veremos en la sección 4.1.3, las cadenas de orden superior se pueden reducir a cadenas de orden 1, por lo cual se pueden estudiar con las mismas técnicas que utilizamos para las de orden 1.

En lo sucesivo, se entenderá siempre que nos referimos a cadenas de orden 1 salvo que se indique explícitamente lo contrario.



→ 4



# Cadenas de Markov simples

## 4.1 Matriz de una cadena de Markov

### 4.1.1 Matrices de probabilidad de transición

Para el caso particular de una cadena de Markov de orden 1, la probabilidad condicional tiene esta forma:

$$p\{E_t / E_{t-1}\}$$

Si identificamos mediante unos índices  $i, j$  los estados del conjunto, podemos escribir esta probabilidad como  $p_{ij}$ , lo cual permite representar la ley condicional en forma de matriz.

En esta notación,  $p_{ij}$  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  $j$  solo depende del estado presente  $i$ .

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 denominada *matriz de probabilidades de transición*  $P$  o, más sencillamente, *matriz de la cadena*.

### 4.1.2 Matrices estocásticas

Las matrices que definen una cadena de Markov son cuadradas, 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.

Puesto que el sistema debe evolucionar al producirse la transición a alguno de los  $n$  estados posibles, las probabilidades de transición cumplirán la propiedad siguiente:



$$\sum_j p_{i,j} = 1$$

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

$$p_{i,j} \geq 0 \quad \forall i, j$$

Las matrices que cumplen estas propiedades se denominan *estocásticas*; de hecho, estas propiedades son precisamente la definición de *matriz estocástica*.

Es interesante recordar dos propiedades de este tipo de matrices, que serán útiles más adelante:

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

#### 4.1.3 Cadenas de orden superior

La limitación del estudio a las cadenas de orden 1 no constituye, en realidad, una limitación para el estudio de las cadenas de orden  $k > 1$ , puesto que se pueden reducir a ella. Veamos aquí cómo es posible.

Sea una cadena de orden  $k$  con  $n$  estados. Definimos una cadena de orden 1 cuyos estados están definidos precisamente por la secuencia de los últimos  $k$  estados de la cadena original; por tanto, se tratará de una cadena de  $n^k$  estados en la cual podremos representar las probabilidades de transición de la cadena original.

Observemos que no serán posibles todas las transiciones en la nueva cadena, sino solamente aquellas que sean compatibles con las transiciones de la cadena original. Dicho de otro modo: el estado original y el final han de ser tales que los primeros  $k-1$  estados del estado inicial coincidan con los  $k-1$  últimos del nuevo estado. Veamos dos ejemplos para aclararlo.

##### *Ejemplo 4.1. Sistema binario de orden 3*

Sea un sistema binario que puede adoptar solamente dos estados, que denominaremos convencionalmente 0 y 1, cuya probabilidad de transición es:



$$p(0) = \frac{1}{z+1}$$

$$p(1) = \frac{z}{z+1}$$

donde  $z$  representa la suma de los tres últimos estados. Se trata, pues, de una cadena con  $n = 2$  estados y orden  $k = 3$ . Definamos la cadena equivalente de orden 1 y  $2^3 = 8$  estados:

| Estado | Definición |
|--------|------------|
| $X_0$  | 000        |
| $X_1$  | 001        |
| $X_2$  | 010        |
| $X_3$  | 011        |
| $X_4$  | 100        |
| $X_5$  | 101        |
| $X_6$  | 110        |
| $X_7$  | 111        |

Tabla 4.1.  
Definición de los  
estados equivalentes  
del sistema binario

donde los valores indicados en la definición representan el estado de la cadena original en  $t-1$ ,  $t-2$  y  $t-3$ , respectivamente.

En estas condiciones, la matriz que define la cadena equivalente será:

|       | $X_0$ | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $X_0$ | 1     |       |       |       | 0     |       |       |       |
| $X_1$ | 1/2   |       |       |       | 1/2   |       |       |       |
| $X_2$ |       | 1/2   |       |       |       | 1/2   |       |       |
| $X_3$ |       | 1/3   |       |       |       | 2/3   |       |       |
| $X_4$ |       |       | 1/2   |       |       |       | 1/2   |       |
| $X_5$ |       |       | 1/3   |       |       |       | 2/3   |       |
| $X_6$ |       |       |       | 1/3   |       |       |       | 2/3   |
| $X_7$ |       |       |       | 1/4   |       |       |       | 3/4   |

Así, por ejemplo, desde el estado  $X_3 = 011$ , solo puede producirse la transición a los estados  $X_1 = 001$  (si aparece un 0) o  $X_5 = 101$  (si aparece un 1), cuyas probabilidades son  $1/3$  y  $2/3$ , respectivamente, dado que  $z = 2$ .

Las casillas en blanco corresponden, pues, a transiciones imposibles que matemáticamente han de tratarse como 0, ya que la probabilidad de que se produzcan es nula.

Observemos, sin embargo, que la casilla  $(X_0, X_4)$  contiene un 0, ya que se trata de un caso sustancialmente distinto. Dicha transición corresponde a pasar del estado  $X_0 = 000$  a  $X_4 = 100$ . Se trata de una transición teóricamente posible, ya que corresponde a la aparición de un 1 cuando los tres últimos estados han sido 0. Pero resulta que, según la ley de probabilidad de la cadena original,



dicha probabilidad es 0, ya que  $z = 0$ , es decir, tampoco puede observarse esta eventualidad. Por tanto, se trata de una transición que la definición de la cadena equivalente no prohíbe, pero cuya probabilidad es 0, debido a la definición de la cadena original.

Al haber solo dos nuevos estados posibles, ello implica que cuando el sistema alcanza el estado  $X_0$  ya no puede salir de él. Se trata de lo que definiremos más adelante como *estado absorbente* y que, en nuestro caso, quiere decir que, si la cadena ha permanecido durante tres transiciones en el estado 0, ya no podrá salir de él. El comportamiento a largo plazo de la cadena será, pues, una sucesión infinita de 0, ya que tarde o temprano acabará sucediendo dicha eventualidad. Las técnicas para el estudio de la cadena equivalente de orden 1 nos permitirán saber el valor esperado del número de transiciones antes de que ello ocurra.

#### Ejemplo 4.2. Evolución bursátil

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 aumente la cotización (estado A), la mantenga (estado B) o la baje (estado C) durante una semana dependerá de su comportamiento en las dos semanas anteriores. Las probabilidades se indican en la tabla siguiente:

Tabla 4.2.  
Probabilidades de  
evolución del valor  
durante una semana

| $E_{t-1}$ | $E_t$    | $p(A)$ | $p(B)$ | $p(C)$ |
|-----------|----------|--------|--------|--------|
| <b>A</b>  | <b>A</b> | 0,7    | 0,1    | 0,2    |
| <b>A</b>  | <b>B</b> | 0,3    | 0,5    | 0,2    |
| <b>A</b>  | <b>C</b> | 0,5    | 0,3    | 0,2    |
| <b>B</b>  | <b>A</b> | 0,4    | 0,1    | 0,5    |
| <b>B</b>  | <b>B</b> | 0,2    | 0,7    | 0,1    |
| <b>B</b>  | <b>C</b> | 0,1    | 0,3    | 0,6    |
| <b>C</b>  | <b>A</b> | 0,3    | 0,3    | 0,4    |
| <b>C</b>  | <b>B</b> | 0,3    | 0,1    | 0,6    |
| <b>C</b>  | <b>C</b> | 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 solo a otros 3 en el instante siguiente, tendremos 27 transiciones posibles. En definitiva, obtendremos la matriz de transición que se muestra en la tabla 4.3, cuyos 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 |

Tabla 4.3.  
Matriz de transición  
del ejemplo de la  
evolución bursátil

En este caso, cada fila tiene tres valores no vacíos, ya que el número de estados de la cadena original es 3.

## 4.2 Probabilidades estacionarias y tipos de cadenas de Markov

### 4.2.1 Probabilidades de transición de $k$ pasos. Teorema de Chapman-Kolmogorov

Puesto que las probabilidades de transición son estables en el tiempo, podemos interesarnos por 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 razonamiento siguiente: si, al cabo de  $m < k$  pasos, nos encontramos en el estado  $e$ , la probabilidad de alcanzar el estado  $j$  después de  $k$  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:

$$p_{i,j}^{(k)} = \sum_e p_{i,e}^{(m)} \cdot p_{e,j}^{(k-m)} \quad \forall m$$

expresión que, por supuesto, debe dar el mismo resultado para cualquier  $m < k$ .



En particular, 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  $k-1$ .

$$p_{i,j}^{(k)} = \sum_e p_{i,e} \cdot p_{e,j}^{(k-1)}$$

y, en particular:

$$p_{i,j}^{(2)} = \sum_e p_{i,e} \cdot p_{e,j}$$

lo cual indica que pueden obtenerse las matrices  $P^{(k)}$  de transición en  $k$  pasos a partir de las potencias de la matriz  $P$ .

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

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

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

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

#### 4.2.2 Probabilidades estacionarias

Calcular las potencias sucesivas de la matriz  $P$  nos permite conocer las probabilidades de los estados del sistema a condición de que se conozca el número de transiciones a calcular (horizonte finito). Pero, si dicho número es muy grande (horizonte infinito) o es desconocido (horizonte indeterminado), la situación es distinta. Distinguiremos, pues, entre el corto plazo ( $k$  conocido) y el largo plazo ( $k$  desconocido o que tiende a infinito).

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* y se definen como

$$P^* = \lim_{k \rightarrow \infty} P^k$$

suponiendo que dicho límite exista (lo cual, como se verá, no ocurre siempre).



Dichas probabilidades se denotan  $\pi_{ij}$  y la matriz de probabilidades de estado estable (o matriz a largo plazo),  $P^*$ . Esta matriz se conoce también como *matriz de probabilidades a priori*, ya que indica la probabilidad de hallar el sistema en un estado u otro cuando no se tiene ninguna información sobre su historia previa.

Puesto que, como hemos visto, las potencias de  $P$  definen las probabilidades en un número cualquiera de transiciones, podemos aproximarnos al estudio de las probabilidades estacionarias viendo qué ocurre si calculamos las potencias elevadas de algunas matrices  $P$ .

Dado que el número de estados es finito y el de transiciones es infinito, el sistema pasará por cada uno de los estados un número infinito de veces, por lo cual se puede definir la frecuencia con que se observará cada estado a largo plazo, que es precisamente la matriz  $P^*$ . Así pues, el comportamiento del sistema tiende a estabilizarse, en el sentido de que pasará por cada uno de los estados con una frecuencia regular, eso es, independiente del número de transiciones transcurridas. Veamos algunos de los casos que pueden darse.

#### 4.2.3 Cadenas ergódicas

Sea la cadena de Markov  $C1$  definida por la matriz  $P1$ , con dos estados y la siguiente matriz de probabilidades de transición:

$$P_1 = \begin{vmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{vmatrix}$$

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

$$P_1^{16} = \begin{vmatrix} 0,429 & 0,571 \\ 0,429 & 0,571 \end{vmatrix}$$

De hecho, estos valores se repiten siempre (decimales aparte) desde bastante antes de la decimosexta potencia. Se observa que las probabilidades de estado estable de los diferentes estados son independientes del estado de origen, razón por la cual la matriz de probabilidades estacionarias tiene todas las filas iguales. Tenemos entonces una cadena de Markov *regular*, en que las probabilidades estacionarias no dependen del estado inicial. Además, ninguna de las probabilidades vale 0. Tenemos entonces una cadena de Markov *ergódica*.



#### 4.2.4 Cadenas semiergódicas

Tenemos ahora una cadena C2 de cuatro estados, cuya matriz de probabilidades de transición es:

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

y calculamos la matriz de la transición decimosexta:

$$\mathbf{P}_2^{16} = \begin{vmatrix} 0,008 & 0,007 & 0,424 & 0,561 \\ 0,005 & 0,004 & 0,426 & 0,565 \\ 0 & 0 & 0,429 & 0,571 \\ 0 & 0 & 0,429 & 0,571 \end{vmatrix}$$

Se observa que todas las filas tienden a ser iguales (aunque no completamente, sobre todo las dos primeras), con una diferencia con respecto a las cadenas ergódicas: existen estados cuya probabilidad de estado estable tiende a 0 (eso es, que no aparecerán en el comportamiento a largo plazo). Por tanto, no se trata de una cadena ergódica. Sin embargo, sigue siendo cierto que todas las filas tienden hacia un mismo valor, por lo cual 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 iguales a 0 se denominan *semiergódicas*. Las cadenas ergódicas pueden considerarse un caso particular de las cadenas semiergódicas, en que no existen probabilidades de estado estable iguales a 0.

#### 4.2.5 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 decimosexta, podemos ver que, mientras algunas filas tienen el mismo comportamiento que las de los casos anteriores, otras tienden a 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 *semirregular*.

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

#### 4.2.6 Cadenas cíclicas

Sea la cadena C4, cuya matriz de probabilidades de transición se muestra a continuación:

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

Al revés de lo que ocurre en los casos anteriores, en que se observa que las sucesivas potencias de  $\mathbf{P}$  tienden rápidamente hacia unos valores determinados, en este caso se observa que, después de un número bastante elevado de transiciones (en este caso, basta con 10 transiciones), presenta un comportamiento distinto del de las cadenas anteriores.

Al ir obteniendo matrices de transición, se observa que estas 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:

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

$$\mathbf{P}_4^{2k+1} = \begin{vmatrix} 0 & 0,12 & 0,40 & 0,48 \\ 0 & 0 & 1 & 0 \\ 0 & 0,2 & 0 & 0,8 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

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



Obsérvese que, en este caso, la primera columna también es siempre 0, de modo que el estado 1 no aparecerá en las probabilidades a largo plazo. Esto quiere decir que la cadena considerada no es ergódica, aunque está claro que pueden existir cadenas cíclicas ergódicas, como veremos en ejemplos posteriores.

También hemos de preguntarnos qué ocurre con las probabilidades estacionarias en las cadenas cíclicas, si las potencias sucesivas de  $P$  no tienden hacia unos valores determinados. Más adelante, cuando estudiemos el cálculo sistemático de  $P^*$ , responderemos a esta pregunta.

#### 4.2.7 Clasificación de las cadenas de Markov

Los distintos casos examinados son solo algunos (los más importantes) de las diferentes situaciones que se pueden presentar a la hora de clasificar las cadenas de Markov. Más adelante se proporcionarán todas las definiciones, de forma completa y sistematizada.

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 siguiendo dos metodologías:

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

### 4.3 Análisis topológico de las cadenas de Markov

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



Esta metodología de análisis parte del grafo asociado a la matriz  $P$ , por lo cual será útil tener bien presentes los conceptos de la teoría de grafos desarrollados en otro volumen de esta colección.

### 4.3.1 Propiedades de estado

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

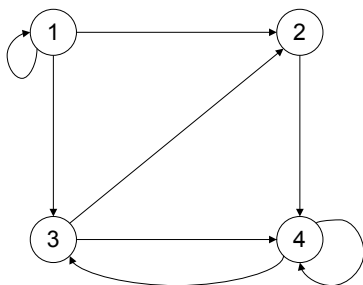
- El estado  $j$  es descendiente de  $i$  si, cuando iniciamos el proceso en  $i$ , existe una probabilidad no nula de que el proceso llegue a  $j$ . En este caso, diremos que existe un camino entre los estados  $i$  y  $j$ .
- Los estados  $i$  y  $j$  se comunican si  $i$  es descendiente de  $j$  y  $j$  es descendiente de  $i$ .
- Existirá un *ciclo* dentro de una cadena de Markov si existe un camino en la cadena que comunique el estado  $i$  consigo mismo. Dicho circuito se caracterizará por el número mínimo de transiciones que necesitará el sistema para regresar 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 al menos se puede acceder a él en 0 transiciones (circuito de longitud 0), con independencia de que además existan otros circuitos de mayor longitud.

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 un arco entre los vértices que representan los estados anterior y posterior de la misma. De este modo, en el diagrama se representan todas las situaciones en que un estado  $i$  es descendiente con respecto a  $j$ .

En la figura 4.1, se muestra el diagrama de transición de una cadena de Markov, junto con su matriz de transición.



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

Figura 4.1  
Diagrama de  
transiciones de  
estados



Algunas propiedades de estado son:

- 4 es descendiente de 1 (y no al revés).
- 3 y 4 se comunican.
- Pueden observarse los ciclos {2 4 3 2}, {3 4 3}, {4 4}.

#### 4.3.2 Propiedades de clase

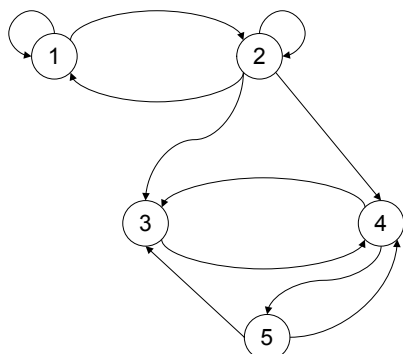
Dado que, por convenio, hemos establecido que un estado está siempre comunicado consigo mismo, la relación entre estados, de *estar comunicados*, es reflexiva, simétrica y transitiva, de modo que se trata de una relación de equivalencia. Por este motivo, podemos decir que un conjunto de estados comunicados 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 ningún estado de la clase tiene descendentes fuera de la misma. Ello significa que, 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. Los estados que pertenecen a alguna clase final se denominan *estados finales*, ya que son los únicos que pueden aparecer una vez el sistema ha alcanzado el régimen permanente.
- Las clases de equivalencia que no son clases finales se denominan *clases de paso (o transitorias)*. Los estados que pertenecen a las clases de paso son los *estados transitorios*, ya que a largo plazo desaparecerán sea cual sea el comportamiento de la cadena, de modo que solo pueden aparecer durante el régimen transitorio del sistema. El estudio de los estados transitorios debe realizarse de forma individual para cada uno de ellos, así que las clases de paso, como tales, tienen un interés muy limitado en el estudio de las cadenas de Markov.

Puesto que el sistema ha de ser capaz de evolucionar indefinidamente entre un número finito de estados, toda cadena ha de 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 4.2 muestra, sobre una cadena de Markov, una clase final y una clase de paso.



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

Figura 4.2.  
Clases de paso y final,  
y su matriz asociada

Las clases de equivalencia de la figura 4.2 son:

- La clase de equivalencia {1 2} es una clase de paso.
- La clase de equivalencia {3 4 5} es una clase final.

#### 4.3.3 Periodicidad de una clase final

Se define el *período de una clase final* como el máximo común divisor de las longitudes de los ciclos que pueden encontrarse en esa clase. Tenemos, entonces, dos tipos de clases finales:

- *Clases cíclicas* (o *periódicas*), que son las que tienen un período de valor  $p > 1$ .
- *Clases acíclicas* (o *aperiódicas*), que son las que tienen un período de valor  $p = 1$ .

La figura 4.3 muestra varias cadenas de Markov. La cadena (a) tiene ciclos de longitud 2 y 3, por lo cual es aperiódica, ya que  $\text{MCD}(2,3) = 1$ . En la cadena (b), se pueden observar ciclos de longitud 2 y 4, por lo cual su período 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 8; sin embargo, todos estos ciclos serán múltiplos de dos). Sin embargo, la cadena (c), similar a la anterior, tiene ciclos de longitud 1, 2 y 4, por lo cual es también aperiódica. La mera existencia de un ciclo de longitud 1 (en este caso, el 1-1) implica que la clase es aperiódica.

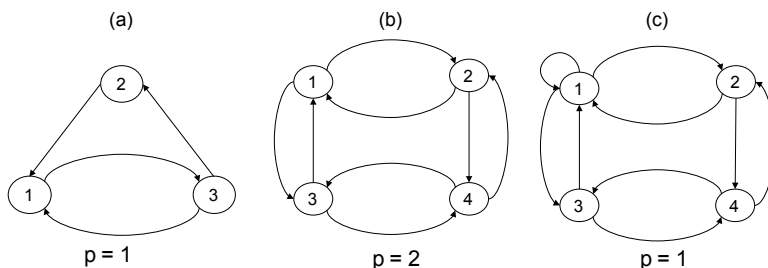


Figura 4.3.  
Ciclicidad en las  
clases finales de  
las cadenas de  
Markov



#### 4.3.4 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 el número de clases finales y la periodicidad o no de sus clases finales:

##### Clasificación según el número de clases finales

Podemos tener las situaciones siguientes 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 semiérgódica*.
- La cadena tiene más de una clase final y una o varias clases de paso. La cadena es, entonces, *no ergódica*.

En teoría, nos podemos encontrar con cadenas con varias clases finales, sin que existan clases de paso. En este caso, más que una única cadena, tenemos varias cadenas ergódicas sin relación entre ellas. Por esta razón, no se han 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 semiérgódica, según si tiene clases de paso o no.
- Si tenemos varias clases finales, todas ellas aperiódicas, tendremos una cadena de Markov *semirregular*.
- Si tenemos una sola clase final cíclica, se trata de una cadena *cíclica*.
- Si tenemos varias clases finales y todas ellas son periódicas, se trata 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, pues, los tres casos posibles de cadenas no ergódicas.

## Relación entre ambas clasificaciones

Evidentemente, la clasificación según el número de clases finales no es independiente de la clasificación según su periodicidad. Determinadas combinaciones reciben nombres especiales, como se observa en la tabla 4.4.

|                      | 1 CF             |              | > 1 CF       |
|----------------------|------------------|--------------|--------------|
|                      | Sin CP           | Con CP       |              |
|                      | Ergódica         | Semiergódica | No ergódica  |
| Todas $p = 1$        | Regular positiva | Regular      | Semirregular |
| Todas $p > 1$        | Cíclica          | Cíclica      | Policíclica  |
| $p = 1$ y<br>$p > 1$ | (No posible)     | (No posible) | Mixta        |

Tabla 4.4.  
Clasificación de las  
cadenas de Markov

En la figura 4.4, se muestran las cadenas (a), (b), (c) y (d) y, a continuación, se clasifican a partir del análisis topológico, mostrando las clases finales y las clases de paso.

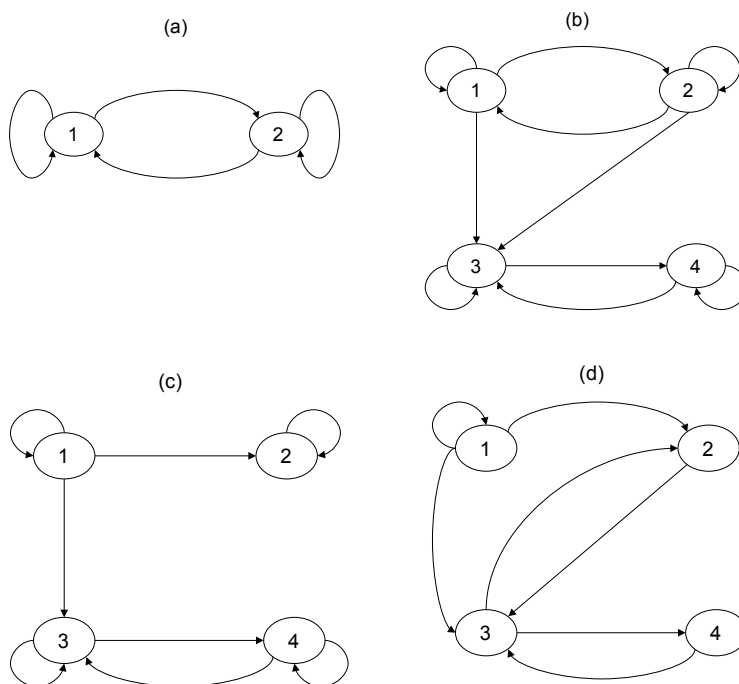


Figura 4.4.  
Ejemplos de  
clasificación de las  
cadenas de  
Markov

La cadena (a) tiene una sola clase final de  $p = 1$ , de modo que es una cadena regular positiva ergódica.



La cadena (b) tiene una clase final de  $p = 1$ :  $\{3\ 4\}$  y una clase de paso:  $\{1\ 2\}$ , por lo cual es una cadena regular semiérgica.

La cadena (c) tiene dos clases finales de  $p = 1$ :  $\{2\}$   $\{3\ 4\}$  y una clase de paso:  $\{1\}$ , así que se trata de una cadena semirregular.

Finalmente, la cadena (d) tiene una clase final de  $p = 2$ :  $\{2\ 3\ 4\}$  y una clase de paso:  $\{1\}$ , por lo cual es una cadena cíclica semiérgica de  $p = 2$ .

## 4.4 Análisis espectral

### 4.4.1 Propiedades del espectro de una matriz estocástica

El *análisis espectral* permite clasificar las cadenas de Markov a partir del espectro de su matriz de probabilidades, es decir, del conjunto de valores propios de la matriz  $P$ .

Para ello, es necesario enunciar algunas propiedades del espectro de toda matriz estocástica, que pondremos en relación con los conceptos que ya hemos visto en el análisis topológico:

- El número de valores propios de la matriz es igual al de filas y columnas, es decir, al número de estados del sistema.
- El módulo de todos los valores propios es inferior o igual a 1.
- Toda clase final de período  $p$  genera, como valores propios, las raíces  $p$ -ésimas de la unidad, cuyo cardinal es igual a la periodicidad de la clase.
- Solo los valores propios generados según la regla anterior pueden ser de módulo unitario.

Ello nos permite conocer el número y el tipo de clases finales de la cadena si conocemos los valores propios de su matriz de probabilidades de transición.

### 4.4.2 Clasificación según el análisis espectral

Para detectar el número de clases finales y su periodicidad, el análisis espectral parte de las propiedades anteriores, lo cual permite establecer las reglas siguientes:

- a) Cada clase final de período  $p$  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.
- b) El número de clases finales será igual a la multiplicidad de 1 como valor propio.



- c) Si existen las demás raíces  $p$ -ésimas de 1, indican la existencia de una clase final del período  $p$ .

Así pues:

- Una cadena de Markov *regular*, tanto *ergódica* como *semi-ergódica*, tendrá un único valor propio de valor 1. El resto de valores propios serán de módulo inferior a uno.
- Una cadena de Markov *semirregular* tendrá un valor propio 1 múltiple y el resto de los 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 *policíclica*) se caracterizará por tener, entre sus valores propios, uno o varios conjuntos de raíces  $p$ -ésimas de 1. Cada uno de estos conjuntos revelará la presencia, en la cadena, de una clase cíclica de período  $p$ .

Podemos tener situaciones como las descritas en los ejemplos 4.3 y 4.4:

#### Ejemplo 4.3

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

$$\{-0,8 \quad 0,4 \quad +1 \quad -1 \quad +1 \quad -0,3 \quad 0,01 \quad 0,5\}$$

es una cadena de Markov no ergódica mixta: cuenta con una clase final de período 1 y con una clase final de período 2.

Dado que existen ocho valores propios, la cadena ha de 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 el hecho de ser una cadena mixta, se trata de un caso no ergódico.

#### Ejemplo 4.4

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

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

es 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 período 4, mientras que la segunda es de período 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 potencias pares e impares



de  $P$ , mientras que en la clase de período 4 se repetirán cada 4 potencias; por tanto, la matriz  $P^k$  se repetirá cada 4 potencias.

#### 4.4.3 Comparación entre los dos tipos de análisis

El cálculo de los valores propios de una matriz no es sencillo si no se dispone del software matemático adecuado. Por ello, en general y para casos simples, será preferible utilizar manualmente el análisis topológico, el cual, además, proporciona más información que el espectral.

El análisis espectral no aporta información sobre las clases de paso ni tampoco sobre la naturaleza de cada estado y, por tanto, tampoco informa si se trata de una cadena ergódica o semiergódica. 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.

Si el caso es bastante complejo, es decir, si la cadena tiene decenas o centenares de estados, ninguna de ambas técnicas podrá abordarse sin contar con los medios adecuados; para el análisis topológico, es preciso disponer de un software de grafos.

Sin embargo, no debe por ello menospreciarse el análisis espectral, puesto que, una vez sabemos cuántas clases finales existen y con qué periodicidad, suele bastar con calcular una potencia de  $P$  lo suficientemente elevada para obtener una información equivalente. En general, elevando la matriz al cuadrado tantas veces como estados tenga la matriz, podremos obtener ya información suficiente, sin olvidar que existen técnicas para calcular la matriz  $P^*$  directamente, como se explicará en el apartado 4.6.

### 4.5 Significado de la clasificación

#### 4.5.1 Significado de la ergodicidad

La ergodicidad define cómo será (o cómo puede ser) el comportamiento del sistema a largo plazo.

En una *cadena ergódica*, todos los estados se pueden dar a largo plazo. Es decir, 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, y solo durante este se conserva la memoria del estado inicial.

En las *cadenas semiergódicas*, los comportamientos a largo y a corto plazo son cualitativamente distintos. Existen unos estados transitorios que desaparecen a largo plazo y, por tanto, solo pueden aparecer mientras el sistema no ha



alcanzado su régimen permanente. Una vez alcanzado este, 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 que vaya a parar el sistema. Ello significa que la probabilidad estacionaria de cada estado depende de su situación inicial, que puede llegar a determinar que ciertos estados no se puedan alcanzar jamás si se parte de un cierto conjunto de estados.

#### 4.5.2 Significado de la periodicidad

Una clase final acíclica representa una situación en 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  $P$ ). Podríamos decir que se trata de un sistema estocástico “puro”, en el sentido de que su comportamiento no se puede describir de forma más precisa que las meras probabilidades. El significado de las probabilidades estacionarias representa, en las clases finales, la probabilidad real de que en una transición cualquiera el sistema se halle en un estado u otro.

Por tanto, las cadenas regulares y semirregulares representan sistemas que acaban desembocando en una situación de este tipo.

Un caso particular interesante dentro de las clases finales acíclicas es el de los denominados *estados absorbentes*: se trata de estados que constituyen, por sí mismos, una sola clase final, puesto que la única transición posible es ir de nuevo al mismo. Matemáticamente, significa que la fila correspondiente de  $P$  está toda con 0 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, en el que ya no puede evolucionar más.

En las clases finales cíclicas, se puede establecer una partición de sus estados en  $p$  subclases, de modo que la evolución de la cadena dentro de esta clase consistirá en que pasará sucesivamente por un estado de la subclase 1, luego por uno de la subclase 2, etc., hasta llegar a la subclase  $p$ , después del cual regresará 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 son un caso intermedio entre un sistema estocástico puro, en 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 tanto, es predecible y calculable. En este caso, no se conoce exactamente cuál será el estado siguiente, 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 un estado u otro es la probabilidad estacionaria, pero ello solo es fruto de nuestro desconocimiento del número de transiciones realizadas. Si conociéramos este número (o, más precisamente, su módulo  $p$ ), las probabilidades serían distintas, puesto que las potencias de  $P$  evolucionan cíclicamente.

Una observación final sobre las cadenas policíclicas. Hemos dicho que las potencias de  $P$  se comportarán cíclicamente, al repetirse en un ciclo igual al mínimo común múltiplo de los diferentes valores de  $p$ . Sin embargo, ello no tiene nada que ver con el comportamiento a largo plazo del sistema que modeliza. Por ejemplo, en una cadena con dos clases finales de  $p = 2$  y  $p = 3$ , las potencias se repetirán cada 6 iteraciones, pero el sistema real se comportará con un ciclo de 2 o 3 iteraciones, según cuál sea la clase final donde vaya a parar finalmente.

## 4.6 Cálculo de las probabilidades estacionarias

### 4.6.1 Probabilidades estacionarias y comportamiento a largo plazo

La clasificación de la cadena de Markov permite identificar la forma de la matriz de probabilidades estacionarias  $P^*$  y facilitar así su cálculo exacto, sin necesidad de obtener una potencia elevada de la matriz  $P$  (procedimiento, por otra parte, desaconsejable 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  $P$  es de una cadena ergódica o semi-ergódica,  $P^*$  es una matriz con todas las filas iguales, lo cual facilita enormemente su cálculo. Para las cadenas no ergódicas, al ser algunas de las filas de  $P^*$  diferentes del resto, su cálculo será más complejo.

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

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

de modo que basta con hallar este límite para conocer  $P^*$  si la cadena no tiene clases finales cíclicas.

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



#### 4.6.2 Probabilidades estacionarias para las cadenas ergódicas y semiergódicas

##### Cadenas regulares positivas

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, de modo que la probabilidad de que el sistema se halle en uno u otro estado no dependerá de dicho estado. Todas las filas de  $P^*$  en una cadena regular son iguales, por lo cual la matriz adopta la forma:

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

Observemos que todos los estados son posibles a largo plazo y, por tanto, todos los valores de  $P^*$  serán estrictamente positivos; de ahí el nombre de “regular positiva” para esta clase de cadenas.

Dado que la matriz  $P^*$  es el límite de las potencias de  $P$ , ha de cumplir la ecuación:

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

Desarrollando esta ecuación matricial para 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 cual hemos de añadir la ecuación, propia de toda matriz estocástica:

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

así que tenemos  $n+1$  ecuaciones para  $n$  incógnitas. Sin embargo, las ecuaciones procedentes del producto de matrices no son linealmente independientes, por lo cual están indeterminadas: podemos tener otras soluciones a partir de un conjunto de soluciones inicial, multiplicando todos los valores de este por un escalar. Es la última ecuación la que hace que  $P^*$  sea estocástica. En consecuencia, deberá eliminarse alguna de las procedentes del producto de matrices, pero nunca la que hace que la suma de las  $\pi_i$  sea igual a 1.

**Ejemplo 4.5. Probabilidades estacionarias para una cadena regular ergódica**

La cadena C1, de matriz de probabilidades de transición:

$$P_1 = \begin{vmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{vmatrix}$$

Tiene una matriz de probabilidades estacionarias de la forma:

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

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

$$0,2 \cdot \pi_1 + 0,6 \cdot \pi_2 = \pi_1$$

$$0,8 \cdot \pi_1 + 0,4 \cdot \pi_2 = \pi_2$$

$$\pi_1 + \pi_2 = 1$$

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

$$\pi_1 = 3/7 = 0,429$$

$$\pi_2 = 4/7 = 0,571$$

**Cadenas semiergódicas**

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  $\pi\{\text{clases de paso}\} = 0$ , ya que, con toda seguridad, tras infinitas transiciones el sistema se encontrará en una clase final.

**Ejemplo 4.6. Probabilidades estacionarias para una cadena regular semiergódica**

La matriz de probabilidades estacionarias de la cadena C2 es:

$$P_2 = \begin{vmatrix} 0,5 & 0,4 & 0,1 & 0 \\ 0,3 & 0,3 & 0,4 & 0 \\ 0 & 0 & 0,2 & 0,8 \\ 0 & 0 & 0,6 & 0,4 \end{vmatrix}$$



Puesto que hay una sola clase final formada por los estados 3 y 4, el sistema a largo plazo estará necesariamente en uno de ellos. Por tanto, tenemos que:

- Todas las filas serán iguales, ya que la clase final es única.
- Los estados 1 y 2 son transitorios, de modo que sus probabilidades estacionarias serán 0.

Basta, pues, con calcular las probabilidades estacionarias de la cadena ergódica formada por los estados 3 y 4, que ya hemos visto en el ejemplo 4.5, y añadir 0 en las columnas de los estados transitorios. Así pues, cada una de las filas de la  $P^*$  será:

$$P^*_2 = [0 \ 0 \ 3/7 \ 4/7] = [0 \ 0 \ 0,429 \ 0,571]$$

Por supuesto, podríamos haber obtenido este mismo resultado con el procedimiento descrito anteriormente: formar un sistema con la ecuación  $\sum \pi_i = 1$  y  $n-1$  ecuaciones como resultado de desarrollar la igualdad matricial  $P^* \cdot P = P^*$ , pero es evidente que resulta mucho más sencillo el procedimiento que se indica aquí.

### Cadenas cíclicas

En las cadenas regulares y semirregulares, se cumple:

$$P^* = \lim_{k \rightarrow \infty} P^k$$

Sin embargo, en las cadenas cíclicas no existe este límite, en el sentido de Cauchy. Las potencias de  $P$  evolucionan cíclicamente entre tantas matrices de transición a largo plazo como el período  $p$  de la cadena, de modo que tendremos un conjunto de  $p$  matrices a largo plazo:

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

Recordemos que el significado de  $P^*$  es el de las probabilidades de hallar el sistema en uno u otro estado cuando lleva funcionando un tiempo indefinidamente largo. Puesto que el número transiciones (módulo  $p$ ) es desconocido, podemos definir  $P^*$  de esta forma:

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

Afortunadamente, el proceso de cálculo que ya hemos explicado para las cadenas regulares y semirregulares sirve también aquí, así que no hace falta repetirlo. Matemáticamente, la diferencia es que ahora  $P^*$  ya no es el límite de las potencias de  $P$ , sino el promedio de las matrices de evolución a largo plazo de cada uno de los períodos; pese a ello, el significado es el mismo.



Evidentemente, aunque el número exacto de transiciones sea desconocido, si conocemos su valor módulo  $p$  ya no será válida esta  $P^*$ , sino la  $P^{*(k)}$  que corresponda.

**Ejemplo 4.7. Probabilidades estacionarias para una cadena cíclica**

Como se ha visto anteriormente, la cadena C4, cuya matriz de probabilidades de transición de un paso se muestra a continuación, tiene una clase cíclica de  $p = 2$ .

$$P_4 = \begin{vmatrix} 0,5 & 0,4 & 0,1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0,2 & 0 & 0,8 \\ 0 & 0 & 1 & 0 \end{vmatrix}$$

Para hallar la matriz de probabilidades estacionarias de  $P_4$ , procederemos formalmente como si la clase final no fuese cíclica y, puesto que se trata de una cadena semiérgódica, ya sabemos que tendrá las filas iguales, con  $\pi_1 = 0$ .

$$\begin{aligned}\pi_2 &= 0,2 \pi_3 \\ \pi_3 &= \pi_2 + \pi_4 \\ \pi_4 &= 0,8 \pi_3 \\ \pi_2 + \pi_3 + \pi_4 &= 1\end{aligned}$$

Resolviendo el sistema:  $P^*_4 = [0 \ 0,1 \ 0,5 \ 0,4]$

La diferencia es que ahora no se trata del límite de las potencias de  $P_4$ , sino del promedio de las dos matrices de probabilidades a que tiende el comportamiento del sistema:

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

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

Para determinar las matrices  $P^{nk}$ , basta con tener en cuenta las igualdades siguientes:

$$\begin{aligned}P^* &= 1/p (P^k + P^{k+1} + \dots + P^{k-1}) \\ P^{k+1} &= P^k \cdot P; P^{k+2} = P^{k+1} \cdot P; \dots; P^k = P^{k-1} \cdot P\end{aligned}$$



Nótese la posibilidad de error que existe si se clasifica la cadena simplemente calculando potencias de  $P$ , tentación en que es fácil de caer si se dispone de un ordenador con una hoja de cálculo. Lo más fácil es elevar  $P$  al cuadrado, hacer lo mismo con  $P^2$ , y así sucesivamente. Entonces, no se verá la aparición de ciclos de longitud 2, como en este caso. Compruébese que en la potencia  $P^{16}$  ya se obtiene  $P^{2^k}$ , por lo cual esta matriz puede confundirse con  $P^*$ , lo cual evidentemente es erróneo.

#### 4.6.3 Probabilidades estacionarias para las cadenas no ergódicas

Aunque las cadenas semirregulares son más complejas, la naturaleza de las distintas clases facilita también el cálculo de la matriz  $P^*$ . He aquí unas reglas muy simples, que nos permiten obtener un primer "esquema" de  $P^*$

- Las columnas de  $P^*$  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 0. 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  $\pi_i$  de las transiciones con destino en estados pertenecientes a clases de paso serán 0, y los  $\pi_i$  de las transiciones con destino en las clases finales serán diferentes de 0. Cada estado perteneciente a una clase de paso tendrá una fila diferente en  $P^*$ .

Así pues, hay que proceder siguiendo estos dos pasos:

1. Se hallan las matrices  $P^*$  correspondientes a las cadenas regulares positivas definidas por cada una de las clases finales (según el procedimiento descrito en 4.6.2) y se componen con ellas las correspondientes filas de sus estados.
2. 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 clases de paso tenga la cadena.

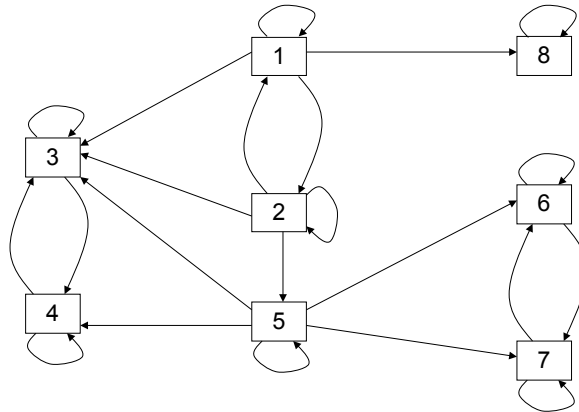
**Ejemplo 4.8. Matriz  $P^*$  para una cadena semirregular**

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

$$P = \begin{pmatrix} 0,7 & 0,1 & 0,1 & 0 & 0 & 0 & 0 & 0,1 \\ 0,4 & 0,3 & 0,1 & 0 & 0,2 & 0 & 0 & 0 \\ 0 & 0 & 0,7 & 0,3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,6 & 0,4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,1 & 0,3 & 0,3 & 0,1 & 0,2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,5 & 0,5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,4 & 0,6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

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

Figura 4.5.  
Análisis topológico  
de la cadena



Igual que hemos hecho para el caso semiergódico, empezamos calculando las  $P^*$  de las matrices ergódicas regulares de las tres clases finales, lo cual nos da las filas correspondientes a todos los estados finales.

Escribimos ahora una primera versión de la matriz  $P^*$ :

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



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

Observemos que las distintas filas se han completado con valores de 0 en:

- Las columnas correspondientes a los estados transitorios
- Las columnas de los estados finales en filas que corresponden a estados de otra clase final

ya que todos ellos representan situaciones inalcanzables a largo plazo, si el estado inicial es el correspondiente a la fila indicada. Falta tan solo calcular los valores de  $P^*$  para las filas de los estados transitorios en las columnas de los estados finales.

En este caso, del diagrama de transiciones se puede deducir fácilmente que desde el estado 5 no se puede ir a la clase {8}, así que  $t = 0$ .

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

Por tanto, nos quedan 14 incógnitas, de modo que hemos de establecer 14 ecuaciones para hallar estos valores. Escribimos, en primer lugar, las siempre obligadas:

$$a + b + c + d + e = 1$$

$$i + j + k + l + m = 1$$

$$p + q + r + s = 1$$

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

$$P P^* = P^* P = P^*$$

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 son simples identidades que no sirven para nuestro propósito, y también existen combinaciones lineales entre ellas que no podemos utilizar.

Aunque la tarea parece complicada, de las propiedades de estas matrices se pueden deducir algunas relaciones que permiten establecer fácilmente algunas de las ecuaciones necesarias. Cabe señalar que todas ellas se pueden hallar a partir de las expresiones matriciales indicadas, pero de esta manera se obtienen de forma más sencilla, 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 denominaremos  $P_{1,34}$ . Una vez el sistema ha ido a parar a dicha clase, la probabilidad de 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 cual se concluye que  $a$  y  $b$  son proporcionales a  $2/3$  y  $1/3$ .

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:

$$a / b = i / j = p / q = (2/3) / (1/3) = 2$$

$$c / d = k / l = r / s = 4 / 5$$

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 solo puede hacer dos cosas: o bien se mantiene en él o bien va directamente a una clase final. Si seguimos la notación que hemos establecido anteriormente, es evidente que  $P_{5,34} + P_{5,67} = 1 - 0,3 = 0,7$  y, como  $p + q = P_{5,34} = (0,1 + 0,3) / 0,7$ , análogamente para  $P_{5,67} = r + s$ . Así pues, podemos escribir:

$$p + q = 4/7$$

$$r + s = 3 / 7$$

Estas igualdades no proporcionan dos ecuaciones más, sino solo una, ya que hay combinación lineal con las anteriores. No obstante, basta con esta 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 solo a las 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 de este estilo nos permitiría establecer expresiones semejantes para los estados 1 y 2, pero ello conduce, en general, a situaciones más complejas que usualmente no vale 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  $P^* = P \cdot P^*$  en grupos de la misma columna:

Empecemos por la pareja  $a, i$ :

$$a = 0,7a + 0,1i + 0,1 \cdot (2/3) \quad \Rightarrow \quad 3a - i = 2/3$$

$$i = 0,4a + 0,3i + 0,1 \cdot (2/3) + 0,2p \quad \Rightarrow \quad 7i - 4a = 1/7$$

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

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

Repetamos la técnica con  $c$  y  $k$ :

$$c = 0,7c + 0,1k \quad \Rightarrow \quad 3c - k = 0$$

$$k = 0,4c + 0,3k + 0,2r \quad \Rightarrow \quad 7k - 4c = 8/21$$

de modo que  $c = 8/357; d = 10/357; k = 8/119; l = 10/119$

Y, finalmente, para  $e$  y  $m$ :

$$e = 0,7e + 0,1m + 0,1 \cdot 1 \quad \Rightarrow \quad 3e - m = 1$$

$$m = 0,4e + 0,3m \quad \Rightarrow \quad 7m - 4e = 0$$

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:



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

De esta matriz, es inmediato deducir las probabilidades de que el sistema vaya a una u otra clase final, en función del 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:

- En este caso la existencia de un estado de paso del cual solo se puede salir para ir a alguna clase final permite el cálculo directo de la fila correspondiente; naturalmente, eso no tiene por qué ocurrir en general.
- La utilización en cada paso de los valores ya obtenidos permite reducir los sistemas de ecuaciones resultantes y, por tanto, facilita el cálculo incluso manual.
- La técnica de escribir las ecuaciones para los elementos de la misma columna permite, en general, trocear el sistema en otros más pequeños y así reducir notablemente la dificultad del cálculo.

#### **Ejemplo 4.9. Matriz $P^*$ para una cadena de Markov semirregular**

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

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



El lector puede comprobar fácilmente, mediante un análisis topológico, que se trata de una cadena semirregular, en la cual 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:

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

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:

$$\pi_1 = \pi_3 = 1/3$$

$$\pi_2 = \pi_4 = 2/3$$

Para hallar las dos primeras filas de  $\mathbf{P}^*$ , se trata de plantear las dos primeras filas de la ecuación matricial:

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

El resultado es la matriz:

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

#### **Ejemplo 4.10. Matriz $\mathbf{P}^*$ para una cadena mixta**

Veamos finalmente un caso, más complejo, de cadena mixta. Sea la cadena de Markov definida por la siguiente matriz de probabilidades de transición:



$$P = \begin{bmatrix} 0,7 & 0,2 & 0,1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0,9 & 1/20 & 0 & 0 & 1/20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,8 & 0,1 & 0,1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3/4 & 0 & 0,1 & 0 & 0 & 0 & 0 & 0,1 & 0 & 1/20 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2/3 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,7 & 0,3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,4 & 0,6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

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 que la clase {11,12} es aperiódica, mientras que las otras dos tienen ciclos de longitud 3 y 2, respectivamente.

El lector debe efectuar los cálculos por su cuenta para determinar los valores de  $P^*$  y comprobará que obtiene los valores siguientes (se han suprimido las columnas de clases de paso).

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

Sin embargo, el significado de estas probabilidades es diferente según el estado (fila) de que se trate: las filas 5, 11 y 12 corresponden a clases aperiódicas y, por tanto, son probabilidades en sentido estricto. Pero las clases {6, 7, 8, 9, 10} y {13, 14, 15} tienen periodicidades 3 y 2, respectivamente. Veamos cómo podemos calcular las matrices correspondientes a las iteraciones de estas clases cíclicas.



La clase {13, 14, 15} tiene periodicidad 2 y sus dos subclases son {13} y {14, 15}. Por tanto, si el número de iteraciones es par, el sistema se hallará en la misma subclase inicial y, si es impar, en la otra. Por otra parte, si se halla en la subclase {14, 15}, las probabilidades de uno u otro estado son, precisamente, las indicadas por la matriz P. Por tanto:

$$P^{2k*} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0,4 & 0,6 \\ 0 & 0,4 & 0,6 \end{vmatrix} \quad P^{2k+1*} = \begin{vmatrix} 0 & 0,4 & 0,6 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{vmatrix}$$

cuyo promedio es, efectivamente,  $P^*$ .

Análogamente, obtenemos los valores para la clase {6, 7, 8, 9, 10}:

$$P^{3k*} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2/3 & 1/6 & 1/6 & 0 \\ 0 & 2/3 & 1/6 & 1/6 & 0 \\ 0 & 2/3 & 1/6 & 1/6 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} \quad P^{3k+1*} = \begin{vmatrix} 0 & 2/3 & 1/6 & 1/6 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{vmatrix}$$

$$P^{3k+2*} = \begin{vmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 2/3 & 1/6 & 1/6 & 0 \end{vmatrix}$$

Si se introduce la matriz P en una hoja de cálculo y se eleva al cuadrado unas cuantas veces (con seis veces será suficiente), se observa que la fila 5 no varía y que 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  $P^2$ . En cambio, en las filas 6-10, se observa que los mismos valores se van "moviendo" cíclicamente de unas filas a otras sin que, sin embargo, lleguen a observarse todos los valores en todas las filas.

Este es un efecto de los ciclos de las clases finales, que no observamos fácilmente si nos limitamos a calcular potencias muy elevadas de P simplemente elevando al cuadrado. Solamente si se multiplica por P para obtener la potencia siguiente, se puede ver cómo, 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 que hemos hecho 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 que el significado de  $P^*$  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.

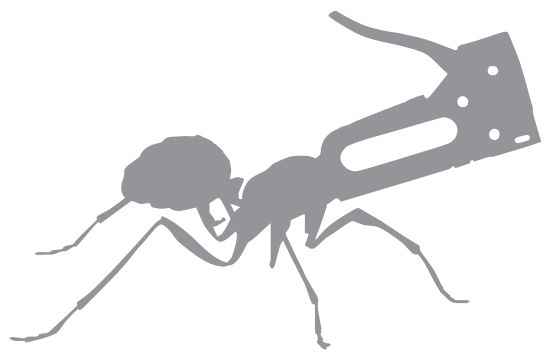
#### 4.6.4 Modelización con cadenas de Markov simples

Con el cálculo de las probabilidades estacionarias, hemos completado las aportaciones de las cadenas de Markov simples. El estudio de los sistemas descritos con estos modelos se limita a analizar las probabilidades de transición a corto y a largo plazo.

El resumen de lo presentado hasta ahora es:

- Para conocer la probabilidad de transición en un número finito (y pequeño) de transiciones, basta con calcular la potencia correspondiente de  $P$ . En estas condiciones, la probabilidad de cada estado dependerá del estado inicial. En esta situación, nos hallamos todavía en el régimen transitorio, en que el sistema conserva la memoria del estado de partida.
- Para conocer la probabilidad a largo plazo, debe hallarse la matriz  $P^*$  correspondiente. Si el sistema es no ergódico, puede conservar parcialmente la memoria del estado inicial si este era de tal naturaleza que no permitía ir a parar a determinadas clases finales.
- En las clases finales periódicas, debe tenerse en cuenta que, si conocemos el número de transiciones realizadas (o, al menos, su valor módulo  $p$ ), podremos obtener mayor información sobre las probabilidades de los estados. En este sentido, las cadenas cíclicas representan una situación intermedia entre un sistema determinista, en que todo es conocido, y un sistema puramente aleatorio, en que no podemos predecir más que las probabilidades.

Puede que el lector haya quedado un tanto insatisfecho con estas conclusiones. Hay una pregunta bastante inmediata a la cual no hemos dado respuesta: ¿Cuándo finaliza el corto plazo? O, dicho de otro modo, si existen estados transitorios, ¿cuántas iteraciones cabe esperar que transcurran antes de caer en una clase final? Para responder a estas preguntas, hemos de estudiar las cadenas de Markov con remuneración, a las cuales dedicamos el capítulo siguiente.



→ 5



## Cadenas de Markov con remuneración

### 5.1 Introducción

Las cadenas de Markov con remuneración (CMR) se caracterizan por que, junto a la matriz  $P$  de probabilidades de transición, existe otra matriz  $R$  de remuneraciones.

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

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

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

El problema a que nos enfrentamos ahora es encontrar la remuneración esperada de un proceso estocástico representable por una cadena de Markov, tanto a largo plazo como a corto plazo. Dado que el comportamiento del sistema es aleatorio, no podremos calcular el valor de la remuneración que se obtendrá; solamente podemos calcular el valor esperado, o sea, su esperanza matemática.



## 5.2 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 técnica conocida como *iteración en el espacio de los estados*. También se introducirá la *normalización*, procedimiento que ayuda a simplificar el proceso de cálculo.

### 5.2.1 Iteración en el espacio de los estados

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

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

donde:

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

En estas expresiones, hemos supuesto que, como es habitual, la remuneración es un concepto sumable, es decir, que la remuneración total es la suma de las remuneraciones obtenidas en cada transición. En realidad, bajo ciertas condiciones, se puede utilizar cualquier otro tipo de operación, como veremos.

En las dos expresiones anteriores,  $j$  representa los  $n$  estados posibles que podrá tomar el sistema en la transición siguiente, en la cual faltarán  $N-1$  transiciones para acabar. En cuanto a la  $q_i$ , nótese que se obtiene cada uno de sus componentes como la suma de los productos de la fila  $i$  de las matrices de probabilidad de transición  $P$  y de remuneraciones  $R$ , es decir, que no se trata de ningún producto matricial (al menos en el sentido habitual). Representa la remuneración esperada en la primera transición, si el sistema se halla en el 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  $v(0)$ .

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

### 5.2.2 Normalización

Si las remuneraciones no son valores de signos distintos (e incluso la mayoría de las veces aunque haya valores de signos distintos), los valores de las  $v(N)$  tenderán a crecer indefinidamente, lo cual puede suponer, como mínimo, un engorro para los cálculos y para la interpretación de los resultados. La normalización permite evitar estos inconvenientes.

La técnica de la *normalización* consiste en restar sistemáticamente a los valores obtenidos en cada transición una cantidad determinada y operar en lo sucesivo con la diferencia. Por ello, las ecuaciones de recurrencia quedan de la forma siguiente:

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

$$\mathbf{v}'(N) = \mathbf{v}(N) - \mathbf{g}_N$$

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

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

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

Esta técnica tiene dos ventajas: en primer lugar, evita que los valores absolutos de  $v(N)$  crezcan de forma continuada, 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 en 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 cómo se fija la  $\mathbf{g}_N$ , valor que tienen 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:

$$g_N = v_k(N), \text{ donde } k \text{ es un estado elegido previamente.}$$

Ambas reglas conducen a los mismos resultados a largo plazo (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  $v'(N)$  son siempre mayores o iguales a 0, lo cual evita errores en los cálculos manuales, mientras que la regla 2 facilita la comparación de los vectores  $v'$ .

### *Ejemplo 5.1. Iteración en el espacio de los estados y normalización*

La tabla 5.1 muestra las sucesivas iteraciones realizadas para determinar el comportamiento de la cadena, con las matrices de probabilidades de transición  $P$  y de remuneraciones  $R$  siguientes:

$$\mathbf{P} = \begin{vmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{vmatrix} \quad \mathbf{R} = \begin{vmatrix} 10 & 20 \\ 40 & -20 \end{vmatrix} \quad \mathbf{q} = \begin{vmatrix} 18 \\ 16 \end{vmatrix}$$

donde ya hemos calculado el valor del vector  $q$  que nos será necesario. Para calcular la remuneración, es preciso conocer las condiciones de contorno, esto es, el vector  $v(0)$ . Este vector representa la remuneración obtenida *después* de la última transición, es decir, una vez finalizada la evolución del sistema, que supondremos que es 0.

Apliquemos la expresión de recurrencia con normalización a nuestro caso:

$$v(N) = q + P \cdot v'(N-1)$$

$$v'(N) = v(N) - g_N$$

y obtendremos la tabla siguiente para las doce últimas iteraciones:



|                      | <b>v(0)</b> | <b>v(1)</b> | <b>v'(1)</b> | <b>v(2)</b> | <b>v'(2)</b> | <b>v(3)</b> | <b>v'(3)</b> | <b>v(4)</b> | <b>v'(4)</b> | <b>v(5)</b> | <b>v'(5)</b> | <b>v(6)</b> | <b>v'(6)</b> |
|----------------------|-------------|-------------|--------------|-------------|--------------|-------------|--------------|-------------|--------------|-------------|--------------|-------------|--------------|
| <b>v<sub>1</sub></b> | 0,00        | 18,00       | 2,00         | 18,40       | 1,20         | 18,24       | 1,52         | 18,30       | 1,39         | 18,28       | 1,44         | 18,29       | 1,42         |
| <b>v<sub>2</sub></b> | 0,00        | 16,00       | 0,00         | 17,20       | 0,00         | 16,72       | 0,00         | 16,91       | 0,00         | 16,84       | 0,00         | 16,87       | 0,00         |
| <b>g<sub>N</sub></b> |             |             | 16           |             | 17,2         |             | 16,72        |             | 16,91        |             | 16,84        |             | 16,87        |

Tabla 5.1.  
Iteraciones para  
determinar el  
comportamiento  
de la cadena

|                      | <b>v(7)</b> | <b>v'(7)</b> | <b>v(8)</b> | <b>v'(8)</b> | <b>v(9)</b> | <b>v'(9)</b> | <b>v(10)</b> | <b>v'(10)</b> | <b>v(11)</b> | <b>v'(11)</b> | <b>v(12)</b> | <b>v'(12)</b> |
|----------------------|-------------|--------------|-------------|--------------|-------------|--------------|--------------|---------------|--------------|---------------|--------------|---------------|
| <b>v<sub>1</sub></b> | 18,28       | 1,43         | 18,29       | 1,43         | 18,29       | 1,43         | 18,29        | 1,43          | 18,29        | 1,43          | 18,29        | 1,43          |
| <b>v<sub>2</sub></b> | 16,85       | 0,00         | 16,86       | 0,00         | 16,86       | 0,00         | 16,86        | 0,00          | 16,86        | 0,00          | 16,86        | 0,00          |
| <b>g<sub>N</sub></b> |             | 16,85        |             | 16,86        |             | 16,86        |              | 16,86         |              | 16,86         |              | 16,86         |

Como podemos observar, a partir de  $N = 8$ , ya se repiten los resultados: podemos decir que el sistema se halla en régimen permanente (esto no es estrictamente cierto, pero las diferencias son tan pequeñas que ya no se aprecian con solo dos decimales).

No hemos indicado qué regla de normalización se ha aplicado: el resultado es coherente con estas dos:

- Normalizar siempre haciendo 0 el estado 2
- Normalizar siempre haciendo 0 el menor  $v_i$

porque ambas reglas coinciden a largo plazo y, en este caso, también a corto.

Para conocer la remuneración esperada cuando quedan 12 transiciones para terminar, basta deshacer las normalizaciones:

$$v_1^R(12) = 1,43 + 16,86 + 16,86 + \dots + 17,2 + 16 = 203,10$$

$$v_2^R(12) = 0,00 + 16,86 + 16,86 + \dots + 17,2 + 16 = 201,67$$

Si estamos en el estado 1, el valor esperado es 1,43 unidades superior al valor esperado si estamos en el estado 2.

Ello nos permite obtener la expresión de la remuneración esperada a largo plazo:

$$\lim_{N \rightarrow \infty} v(N) \approx 16,86 N + \begin{bmatrix} 1,43 \\ 0 \end{bmatrix}$$

El valor exacto de la remuneración esperada no está fijado, ya que el número de transiciones tiende a infinito y, por tanto, también la remuneración total. Lo que sí está fijado es la remuneración media por transición ( $g$ ) y la diferencia entre ambos estados iniciales:

$$g = 16,86$$

$$\lim_{N \rightarrow \infty} (v_1 - v_2) = 1,43$$



Evidentemente, dado que  $N$  es muy elevado, la diferencia de ambos valores es irrelevante frente al valor de  $Ng$ .

## 5.3 Remuneración esperada a largo plazo

### 5.3.1 Paso del corto al 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 con repetir dicha iteración hasta alcanzar el 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 media (la 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 cual finalmente vaya a parar la evolución del sistema, ya que es en ella donde se producirán la mayoría de las transiciones (suponiendo que el número de estas sea suficientemente grande); sin embargo, es obvio que dicho valor medio se verá afectado por un factor de corrección que dependerá del estado  $i$  en que se inicie el proceso. Podremos representar este hecho haciendo que, para un número de transiciones bastante elevado:

$$\text{Si } N \rightarrow \infty \quad v_i(N) \rightarrow Ng_i + w_i$$

$$v(N) \rightarrow Ng + w$$

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, obtenemos el resultado siguiente:

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

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

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



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

Ello nos permite obtener la ecuación siguiente, 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 los estados  $i$ :

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

Se trata de un sistema indeterminado, ya que tiene  $2 \cdot n$  incógnitas y solo  $n$  ecuaciones, pero las propiedades de la cadena nos permiten eliminar algunas indeterminaciones:

- El vector  $\mathbf{g}$  tendrá el mismo formato que las filas de  $\mathbf{P}^*$  ya que, para cada clase final, existe un único valor de  $g_i$ , lo cual se deduce de:

$$\mathbf{g} = \mathbf{P} \cdot \mathbf{g}$$

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

$$w_i - w_j = \lim_{N \rightarrow \infty} (v_i - v_j)$$

lo cual significa que, para cada clase final, tan solo se puede fijar una de las  $w_i$ .

Por tanto, en una cadena con  $n$  estados y  $r$  clases finales, tenemos  $r$  grados de libertad, correspondientes a fijar arbitrariamente el valor de una de las  $w_i$  para cada clase.

Si disponemos de la matriz  $\mathbf{P}^*$ , se puede calcular el vector  $\mathbf{g}$  mediante esta expresión:

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

Esta alternativa es útil si solo deseamos conocer los valores de  $g_i$  y no los de las  $w_i$ , si bien debe valorarse qué resulta más sencillo: calcular la  $\mathbf{P}^*$  o resolver el sistema.

**Ejemplo 5.2. El largo plazo**

Apliquemos esto a la matriz del ejemplo 5.1.

$$P = \begin{vmatrix} 0,2 & 0,8 \\ 0,6 & 0,4 \end{vmatrix} \quad R = \begin{vmatrix} 10 & 20 \\ 40 & -20 \end{vmatrix} \quad q = \begin{vmatrix} 18 \\ 16 \end{vmatrix}$$

La cadena es ergódica regular, de modo que la  $P^*$  tendrá las dos filas iguales:

$$P^* = [\pi_1 \quad 1 - \pi_1]$$

$$0,2 \pi_1 + 0,6 (1 - \pi_1) = \pi_1 \Rightarrow \pi_1 = 3/7 \Rightarrow P^* = [3/7 \quad 4/7]$$

Por tanto, la ganancia media por transición será:

$$g = P^* q = 118/7 \approx 16,86$$

como habíamos obtenido.

Empecemos ahora de nuevo suponiendo que no conocemos la  $P^*$ . Tendremos que plantear todo el sistema; supongamos que no conocemos el valor de las  $g_i$ , pero sí su formato, ya que hemos clasificado la cadena, es decir,  $g_1 = g_2 = g$ :

$$w_1 + g = 18 + 0,2 w_1 + 0,8 w_2$$

$$w_2 + g = 16 + 0,6 w_1 + 0,4 w_2$$

Se trata de un sistema con 2 ecuaciones y 3 incógnitas; el grado de libertad disponible debe aplicarse al valor de una de las  $w_i$ ; fijemos  $w_2 = 0$ . Resulta:

$$g = 118/7; w_1 = 10/7 \approx 1,43$$

como era de esperar. Si hubiésemos utilizado de forma distinta el grado de libertad disponible, los resultados habrían cumplido siempre que  $w_1 - w_2 = 10/7$  y el valor de  $g$  sería el mismo.

**5.3.2 Remuneración esperada a largo plazo en las cadenas regulares**

Si la cadena es regular, todas las filas de  $P^*$  son iguales, así que todos los componentes del vector  $g$  son iguales:

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

Por 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 obtener así las demás ecuaciones.

El hecho de fijar a 0 uno u otro de los valores no tiene una significación arbitraria, ya que equivale a la determinación del estado para el cual se iguala a 0 el valor  $v'(N)$  al normalizar. En efecto, aunque los valores de  $w_i$  no están determinados por el 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:

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

esto puede entenderse de la forma siguiente:  $g_i$  representa el valor medio 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 (ello es obligado si la cadena es ergódica). 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 no es 0, 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 5.3. Ganancia media a largo plazo***

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

- 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  $j - i$ .

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

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

Observemos que, en la matriz  $R$ , se han dado valores a todas las transiciones posibles, aunque su probabilidad sea 0. Desde luego es discutible si ello puede hacerse, pues en general no tiene sentido dar un valor a una posibilidad inexistente. No obstante, nada lo impide; basta con interpretar que ese sería el valor de la remuneración si dicha transición fuese posible. En cualquier caso, multiplicar el valor que sea por 0 da siempre 0, o sea, no hay ninguna dificultad; las dificultades podrían venir más bien por el lado operativo, ya que acaso multiplicar un valor "vacío" por su probabilidad (aunque sea 0) pueda ocasionar dificultades informáticas. Por ello, no es infrecuente dar un valor predeterminado (usualmente, 0) a las transiciones no posibles.

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

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

Finalmente, a partir de la ecuación matricial:

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



obtenemos el conjunto de ecuaciones siguiente, en que, dado que se trata de una cadena ergódica, sabemos que todos los valores de las componentes de  $g$  son iguales a un valor  $g$ :

$$w_1 + g = 2,3 + 0,2w_2 + 0,4w_3 + 0,3w_4 + 0,1w_5$$

$$w_2 + g = 0,5 + 0,5w_1 + 0,5w_2$$

$$w_3 + g = 1,8 + 0,2w_2 + 0,4w_3 + 0,4w_5$$

$$w_4 + g = -1,1 + 0,7w_2 + 0,3w_5$$

$$w_5 + g = -1,2 + 0,8w_3 + 0,1w_4 + 0,1w_5$$

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

$$w_1 = 0,5545$$

$$w_2 = 0$$

$$w_3 = 0,1178$$

$$w_4 = -2,5913$$

$$w_5 = -2,3801$$

$$g = 0,7772$$

Los resultados indican que la ganancia obtenida a largo plazo, tras  $N$  transiciones, es aproximadamente de  $0,7772 \cdot N$  (téngase en cuenta que, como  $N$  es muy grande, cualquier variación finita es irrelevante ante este valor: para un número 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.

### 5.3.3 Remuneración esperada en las cadenas cíclicas

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

**Ejemplo 5.4. Remuneración esperada para una cadena cíclica de  $p = 2$** 

Consideremos ahora una cadena con remuneración, como la siguiente:

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

El lector comprobará que se trata de una cadena cíclica ergódica con  $p = 2$ , cuyo vector  $\mathbf{q}$  es:

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

Planteemos el problema de obtener  $g$  y  $w$  prescindiendo del hecho de que la cadena sea cíclica; puesto que tiene una única clase final, podemos plantear las ecuaciones:

$$w_1 + g = 2 + (w_2 + w_4)/2$$

$$w_2 + g = -1 + (w_1 + w_3)/2$$

$$w_3 + g = 1 + (w_2 + w_4)/2$$

$$w_4 + g = 3 + (w_1 + w_3)/2$$

de donde obtenemos:

$$g = 5/4$$

$$w_1 = 0$$

$$w_2 = -11/4$$

$$w_3 = -1$$

$$w_4 = 5/4$$

Estos valores, al igual que en el caso regular, describen el comportamiento medio del sistema a 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 podrá comprobar, por inducción, que la ganancia total del sistema cuando faltan  $N$  etapas para terminar tiene el comportamiento siguiente:

|                | $N = 2k$       | $N = 2k+1$     |
|----------------|----------------|----------------|
| $\mathbf{v}_1$ | $(5/4)N + 1/2$ | $(5/4)N + 3/4$ |
| $\mathbf{v}_2$ | $(5/4)N - 2$   | $(5/4)N - 9/4$ |
| $\mathbf{v}_3$ | $(5/4)N - 1/2$ | $(5/4)N - 1/4$ |
| $\mathbf{v}_4$ | $(5/4)N + 2$   | $(5/4)N + 7/4$ |

Tabla 5.2.  
Ganancia total del sistema.

Esta tabla proporciona los valores exactos de la remuneración esperada a  $N$  etapas del final y se comprueba que, como ya habíamos hallado,  $g = 5/4$ .

Las componentes de  $\mathbf{v}$  se comportan de modo diferente en los períodos pares que en los períodos impares. Para compararlas, hagamos  $w_1 = 0$ , tal como hicimos al resolver el sistema de ecuaciones, y veamos el comportamiento que muestran las  $v_i$  obtenidas así y su valor medio a lo largo de un ciclo de dos iteraciones:

|                               | $N = 2k$ | $N = 2k+1$ | $w_i - w_1$ |
|-------------------------------|----------|------------|-------------|
| $\mathbf{v}_1 - \mathbf{v}_1$ | 0        | 0          | 0           |
| $\mathbf{v}_2 - \mathbf{v}_1$ | $-5/2$   | $-3$       | $-11/4$     |
| $\mathbf{v}_3 - \mathbf{v}_1$ | $-1$     | $-1$       | $-1$        |
| $\mathbf{v}_4 - \mathbf{v}_1$ | $3/2$    | $1$        | $5/4$       |

Tabla 5.3.  
Valor medio de  $v_i$  en un ciclo.

Como vemos, obtenemos los valores de las  $w_i$  obtenidas con el sistema de ecuaciones.

Si se hubiera tratado de una cadena con otro valor de  $p$ , habríamos tenido  $p$  patrones distintos del comportamiento de la  $v_i$ , y la diferencia de las  $w_i$  sería igualmente el valor medio de las  $p$  diferencias.

### 5.3.4 Remuneración esperada en las cadenas no ergódicas

En general, una cadena no ergódica con  $n$  estados tiene:

- $CF_1, CF_2, \dots, CF_m$   $m$  clases finales (ya sean cíclicas o acíclicas)
- $ecp_1, ecp_2, \dots, ecp_r$   $r$  estados transitorios, 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  $g$  que, en principio, es 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. Lamentablemente, solo podemos fijar libremente  $m$  valores de  $w_i = 0$  (uno para cada clase final), por lo que el sistema no nos permite obtener toda la información.

Para ello, hemos de calcular el valor de las  $g_{ecpi}$  para los estados transitorios, que será el promedio de los valores correspondientes a las clases finales:

$$g_{ecpi} = \sum_{CF_j} \Pr\{i \rightarrow CF_j\} \times g_{CF_j}$$

que podemos obtener fácilmente mediante la propiedad:

$$\mathbf{P} \cdot \mathbf{g} = \mathbf{g}$$

Una vez calculados dichos valores, ya podemos resolver el sistema para hallar las  $w_{eci}$ .

Debe observarse, sin embargo, que estos valores solo tienen sentido en relación con los valores fijados de las  $w_i$  para cada clase final, con lo cual su interpretación puede ser bastante discutible.

### *Ejemplo 5.5. Remuneración esperada para una cadena semirregular*

Consideremos la matriz de transición  $P$  de una cadena semirregular y las componentes activas de la matriz de remuneraciones  $R$  siguientes:

$$\mathbf{P} = \begin{vmatrix} 0,1 & 0,2 & 0,3 & 0,4 & 0 \\ 0 & 0,4 & 0,6 & 0 & 0 \\ 0 & 0,7 & 0,3 & 0 & 0 \\ 0 & 0 & 0 & 0,3 & 0,7 \\ 0 & 0 & 0 & 0,8 & 0,2 \end{vmatrix} \quad \mathbf{R} = \begin{vmatrix} 4 & 7 & -8 & 9 & - \\ - & 8 & 3 & - & - \\ - & 2 & -13 & - & - \\ - & - & - & 6 & -4 \\ - & - & - & 6 & -4 \end{vmatrix}$$

El sistema tiene dos clases finales,  $A = \{2,3\}$  y  $B = \{4,5\}$ , y un estado de paso (1). Por tanto, el vector de ganancias medias por transición  $g$  tendrá tres componentes diferentes:



- Un componente  $g_1$  asociado al estado de paso 1
- Un componente  $g_A$  asociado a la clase final  $\{2,3\}$
- Un componente  $g_B$  asociado a la clase final  $\{4,5\}$

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

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

Planteemos las ecuaciones  $g + w = q + Pw$  para la clase final A:

$$w_2 + g_A = 5 + 0,4w_2 + 0,6w_3$$

$$w_3 + g_A = -2,5 + 0,7w_2 + 0,3w_3$$

De estas ecuaciones, obtenemos:

$$g_A = 20/13$$

$$w_2 = 0; \quad w_3 = -75/13$$

Planteando las ecuaciones para la clase final B, tenemos:

$$w_4 + g_B = -1 + 0,3w_4 + 0,7w_5$$

$$w_5 + g_B = 4 + 0,8w_4 + 0,2w_5$$

$$g_B = 4/3$$

$$w_4 = 0; \quad w_5 = 10/3$$

Ahora quedan por encontrar los valores de la clase de paso,  $g_1$  y  $w_1$ . Para ello, aplicamos 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 tanto, necesitamos otra ecuación. Dicha ecuación puede encontrarse a partir de las ecuaciones matriciales que  $\mathbf{g}$  satisface:

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

$$\mathbf{P} \cdot \mathbf{g} = \mathbf{g}$$

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  $g$  y  $w$  para los estados de paso se puede deducir de las expresiones utilizadas para hallarlos. En primer lugar,  $g_1$  en la ecuación anterior es un promedio de los valores  $g_A$ ,  $g_B$  de las diferentes clases finales. Ello proviene de que, a largo plazo, lo que cuenta 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 indica 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 que ya se ha visto en las cadenas ergódicas y semiergódicas. 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  $v_i = Ng + w_i$ , sino solamente que  $v_i - v_j = w_i - w_j$ .

## 5.4 Duración del régimen transitorio

Un tipo de problema que se presenta habitualmente en las cadenas sin remuneración es saber cuántas transiciones dura el régimen transitorio. Por supuesto, se trata de una variable aleatoria de la cual solamente podemos calcular el valor esperado.



Para ello, basta con definir una remuneración que “cuente” las transiciones que no forman parte del régimen permanente, es decir, que solamente exista en los estados de paso. Veamos la técnica mediante el ejemplo siguiente:

### *Ejemplo 5.6. Duración del régimen transitorio*

Sea la cadena de orden 1 equivalente al sistema binario de orden 3 estudiado en el ejemplo 4.1, cuya matriz P es:

|       | $X_0$ | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $X_0$ | 1     |       |       |       | 0     |       |       |       |
| $X_1$ | 1/2   |       |       |       | 1/2   |       |       |       |
| $X_2$ |       | 1/2   |       |       |       | 1/2   |       |       |
| $X_3$ |       | 1/3   |       |       |       | 2/3   |       |       |
| $X_4$ |       |       | 1/2   |       |       |       | 1/2   |       |
| $X_5$ |       |       | 1/3   |       |       |       | 2/3   |       |
| $X_6$ |       |       |       | 1/3   |       |       |       | 2/3   |
| $X_7$ |       |       |       | 1/4   |       |       |       | 3/4   |

Dicha matriz define una cadena semiergódica, con una única clase final formada por el estado absorbente  $X_0$ . Deseamos saber cuántas iteraciones podemos esperar que dure el sistema antes de caer en el mismo.

Para ello, definimos una remuneración con el vector q siguiente:

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

Dos observaciones acerca de esta definición:

- No hemos definido la matriz R de remuneraciones, sino directamente el vector q. Esto es habitual cuando la remuneración solo depende del estado inicial del sistema y no del estado final. De hecho, es lo que nos interesa, a efectos de cálculo.
- La definición está pensada para contar las transiciones en régimen transitorio. En efecto, la cadena tiene una sola clase final, formada por el estado absorbente  $\{X_0\}$ , mientras que el resto son estados transitorios; por tanto, la remuneración total sumará una unidad cada vez que el sistema pase por uno de estos estados.



Dadas las características de la cadena, es evidente que se cumple:

$$P^* = [1 \ 0 \dots 0] \quad g = 0$$

y las  $w_i$  indican el número esperado de transiciones *antes* de que el sistema llegue a su estado final. Planteemos el sistema correspondiente:

$$w_0 + g = 0 + w_0$$

$$w_1 + g = 1 + (w_0 + w_4)/2$$

$$w_2 + g = 1 + (w_1 + w_5)/2$$

$$w_3 + g = 1 + (w_1 + 2w_5)/3$$

$$w_4 + g = 1 + (w_2 + w_6)/2$$

$$w_5 + g = 1 + (w_2 + 2w_6)/3$$

$$w_6 + g = 1 + (w_3 + 2w_7)/3$$

$$w_7 + g = 1 + (w_3 + 3w_7)/4$$

en que resulta natural hacer  $w_0 = 0$ , con lo cual resulta:

$$w_1 = 12,50; w_2 = 19,21; w_3 = 21,12; w_4 = 23,00; w_5 = 23,93;$$

$$w_6 = 24,79; w_7 = 25,12$$

que es el número esperado de transiciones del sistema antes de caer en el estado absorbente, según el estado en que se halle.

### *Ejemplo 5.7. Duración del régimen transitorio en una cadena no ergódica*

Sea la cadena de orden 1 cuya matriz P es:

$$P = \begin{array}{c|cccccc} & 0,7 & 0,2 & 0,1 & & & \\ & & 0,5 & 0,3 & 0,1 & 0,1 & \\ & & & \boxed{\begin{array}{cc} 0,5 & 0,5 \\ 0,5 & 0,5 \end{array}} & & & \\ & & & & & & \boxed{\begin{array}{c} 1 \\ 1 \end{array}} \end{array}$$

Dicha matriz define una cadena mixta con dos clases finales, la  $\{3,4,5\}$  acíclica y otra con  $p = 2$ , la  $\{6,7\}$ . Igual que en el ejemplo anterior, deseamos saber cuántas iteraciones podemos esperar que dure el sistema antes de caer en alguna de las dos. Para ello, definimos una remuneración con el vector  $q$  siguiente:



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

Dadas las características de la cadena, está claro que se cumple  $g_{345} = g_{67} = 0$ , pero la  $P^*$  no es tan inmediata como antes. Sin embargo, si solo nos interesa la duración del transitorio, podemos ahorrarnos su cálculo.

El sistema de ecuaciones correspondiente será:

$$w_1 + g = 1 + 0,7 w_1 + 0,2 w_2 + 0,1 w_3$$

$$w_2 + g = 1 + 0,5 w_2 + 0,3 w_3 + 0,1 w_5 + 0,1 w_6$$

$$w_3 + g = 0 + 0,5 w_4 + 0,5 w_5$$

$$w_4 + g = 0 + 0,5 w_3 + 0,5 w_5$$

$$w_5 + g = 0 + 0,5 w_3 + 0,5 w_4$$

$$w_6 + g = 0 + w_7$$

$$w_7 + g = 0 + w_6$$

Disponemos de dos grados de libertad que nos permiten fijar dos de la  $w_i$ , una para cada clase final; hagamos, por ejemplo,  $w_3 = w_6 = 0$ , con lo cual resulta:

$$w_1 = 4,67; w_2 = 2; w_3 = w_4 = w_5 = w_6 = w_7 = 0$$

que es el número esperado de transiciones del sistema antes de caer en el estado absorbente.

Si deseamos conocer la probabilidad de ir a parar a una u otra clase final, hemos de calcular la  $P^*$ . Para las clases finales, resulta obvio que:

$$P_{345}^* = [1/3 \quad 1/3 \quad 1/3] \quad ; P_{67}^* = [1/2 \quad 1/2]$$

y, para los estados transitorios (ya sin las columnas 1 y 2):

$$P_{12}^* = \begin{vmatrix} 0,2889 & 0,2889 & 0,2889 \\ 0,2667 & 0,2667 & 0,2667 \end{vmatrix} \begin{vmatrix} 0,0667 & 0,0667 \\ 0,1000 & 0,1000 \end{vmatrix}$$



de modo que las probabilidades solicitadas son:

$$P\{1;345\} = 0,8667; P\{1;67\} = 0,1333; P\{2;345\} = 0,8000; P\{1;67\} = 0,2000$$

Para finalizar el capítulo, indicamos que las cadenas de Markov con remuneración que se han presentado aquí son, de hecho, un caso particular de programación dinámica aleatoria. En el volumen de esta misma colección dedicado a la programación dinámica, se desarrolla con más detalle esta metodología, en un contexto más general. El lector puede comparar ambos métodos y comprobará que, en algunos casos, aunque la diferencia teórica es importante, desde el punto de vista de los cálculos, la distinción no va más allá de un cambio de notación.



→6



# Cadenas de Markov con remuneración y decisión

## 6.1 Introducción

Se dice que un proceso markoviano tiene decisión si en cada transición se puede fijar una variable (denominada *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  $P$  como la de remuneraciones  $R$  tienen tres índices o dimensiones: el estado anterior, el posterior y la decisión. Por ello, denotamos sus elementos  $p_{iju}$  y  $r_{iju}$ , donde  $u$  representa la variable de decisión. Así, los valores 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 denomina *política* la regla que fija para cada estado de la cadena el valor de la decisión que hay que adoptar. Podrá expresarse entonces de forma vectorial:

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

donde  $u_i$  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 capítulo anterior. Es importante tener en cuenta que, mientras no se ha fijado una política, 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 tendrá 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 la mejor posible, lo cual puede significar la máxima o la mínima, según la naturaleza de la remuneración. De este modo, se puede hablar de *política óptima*, que es la que permite obtener la mejor remuneración esperada. Téngase en cuenta que, por razón de la naturaleza aleatoria del sistema, no se puede optimizar la remuneración *obtenida realmente* (que es una magnitud aleatoria), sino tan solo 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 de *política óptima a largo plazo* (es decir, con un horizonte ilimitado o con un horizonte finito pero desconocido).

Obsérvese que, en la política óptima a corto plazo, la decisión no solo depende del estado inicial, sino también del número de transición, lo cual no ocurre con la política a largo plazo. La razón de ello es evidente: si no sabemos cuántas transiciones nos faltan para acabar, lo interesante será obtener una buena remuneración, pero sin que ello perjudique las remuneraciones futuras ya que, en principio, el horizonte es infinito; en cambio, a corto plazo puede ser interesante “sacar todo el jugo” del sistema, aunque ello signifique que no vaya durar mucho; el ejemplo más claro es el de una máquina que pronto vamos a sustituir y ya no vale la pena gastar mucho dinero en su mantenimiento.

No debe confundirse esta característica de la política a corto plazo con el hecho de que el sistema sea homogéneo o no en el tiempo. Un sistema es homogéneo en el tiempo si su comportamiento no depende de la variable cronológica, lo cual no impide que lo tratemos de forma diferente según sea el valor de dicha variable: una cosa es cómo es el sistema y otra, cómo lo gestionamos.

## 6.2 Procesos markovianos de decisión

Algunos procesos polietápicos de decisión pueden modelizarse como cadenas de Markov de orden 1 con remuneración y decisión. Para que ello sea posible, han de tener las características siguientes:



- El sistema puede evolucionar dentro de un número finito de estados  $n$ .
- El sistema ha de ser homogéneo en el tiempo.
- La evolución del estado en el futuro dependerá exclusivamente:
  - a) Del estado  $i$  en que se encuentre en el momento presente (eso es, el sistema cumple la propiedad markoviana).
  - b) 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 puede definirse como una secuencia de evolución (E) según un proceso markoviano y de decisión (D) que condiciona la evolución a la etapa siguiente. Estos sistemas pueden caracterizarse como sistemas E / D. Esta característica nos facilitará la modelización de sistemas complejos, como veremos más adelante.

Los datos de partida para resolver el sistema son:

- La matriz de probabilidades de transición para cada una de las políticas  $P^{(p)}$ . 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  $R^{(p)}$ , 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  $i$ .

Una vez definido el problema en estos términos, es importante destacar que, solo cuando hayamos establecido una política determinada, tendremos un proceso markoviano con remuneración como se ha estudiado en el capítulo anterior.

Dado que la remuneración obtenida depende de la política aplicada, podemos definir la *política óptima* como aquella que nos permite obtener la mejor remuneración esperada posible (máxima o mínima) y la *remuneración óptima* correspondiente a esta política óptima. Denotaremos ambos conceptos  $u^*(N)$  y  $v^*(N)$ , en que el asterisco indica que nos referimos al óptimo. Dada la estrecha relación existente entre ambas, hallar una implica hallar también la otra, para lo cual disponemos de dos procedimientos:

- La *iteración en el espacio de los estados*, que consiste en encontrar la política que nos da la  $v(N)$  óptima, esto es,  $v^*(N)$ , a partir de  $v^*(N-1)$  y la  $u^*(N)$  correspondiente. Este método sirve para obtener la política óptima para un número finito de etapas (problema de horizonte finito) y, si se prolonga lo suficiente, se observa que, después de un número determinado de etapas, el sistema converge a una determinada política estable, que ya no depende de  $N$ ; es la política óptima a largo plazo (problema de horizonte infinito).



- La *iteración del espacio de las políticas*, por el contrario, solo permite hallar la política y la remuneración óptima a largo plazo, y consiste en encontrar los vectores  $g$  y  $w$  para diferentes políticas posibles del sistema, y obtener de su análisis una política cada vez mejor a largo plazo hasta llegar a la óptima. Por su propia naturaleza, se trata de una metodología para resolver el problema de horizonte infinito.

A continuación, se describen cada una de estas técnicas utilizando diversos ejemplos prototipo.

### 6.3 Iteración en el espacio de los estados

Como se ha expuesto anteriormente, se trata de encontrar  $v^*(N)$ , una vez obtenida  $v^*(N-1)$ . Para ello, se utilizará la ecuación de recurrencia definida para las cadenas de Markov con remuneración, teniendo en cuenta ahora que, para cada uno de los estados de 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 obtenidas las decisiones óptimas para todos los estados, quedará definida la política óptima para esa etapa:

$$v_i^*(N) = \text{OPT}_u \{v_{iu}(N)\}$$

$$u_i^*(N) = \{u_j \mid v_j^*(N)\}$$

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

Como ya vimos al tratar de las cadenas con remuneración, una técnica interesante es la *normalización*, que consiste en restar, a los valores obtenidos de la  $v^*(N)$ , un valor constante que hay que elegir en cada iteración ( $g_N$ ):

$$v'^*(N) = v^*(N) - g_N$$

y utilizar estos valores normalizados en la siguiente iteración. Usualmente, se toma como  $g_N$  el mínimo de los  $v_i^*(N)$ , lo cual permite que el vector  $v'^*(N)$  tenga un componente a 0 y el resto, positivos. Esto, además de facilitar los



cálculos, permite comparar fácilmente entre los valores correspondientes a las diferentes iteraciones y evidencia la aparición del régimen permanente.

Dado que, una vez superado el período transitorio, el sistema converge hacia un régimen permanente, es claro que el procedimiento de iteración en el espacio de los estados nos permite también hallar, si lo iteramos suficientemente, la política óptima a largo plazo, y la normalización es una herramienta muy útil para ello. En particular, sabemos que se ha alcanzado el régimen permanente cuando se cumple:

$$\mathbf{u}^*(N) = \mathbf{u}^*(N-1)$$

$$\mathbf{v}'^*(N) = \mathbf{v}'^*(N-1)$$

$$\mathbf{g}_N = \mathbf{g}_{N-1}$$

La primera condición suele hallarse relativamente pronto, pero las condiciones segunda y tercera, que se cumplen simultáneamente, suelen tardar bastante (en realidad, solo se cumplen asintóticamente), por lo cual el método no suele resultar muy práctico para hallar la política a largo plazo y hay que recurrir a la iteración del espacio de las políticas, teniendo en cuenta que

$$\mathbf{u}^*(N) \rightarrow \mathbf{u}$$

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

$$\mathbf{g}_N \rightarrow \mathbf{g}$$

### *Ejemplo 6.1. Iteración en el espacio de los estados*

Sea un sistema que puede hallarse en tres estados, A, B o C, y disponemos de tres acciones (1, 2 y 3) que permiten variar las probabilidades del nuevo estado y las remuneraciones obtenidas, que supondremos que son ganancias.

Los números 1, 2 y 3 con que identificamos cada política son solo índices que las distinguen, pero no presuponen nada con respecto a su naturaleza: la acción indicada como 1 en el estado A puede ser o no la que se indica con el mismo número en los estados B y C; por ello, no tiene por qué haber la misma cantidad de opciones en cada estado, como ocurre en el este ejemplo.

Estas 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 6.1.



Tabla 6.1.  
Probabilidades,  
remuneraciones y  
ganancias  
esperadas

| Estado | Decisión | $P_{iju}$ |     |     | $R_{iju}$ |    |    | $q_i^{(p)}$ |
|--------|----------|-----------|-----|-----|-----------|----|----|-------------|
|        |          | A         | B   | C   | A         | B  | C  |             |
| 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           |

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

- Del estado del sistema A, B o C en el momento de la decisión: el sistema es de tipo markoviano y de orden 1.
- De la política que hayamos seguido: deseamos obtener la política óptima, que indicará la decisión a tomar según el estado en que se encuentre el sistema. Así, por ejemplo, la política  $p = \{1, 2, 3\}$  significa que debe hacerse la acción 1 si el sistema se halla en el estado A, la 2 si está en el B y la 3 si está en el C. Entonces, tenemos  $3 \times 2 \times 3 = 18$  políticas posibles, por lo cual no resulta operativo explorarlas todas.

Hemos de distinguir entre la política óptima a corto plazo (o transitoria), que depende de N (no es lo mismo si faltan 3 transiciones para terminar la gestión que si faltan 5, por ejemplo), y la política óptima a largo plazo, que ya no depende de N, dado que este número es tan grande que el sistema ya habrá alcanzado el régimen permanente antes de alcanzar el final.

En la tabla 6.1, se ha calculado también el vector  $\mathbf{q}$ , que, como en el caso sin decisión, representa la remuneración obtenida en la primera transición, solo que ahora, además del estado en que se halle el sistema, dependerá también de la decisión.

Dado que la ecuación de recurrencia exige conocer los valores de N-1 para calcular la política en N, hemos de definir un vector inicial  $\mathbf{v}^*(0)$  con el cual iniciar la iteración. Dicho vector representa la remuneración obtenida después de la última transición (eso es, una vez terminada la gestión), y supondremos que es 0.

En primer lugar, resolvemos el problema sin normalización:



|          | <b>v(0)</b> | u=1 | u=2 | u=3 | <b>v*(1)</b> | <b>u*(1)</b> | u=1  | u=2  | u=3  | <b>v*(2)</b> | <b>u*(2)</b> |
|----------|-------------|-----|-----|-----|--------------|--------------|------|------|------|--------------|--------------|
| <b>A</b> | <b>0</b>    | 8   | 3   | 4   | <b>8</b>     | <b>1</b>     | 17,3 | 16,4 | 12,2 | <b>17,3</b>  | <b>1</b>     |
| <b>B</b> | <b>0</b>    | 16  | 14  |     | <b>16</b>    | <b>1</b>     | 23,5 | 28,3 |      | <b>28,3</b>  | <b>2</b>     |
| <b>C</b> | <b>0</b>    | 7   | 4   | 4   | <b>7</b>     | <b>1</b>     | 16,9 | 18,3 | 12,6 | <b>18,3</b>  | <b>2</b>     |

Tabla 6.2.  
Solución del  
ejemplo 6.1 sin  
normalización

|          | u=1  | u=2  | u=3  | <b>v*(3)</b> | <b>u*(3)</b> | u=1  | u=2 | u=3  | <b>v*(4)</b> | <b>u*(4)</b> |
|----------|------|------|------|--------------|--------------|------|-----|------|--------------|--------------|
| <b>A</b> | 27,8 | 28,2 | 23   | <b>28,2</b>  | <b>2</b>     | 39,2 | 40  | 34,6 | <b>40</b>    | <b>2</b>     |
| <b>B</b> | 33,8 | 40,2 |      | <b>40,2</b>  | <b>2</b>     | 45,2 | 52  |      | <b>52</b>    | <b>2</b>     |
| <b>C</b> | 28,1 | 30,2 | 22,6 | <b>30,2</b>  | <b>2</b>     | 39,8 | 42  | 33,8 | <b>42</b>    | <b>2</b>     |

|          | u=1  | u=2  | u=3  | <b>v*(5)</b> | <b>u*(5)</b> | u=1  | u=2  | u=3  | <b>v*(6)</b> | <b>u*(6)</b> |
|----------|------|------|------|--------------|--------------|------|------|------|--------------|--------------|
| <b>A</b> | 51   | 51,8 | 46,4 | <b>51,8</b>  | <b>2</b>     | 62,8 | 63,6 | 58,2 | <b>63,6</b>  | <b>2</b>     |
| <b>B</b> | 57   | 63,8 |      | <b>63,8</b>  | <b>2</b>     | 68,8 | 75,6 |      | <b>75,6</b>  | <b>2</b>     |
| <b>C</b> | 51,6 | 53,8 | 45,6 | <b>53,8</b>  | <b>2</b>     | 63,4 | 65,6 | 57,4 | <b>65,6</b>  | <b>2</b>     |

Dado que se trata de una función a maximizar, las decisiones óptimas se han obtenido como:

$$v_i^* = \max \{ v_i(u=1), v_i(u=2), v_i(u=3) \}$$

Se puede observar que la política a corto plazo evoluciona desde la {1,1,1} para  $N = 1$  (es decir, en la última transición antes de acabar) hasta la {2,2,2}; en este caso, la transición es muy rápida, puesto que se cumple ya a partir de la tercera. Como era de esperar, los valores de  $\mathbf{v}^*(N)$  van creciendo, pero no es fácil observar el crecimiento medio por transición, ni tampoco si se ha alcanzado ya la estabilidad del régimen permanente.

Veamos ahora la misma resolución aplicando la normalización con la técnica habitual:

$$g(N) = \min \{ v_A^*, v_B^*, v_C^* \}$$

$$v_i^{*'} = v_i^* - g(N)$$

|             | <b>v(0)</b> | u=1 | u=2 | u=3 | <b>v*(1)</b> | <b>v'*(1)</b> | <b>u*(1)</b> | u=1  | u=2  | u=3 | <b>v*(2)</b> | <b>v'*(2)</b> | <b>u*(2)</b> |
|-------------|-------------|-----|-----|-----|--------------|---------------|--------------|------|------|-----|--------------|---------------|--------------|
| <b>A</b>    | <b>0</b>    | 8   | 3   | 4   | <b>8</b>     | <b>1</b>      | <b>1</b>     | 10,3 | 9,4  | 5,2 | <b>10,3</b>  | <b>0</b>      | <b>1</b>     |
| <b>B</b>    | <b>0</b>    | 16  | 14  |     | <b>16</b>    | <b>9</b>      | <b>1</b>     | 16,5 | 21,3 |     | <b>21,3</b>  | <b>11</b>     | <b>2</b>     |
| <b>C</b>    | <b>0</b>    | 7   | 4   | 4   | <b>7</b>     | <b>0</b>      | <b>1</b>     | 9,9  | 11,3 | 5,6 | <b>11,3</b>  | <b>1</b>      | <b>2</b>     |
| <b>g(N)</b> |             |     |     |     | <b>7</b>     |               |              |      |      |     | <b>10,3</b>  |               |              |

Tabla 6.3.  
Solución del  
ejemplo 6.1 con  
normalización



Continuación  
de la tabla 6.3.  
Solución del  
ejemplo 6.1 con  
normalización

|             | u=1         | u=2  | u=3 | $v^*(3)$    | $v'^*(3)$ | $u^*(3)$ | u=1          | u=2  | u=3 | $v^*(4)$    | $v'^*(4)$ | $u^*(4)$ |
|-------------|-------------|------|-----|-------------|-----------|----------|--------------|------|-----|-------------|-----------|----------|
| <b>A</b>    | 10,5        | 10,9 | 5,7 | <b>10,9</b> | <b>0</b>  | <b>2</b> | 11           | 11,8 | 6,4 | <b>11,8</b> | <b>0</b>  | <b>2</b> |
| <b>B</b>    | 16,5        | 22,9 |     | <b>22,9</b> | <b>12</b> | <b>2</b> | 17           | 23,8 |     | <b>23,8</b> | <b>12</b> | <b>2</b> |
| <b>C</b>    | 10,8        | 12,9 | 5,3 | <b>12,9</b> | <b>2</b>  | <b>2</b> | 11,6         | 13,8 | 5,6 | <b>13,8</b> | <b>2</b>  | <b>2</b> |
| <b>g(N)</b> | <b>10,9</b> |      |     |             |           |          | <b>11,80</b> |      |     |             |           |          |

|             | u=1          | u=2  | u=3 | $v^*(5)$    | $v'^*(5)$ | $u^*(5)$ | u=1          | u=2  | u=3 | $v^*(6)$    | $v'^*(6)$ | $u^*(6)$ |
|-------------|--------------|------|-----|-------------|-----------|----------|--------------|------|-----|-------------|-----------|----------|
| <b>A</b>    | 11           | 11,8 | 6,4 | <b>11,8</b> | <b>0</b>  | <b>2</b> | 11           | 11,8 | 6,4 | <b>11,8</b> | <b>0</b>  | <b>2</b> |
| <b>B</b>    | 17           | 23,8 |     | <b>23,8</b> | <b>12</b> | <b>2</b> | 17           | 23,8 |     | <b>23,8</b> | <b>12</b> | <b>2</b> |
| <b>C</b>    | 11,6         | 13,8 | 5,6 | <b>13,8</b> | <b>2</b>  | <b>2</b> | 11,6         | 13,8 | 5,6 | <b>13,8</b> | <b>2</b>  | <b>2</b> |
| <b>g(N)</b> | <b>11,80</b> |      |     |             |           |          | <b>11,80</b> |      |     |             |           |          |

Como puede observarse, los valores tanto de la  $g(N)$  como del vector normalizado  $v'^*(N)$  se repiten exactamente desde la cuarta iteración, lo cual nos garantiza que ya se ha alcanzado el régimen permanente. Para obtener el valor real de la  $v^*(N)$ , basta con sumar a la  $v'^*(N)$  correspondiente todos los valores normalizados hasta entonces:

$$\mathbf{v}^{*R}(6) = \mathbf{v}'^*(6) + g(6) + g(5) + \dots + g(1)$$

$$\mathbf{v}^{*R}(6) = \begin{vmatrix} 0 \\ 12 \\ 2 \end{vmatrix} + 3 * 11,80 + 10,9 + 10,3 + 7 = \begin{vmatrix} 63,6 \\ 75,6 \\ 65,6 \end{vmatrix}$$

de donde se deduce claramente que la política  $\{2,2,2\}$  es óptima a largo plazo, con una ganancia media por transición de 11,80, y las  $w_j$  son 0, 12 y 2, respectivamente.

No siempre la convergencia es tan rápida como en este caso. Normalmente, la política se repite relativamente pronto, pero las  $v'^*$  tardan bastante más (en realidad, puede estar "bailando" siempre el  $n$ ésimo decimal), por lo cual este método no suele ser muy eficiente para calcular el régimen permanente.

Se trata, entonces, de llevar a cabo una exploración en el espacio de las políticas, cuya operativa se detalla en el apartado siguiente.



## 6.4 Iteración en el espacio de las políticas

La *iteración del espacio de las políticas* es un procedimiento que permite hallar la política óptima para el sistema a largo plazo. Para cada uno de los  $n$  estados posibles, indica la acción que hay que realizar si el sistema está lo bastante alejado de su final para considerar que se halla en régimen permanente.

El procedimiento, esencialmente, consiste en:

1. Fijar una política inicial cualquiera.
2. Determinar  $\mathbf{g}$  y  $\mathbf{w}$  para la política actual.
3. Comprobar si los valores hallados cumplen la condición de óptimo.
4. Si cumplen la condición, fin del algoritmo.
5. Si no la cumplen, sustituir la política actual por otra mejor y volver al punto 2.

Examinaremos cada paso con detalle.

### *Paso 1: Fijar una política inicial*

En principio, no existe ninguna restricción para fijar esta política. Sin embargo, cuanto mejor sea, menos iteraciones serán necesarias para hallar el óptimo. Por tanto, podemos fijar algunas reglas prácticas:

- Si existe una política que conduzca a una cadena ergódica, ninguna política que conduzca a una cadena no ergódica podrá ser óptima (recordemos que la clasificación de la cadena depende de la política). Ello se debe a que las  $\mathbf{g}$  asociadas a las distintas clases finales son distintas, en general, con lo cual toda política que evite la posibilidad de ir a parar a una clase con una  $\mathbf{g}$  peor será mejor que la que permita dicha posibilidad.
- Si la política óptima a corto plazo se ha calculado mediante la iteración en el espacio de los estados, la mejor opción es tomar la política para la mayor  $N$  calculada; de esta manera, se podrá confirmar si los resultados obtenidos mediante la iteración en los estados son ya óptimos en régimen permanente.
- A falta de mejor información, una buena elección es escoger para cada estado aquella acción cuya  $q_i$  sea mejor.

### *Paso 2: Hallar los vectores $\mathbf{g}$ y $\mathbf{w}$*

Dado que fijar una política equivale a definir una cadena con remuneración, se trata de resolver para la política  $p$  el sistema ya conocido:



$$w(p) + g(p) = q(p) + P(p) w(p)$$

Recordemos que, si la cadena resultante es ergódica, tendremos que fijar una de las  $w_i = 0$ . Si la cadena no es ergódica, bastará con hallar las  $g$  de cada clase final  $y$ , simplemente cambiando las decisiones para asegurar que el sistema siempre irá a parar a la clase con mejor  $g$ , ya se obtiene una política mejor. En este caso, puede resultar más cómodo para hallar los valores de  $g$  obtener la matriz  $P^*$  que resolver el sistema.

*Paso 3: Comprobar si cumple la condición de óptimo*

Si la política es óptima, para cada estado se debe cumplir:

$$w_i = \text{OPT}_u \{ q_i(u) + \sum_j P_{ij}(u) - g_i \}$$

Se trata de comprobar si, en cada estado, la acción definida por la política cumple o no dicha condición.

*Paso 4: Definición de la nueva política o fin del algoritmo*

Si en el paso anterior se ha cumplido la condición en todos los estados, la política es óptima y se termina el algoritmo.

En caso contrario, habrá al menos un estado en que exista por lo menos una decisión ( $u$ ) para la cual el valor  $q_i(u) + \sum_j P_{ij}(u) - g_i$  será mejor (mayor si maximizamos, menor si minimizamos) que el correspondiente a la política ensayada. Cambiando esta decisión, se garantiza que la política definida por este cambio ya será mejor que la anterior. Evidentemente, es aconsejable hacer en cada iteración todos los cambios necesarios en todos los estados, eligiendo sistemáticamente aquella  $u$  cuyo valor sea mejor. De esta manera, se obtiene una nueva política que se debe analizar volviendo al paso 2 del algoritmo.

Obsérvese que el hecho de que en una iteración se haya sustituido una decisión concreta por otra no garantiza que dicha decisión no vuelva a aparecer más adelante en otra iteración, ya que su optimalidad no depende solo de ella misma, sino del conjunto de la política.

### ***Ejemplo 6.2. Iteración en el espacio de las políticas***

Resolvamos el problema del ejemplo 6.1 a largo plazo. Como se ha indicado más arriba, lo mejor para fijar la política inicial sería tomar la  $\{2, 2, 2\}$ , que sabemos que es óptima a corto plazo, con  $N = 4$  (de hecho, ya sabemos que es óptima a largo plazo), pero ello no nos permitiría ver el algoritmo completo; prescindamos, pues, de lo que ya sabemos y partamos de la mínima información: supongamos que tan solo conocemos la tabla con los datos del



enunciado. En estas condiciones, lo mejor es fijar la política  $\{1,1,1\}$ , que es la que garantiza una  $q_i$  mayor en cada estado; la cadena resultante es regular positiva, por lo cual el sistema será:

$$w_A + g = 8 + 0,5 w_A + 0,2 w_B + 0,3 w_C$$

$$w_A + g = 16 + 0,5 w_A + 0,5 w_C$$

$$w_A + g = 7 + 0,2 w_A + 0,3 w_B + 0,5 w_C$$

cuya solución es:  $g = 9,191$ ;  $w_A = 0,426$ ;  $w_B = 7,021$ ;  $w_C = 0$ .

Comprobemos si es óptima:

$$\begin{aligned} w_A: \quad \max \{ & 8 + 0,5 w_A + 0,2 w_B + 0,3 w_C; \\ & 3 + 0,1 w_A + 0,7 w_B + 0,2 w_C; \\ & 4 + 0,3 w_A + 0,1 w_B + 0,6 w_C \} - g = \\ & = \max \{0,426; -0,13; -4,36\} = 0,426 = w_A \Rightarrow \text{OK} \end{aligned}$$

$$\begin{aligned} w_B: \quad \max \{ & 16 + 0,5 w_A + 0,5 w_C; \\ & 14 + 0,1 w_A + 0,8 w_B + 0,1 w_C \} - g = \\ & = \max \{7,021; 10,47\} = 10,47 > w_B \Rightarrow u = 2 \end{aligned}$$

$$\begin{aligned} w_C: \quad \max \{ & 7 + 0,2 w_A + 0,3 w_B + 0,5 w_C; \\ & 4 + 0,1 w_A + 0,8 w_B + 0,1 w_C; \\ & 4 + 0,7 w_A + 0,1 w_B + 0,2 w_C \} - g = \\ & = \max \{0; 0,468; -4,19\} = 0,466 > w_C \Rightarrow u = 2 \end{aligned}$$

Por tanto, la política  $\{1, 2, 2\}$  es mejor a largo plazo. Vamos a comprobarlo:

$$\begin{aligned} w_A + g &= 8 + 0,5 w_A + 0,2 w_B + 0,3 w_C & g &= 11,667; \\ w_A + g &= 14 + 0,1 w_A + 0,8 w_B + 0,1 w_C & \Rightarrow & w_A = -3,333; \\ w_A + g &= 4 + 0,1 w_A + 0,8 w_B + 0,1 w_C & & w_B = 10; w_C = 0 \end{aligned}$$

$$\begin{aligned} w_A: \quad \max \{ & 8 + 0,5 w_A + 0,2 w_B + 0,3 w_C; 3 + 0,1 w_A + 0,7 w_B + 0,2 w_C; \\ & 4 + 0,3 w_A + 0,1 w_B + 0,6 w_C \} - g = \max \{-3,333; -2; -7,667\} = \\ & = -2 > w_A \Rightarrow u = 2 \end{aligned}$$



$$w_B: \quad \max \{16 + 0,5 w_A + 0,5 w_C; 14 + 0,1 w_A + 0,8 w_B + 0,1 w_C\} - g = \\ = \max \{2,667; 10\} = 10 = w_B \Rightarrow \text{OK}$$

$$w_C: \quad \max \{7 + 0,2 w_A + 0,3 w_B + 0,5 w_C; 4 + 0,1 w_A + 0,8 w_B + 0,1 w_C; \\ 4 + 0,7 w_A + 0,1 w_B + 0,2 w_C\} - g = \max \{-2,333; 0; -9\} = 0 = w_C \Rightarrow \text{OK}$$

Debemos, pues, ensayar la política {2,2,2}: la resolución del sistema

$$\begin{aligned} w_A + g &= 3 + 0,1 w_A + 0,7 w_B + 0,7 w_C & g &= 11,8; \\ w_A + g &= 14 + 0,1 w_A + 0,8 w_B + 0,1 w_C & \Rightarrow w_A &= -2; \\ w_A + g &= 4 + 0,1 w_A + 0,8 w_B + 0,1 w_C & w_B &= 10; w_C = 0 \end{aligned}$$

El lector podrá comprobar que es óptima a largo plazo, como ya sabíamos.

Este ejemplo permite hacer algunas observaciones interesantes. Como era de esperar, el valor hallado para  $g$  es el mismo que ya habíamos obtenido en la iteración en el espacio de los estados; pero, aparentemente, no ocurre lo mismo con las  $w_i$ . Recordemos que estos valores no están determinados, pero sí sus diferencias; por tanto, la solución {0; 12; 2} obtenida antes es igual que la obtenida ahora {-2; 10; 0}; simplemente, hemos cambiado el valor de referencia para la normalización.

Más interesante es la evolución de las soluciones ensayadas en este procedimiento. Hemos empezado con la política {1; 1; 1}, que es la que garantiza una mayor ganancia en la primera transición (eso es lo que significa el vector  $q$ ). Es decir, se trata de una política "miope", que solo mira un plazo inmediato, y por ello es óptima a corto plazo para  $N = 1$ ; esta suele ser también una buena política a largo plazo, pero no siempre, y esto es lo que ocurre en este caso. Si observamos la tabla de datos, veremos que las mayores ganancias se obtienen cuando el sistema se halla en el estado B; por tanto, a largo plazo conviene maximizar la probabilidad de que la transición nos conduzca hacia él, que es lo que hace la política {2; 2; 2}. En cambio, la {1; 1; 1} tiene unas probabilidades bajas de llevarnos a B: en un lenguaje coloquial, diríamos que es "pan para hoy, hambre para mañana". Lo confirma el cálculo de las  $P^*$  para cada una de las tres políticas ensayadas:

$$P^* \{1,1,1\} = [0,37; 0,20; 0,43];$$

$$P^* \{1,2,2\} = [0,17; 0,70; 0,13];$$

$$P^* \{2,2,2\} = [0,10; 0,79; 0,11];$$



### *Ejemplo 6.3. El dilema del profesor*

Un profesor de Métodos Cuantitativos debe decidir, dos veces por trimestre, si pone un examen fácil o difícil. Lo ideal sería poner un examen difícil y que hubiera muchos aprobados: ello indicaría que los alumnos han estudiado y dominan la materia impartida. Por el contrario, si hay pocos aprobados y el examen ha sido fácil, es un indicador de que no han estudiado lo suficiente. El profesor considera de valor intermedio las otras dos situaciones posibles: examen difícil y pocos aprobados, o muchos aprobados y examen fácil.

Ahora bien, los alumnos estudian más o menos según si en el examen anterior ha habido pocos o muchos aprobados: si lo han aprobado muchos, estudian poco y, si ha habido muchos suspensos, estudian mucho. Ello afecta la probabilidad de aprobar: si, después de un examen con muchos aprobados, el profesor pone un examen fácil, solo tres de cada cinco veces habrá muchos aprobados. Si, tras una situación con pocos aprobados, pone un examen difícil, habrá muchos o pocos aprobados, con la misma 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 existe una probabilidad del 90 % de que haya muchos aprobados.

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

Aunque este ejemplo está menos estructurado que el ejemplo 6.2 anterior, resulta bastante 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 cadena de Markov de orden 1 (o varias, puesto que podrían variar en función de la política).
- Antes de la evaluación siguiente, el profesor toma la decisión de que el examen sea fácil (F) o difícil (D). Dado que el profesor ha de establecer cuál será su comportamiento en función de que haya habido muchos o pocos aprobados, podemos definir la política, en este caso, como:

{comportamiento si M, comportamiento si P}



Y serán 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 tabla 6.4: 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. El enunciado no dice cómo hemos de valorar las remuneraciones, así que somos libres para fijar su escala numérica: dado el contexto del problema, tomaremos una entre 0 y 10, habitual en la enseñanza; si hubiésemos elegido otra, podríamos pasar de una a otra mediante una sencilla transformación lineal.

Tabla 6.4.  
Matrices de  
probabilidad,  
remuneración y  
ganancia esperada

|        |          | Matriz P |     | Matriz R |   | q   |
|--------|----------|----------|-----|----------|---|-----|
| Estado | Decisión | M        | P   | M        | P |     |
| M      | F        | 0,6      | 0,4 | 5        | 0 | 3   |
|        | D        | 0        | 1   | 10       | 5 | 5   |
| P      | F        | 0,9      | 0,1 | 5        | 0 | 4,5 |
|        | D        | 0,5      | 0,5 | 10       | 5 | 7,5 |

Aunque lo que interesa en este problema es el largo plazo, empezamos estudiando el régimen transitorio mediante la iteración en el espacio de los estados para presentar el caso completo:

Tabla 6.5.  
Solución del ejem-  
plo 6.3 iterando en  
el espacio de los  
estados

|   | v(0) | F   | D   | v*(1) | v'*(1) | u*(1) | F    | D    | v*(2) | v'*(2) | u*(2) |
|---|------|-----|-----|-------|--------|-------|------|------|-------|--------|-------|
| M | 0    | 3   | 5   | 5     | 0      | D     | 4    | 7,5  | 7,5   | 0      | D     |
| P | 0    | 4,5 | 7,5 | 7,5   | 2,5    | D     | 4,75 | 8,75 | 8,75  | 1,25   | D     |
| g |      | 5   |     |       |        |       | 7,5  |      |       |        |       |

|   | F     | D     | v*(3) | v'*(3) | u*(3) | F     | D     | v*(4) | v'*(4) | u*(4) |
|---|-------|-------|-------|--------|-------|-------|-------|-------|--------|-------|
| M | 3,5   | 6,25  | 6,25  | 0      | D     | 3,75  | 6,875 | 6,875 | 0      | D     |
| P | 4,625 | 8,125 | 8,125 | 1,875  | D     | 4,688 | 8,438 | 8,438 | 1,563  | D     |
| g | 6,25  |       |       |        |       | 6,875 |       |       |        |       |

|   | F     | D     | v*(5) | v'*(5) | u*(5) | F     | D     | v*(6) | v'*(6) | u*(6) |
|---|-------|-------|-------|--------|-------|-------|-------|-------|--------|-------|
| M | 3,625 | 6,563 | 6,563 | 0      | D     | 3,688 | 6,719 | 6,719 | 0      | D     |
| P | 4,656 | 8,281 | 8,281 | 1,719  | D     | 4,672 | 8,359 | 8,359 | 1,641  | D     |
| g | 6,563 |       |       |        |       | 6,719 |       |       |        |       |

La política {D, D} es óptima desde el primer momento, lo cual permite afirmar que, casi con toda seguridad, será también óptima a largo plazo (el caso es tan sencillo que podemos afirmarlo sin dudar). Sin embargo, debe verificarse ya



que, aunque los valores tanto de  $v^*$  como de  $g$  son similares, no se repiten exactamente y, por otra parte, falta hallar cuál es exactamente el límite hacia el cual tienden.

Planteemos el sistema.

$$\mathbf{P} = \begin{vmatrix} 0 & 1 \\ 0,5 & 0,5 \end{vmatrix} \quad \mathbf{R} = \begin{vmatrix} 10 & 5 \\ 10 & 5 \end{vmatrix} \quad \mathbf{q} = \begin{vmatrix} 5 \\ 7,5 \end{vmatrix}$$

La cadena inducida por la política es claramente ergódica, con lo cual:

$$w_M + g = 5 + w_P$$

$$g = 20/3 = 6,667;$$

$$w_P + g = 7,5 + 0,5 w_M + 0,5 w_P \Rightarrow w_M = 0; w_P = 5/3 = 1,667$$

Comprobemos si es óptima:

$$w_M: \max \{3 + 0,6 w_M + 0,4 w_P; 5 + w_P\} - g = \max \{-3; 0\} = 0 = w_M \Rightarrow \text{OK}$$

$$w_P: \max \{4,5 + 0,9 w_M + 0,1 w_P; 7,5 + 0,5 w_M + 0,5 w_P\} - g = \\ = \max \{-2; 1,667\} = w_P \Rightarrow \text{OK}$$

Se observa que los valores obtenidos tanto para la  $g$  como para las  $w_i$  a corto plazo, aunque lentamente, tienden a los valores del largo plazo.

- b) La universidad ha fijado unas normas para valorar la calidad de la docencia del profesorado, en 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 de las de la universidad) y tenemos la situación siguiente:

| Estado | Decisión | Matriz P |     | Matriz R |   | Q |
|--------|----------|----------|-----|----------|---|---|
|        |          | M        | P   | M        | P |   |
| M      | F        | 0,6      | 0,4 | 10       | 0 | 6 |
|        | D        | 0        | 1   | 10       | 0 | 0 |
| P      | F        | 0,9      | 0,1 | 10       | 0 | 9 |
|        | D        | 0,5      | 0,5 | 10       | 0 | 5 |

Tabla 6.6.  
Nuevas  
probabilidades,  
remuneraciones y  
ganancias  
esperadas



Podemos ensayar la política "miope" {F,F} para ver si es óptima a largo plazo:

$$w_M + g = 6 + 0,6 w_M + 0,4 w_P \quad g = 90/13 = 6,923;$$

$$w_P + g = 9 + 0,9 w_M + 0,1 w_P \Rightarrow w_M = 0; w_P = 30/13 = 2,308$$

Comprobemos si es óptima:

$$w_M: \max \{6 + 0,6 w_M + 0,4 w_P; 0 + w_P\} - g =$$

$$= \max \{0; -60/13\} = 0 = w_M \Rightarrow \text{OK}$$

$$w_P: \max \{9 + 0,9 w_M + 0,1 w_P; 5 + 0,5 w_M + 0,5 w_P\} - g =$$

$$= \max \{30/13; -10/13\} = w_P \Rightarrow \text{OK}$$

Es decir, al utilizar criterios diferentes de valoración, la política óptima según los criterios del profesor y según los de la universidad es distinta; además, las valoraciones son también muy diferentes.

Si evaluamos la política óptima en cada caso, con ambos criterios de remuneración obtenemos los resultados de la tabla 6.7.

Tabla 6.7.  
Política óptima  
del ejemplo 6.3.b

| Política    | P*         | Valoración según |             |
|-------------|------------|------------------|-------------|
|             |            | Profesor         | Universidad |
| <b>D, D</b> | 1/3; 2/3   | 6,667            | 3,333       |
| <b>F, F</b> | 9/13; 4/13 | 3,462            | 6,923       |



→7



## Problemas resueltos

### 7.1 Revisión de los conceptos teóricos

Una cadena de Markov tiene 7 estados (A, B, ..., G), de los cuales el A es absorbente; B, C, y D son transitorios, y el resto no son ni una cosa ni otra. Indica cuáles de las frases siguientes son ciertas y cuáles son falsas (si se puede saber; si no se puede, explica por qué no se puede):

- a) Las columnas de la matriz  $P$ ... suman todas 1 / pueden sumar todas 1 / hay al menos una que no suma 1.
- b) La cadena tiene... una clase final / dos clases finales / al menos dos clases finales.
- c) La cadena es... ergódica / semi-ergódica / no ergódica.
- d) Hay una clase final cíclica / No hay ninguna clase final cíclica / Hay más de una clase final cíclica.
- e) La matriz  $P^*$ ... tiene por lo menos tres columnas de ceros / tiene exactamente tres columnas de ceros / tiene más de tres columnas de ceros / no tiene ninguna columna de ceros.
- f) Una de las columnas de  $P^*$  es toda a ceros, excepto una celda... que vale 1 / que tiene un valor  $> 0$  y  $< 1$  / no hay ninguna columna así.
- g) Las filas de  $P^*$ ... son todas iguales / son todas diferentes / las hay de tres formatos diferentes / puede haber de hasta seis formatos diferentes / puede haber de hasta cinco formatos diferentes / pueden ser todas iguales.



## Solución

a) Hay al menos una que no suma 1. Nota importante: Se pregunta por las columnas, no por las filas.

La columna A suma más que 1, ya que tiene un 1 en la fila A y algún valor diferente de 0 al menos en una de las filas B, C o D. Y, puesto que el total de los elementos de la matriz ha de sumar 7, debe haber al menos una columna que sume menos que 1.

Por otra parte, las columnas B, C y D solo son  $> 0$  en estas mismas filas, y los valores de esta matriz  $3 \times 3$  no suman 3; por tanto, al menos una de estas columnas ha de sumar menos que 1.

b) La cadena tiene dos clases finales. Si los estados E, F y G no son ni transitorios ni absorbentes, han de formar una sola clase final (si no, uno de ellos sería un estado absorbente).

c) La cadena es no ergódica, ya que tiene dos clases finales.

d) Hay una clase final cíclica / No hay ninguna clase final cíclica / Hay más de una clase final cíclica.

Una de las dos primeras frases es cierta y la otra es falsa, y la tercera opción seguro que es falsa. La clase {A} es acíclica y la clase {E, F, G} puede ser cíclica (con  $p = 2$  o  $p = 3$ ) o acíclica.

e) La matriz  $P^*$  tiene exactamente tres columnas de ceros, que corresponden a los tres estados transitorios, B, C y D.

f) No hay ninguna columna así porque la columna A tiene un 1 en la fila A, pero al menos uno de los valores de las filas B, C o D han de ser diferentes de 0. Las demás columnas son todas 0 (B, C y D) o tienen las filas E, F y G, y al menos una de las columnas B, C y D, diferentes de 0.

g) Puede haber de hasta cinco formatos diferentes. La fila A tiene el formato (1 0 0 0 0 0) y las filas E, F y G son iguales con el formato (0 0 0 0 e f g), pero las filas B, C y D pueden ser iguales o diferentes; tan solo sabemos que serán 0 las columnas B, C y D.



## 7.2 Un poco más de revisión de conceptos

Una cadena de Markov se rige por la siguiente matriz de probabilidades:

|   | A  | B   | C   | D   | E   |
|---|----|-----|-----|-----|-----|
| A |    | 0,3 | 0,2 | 0,1 | 0,1 |
| B |    | 0,0 | 1,0 | 0,0 | 0,0 |
| C | ¿? | 0,0 | 0,5 | 0,4 | 0,1 |
| D |    | 1,0 | 0,0 | 0,0 | 0,0 |
| E |    | 1,0 | 0,0 | 0,0 | 0,0 |

Tabla 7.1.  
Matriz de  
probabilidades sin  
columna A

- Indica los valores que faltan en la columna A de la matriz.
- Clasifica la cadena y calcula la matriz de probabilidades estacionarias.
- Indica la probabilidad de hallar el sistema en cada uno de los estados si sabemos que su evolución ha empezado en el estado A y que han transcurrido 5.400 transiciones. ¿Y si ha empezado en el B?
- Se asocia a esta cadena una remuneración igual a 100 veces la probabilidad de cada transición. ¿Cuál es la ganancia media por transición a largo plazo?

### Solución

- a) Dado que la suma de las filas ha de ser 1, la matriz P será:

|   | A   | B   | C   | D   | E   |
|---|-----|-----|-----|-----|-----|
| A | 0,3 | 0,3 | 0,2 | 0,1 | 0,1 |
| B | 0,0 | 0,0 | 1,0 | 0,0 | 0,0 |
| C | 0,0 | 0,0 | 0,5 | 0,4 | 0,1 |
| D | 0,0 | 1,0 | 0,0 | 0,0 | 0,0 |
| E | 0,0 | 1,0 | 0,0 | 0,0 | 0,0 |

Tabla 7.2.  
Matriz de  
probabilidades  
con la solución  
de la columna A

- b) La figura 7.1 representa el grafo de la matriz P.

Hay un estado transitorio (A) y una sola clase final {B, C, D, E}, con  $p = 1$ .

Se trata, pues de una cadena semiergódica semirregular. La matriz  $P^*$  tendrá todas las filas iguales y la primera columna de ceros.

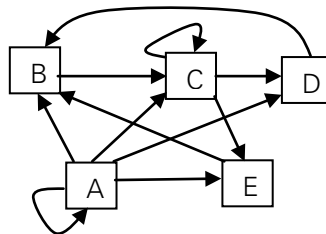


Figura 7.1.  
Grafo correspondiente  
a la matriz P



$$b + c + d + e = 1$$

$$b = d + e$$

$$b = 1/4; c = 1/2; d = 1/5; e = 1/20$$

$$c = b + 0,5 c \quad \Rightarrow$$

$$d = 0,4 c$$

$$P^* = [0; 1/4; 1/2; 1/5; 1/20]$$

$$e = 0,1 c$$

Recordemos que, aunque el sistema de ecuaciones tiene una más (5) que incógnitas (4), en realidad una de las cuatro últimas es redundante con las otras tres, por lo cual debemos prescindir de una de ellas; no así de la primera, que es obligada.

c) Las probabilidades son las indicadas en la  $P^*$ , ya que 5.400 transiciones son suficientes para estar ya en el régimen permanente, y se trata de una cadena acíclica.

d) Calculemos los vectores  $q$  y  $g$ :

$$q = \begin{vmatrix} 30*0,3 + 30*0,3 + 20*0,2 + 20*0,2 + 10*0,1 \\ 100*1 \\ 50*0,5 + 40*0,4 + 10*0,1 \\ 100*1 \\ 100*1 \end{vmatrix} = \begin{vmatrix} 24 \\ 100 \\ 42 \\ 100 \\ 100 \end{vmatrix}$$

$$g = P^* q = 71$$

### 7.3 Saratres Airlines

Saratres Airlines tiene, entre otros, diversos aviones del modelo PI-F1A con los cuales realiza vuelos entre las ciudades A, B, C y D. Dado que la ruta concreta que hará cada avión depende de la programación del momento, la ciudad a la cual se dirigirá cada avión después de cada vuelo es aleatoria, y estadísticamente se han determinado las frecuencias de la tabla 7.3. El tiempo de vuelo en horas entre cada par de ciudades es el que figura en la tabla 7.4 (el tiempo de ida y el de vuelta son diferentes, debido a la presencia de vientos dominantes).

- a) Modeliza la situación con la técnica adecuada, indicando claramente cuáles son las transiciones, los estados, las decisiones y las remuneraciones.



- b) ¿Cuál es la duración media de un vuelo de este avión?

|   | A   | B   | C   | D   |
|---|-----|-----|-----|-----|
| A | 0   | 0,4 | 0,4 | 0,2 |
| B | 0,4 | 0   | 0,4 | 0,2 |
| C | 0,5 | 0,3 | 0   | 0,2 |
| D | 0,3 | 0,3 | 0,4 | 0   |

Tabla 7.3.  
Frecuencias de vuelo  
entre dos ciudades

|   | A   | B   | C   | D   |
|---|-----|-----|-----|-----|
| A | 0   | 3,8 | 5,5 | 6,4 |
| B | 3,7 | 0   | 1,7 | 4,2 |
| C | 5,1 | 2,5 | 0   | 3,5 |
| D | 7,2 | 4,4 | 3,8 | 0   |

Tabla 7.4.  
Horas de vuelo entre  
dos ciudades

La compañía desea establecer unos controles periódicos especiales para estos aviones, para lo cual necesita un taller, que se establecerá en una de las cuatro ciudades donde opera. Se ha decidido que será aquella adonde vuelen los aviones con más frecuencia.

- c) ¿En qué ciudad se establecerá el control? Si el avión pasa el control cada vez que llegue a dicha ciudad, ¿cada cuántos vuelos pasará el control, de media?
- d) ¿Cuántas horas de vuelo habrá realizado por término medio un avión entre cada dos controles?

## Solución

- a) Una transición es un vuelo y el estado, la ciudad donde se halla el avión. No hay decisión, y la remuneración son las horas de vuelo. El sistema es aleatorio, homogéneo en el tiempo y con horizonte no finito; por tanto, debe modelizarse con una cadena de Markov con remuneración.

|   | P   |     |     |     | R   |     |     |     | q |
|---|-----|-----|-----|-----|-----|-----|-----|-----|---|
|   | A   | B   | C   | D   | A   | B   | C   | D   |   |
| A | 0   | 0,4 | 0,4 | 0,2 | 0   | 3,8 | 5,5 | 6,4 | 5 |
| B | 0,4 | 0   | 0,4 | 0,2 | 3,7 | 0   | 1,7 | 4,2 | 3 |
| C | 0,5 | 0,3 | 0   | 0,2 | 5,1 | 2,5 | 0   | 3,5 | 4 |
| D | 0,3 | 0,3 | 0,4 | 0   | 7,2 | 4,4 | 3,8 | 0   | 5 |

Tabla 7.5.  
Probabilidades,  
remuneraciones y  
ganancias esperadas

Se trata de una cadena ergódica acíclica, es decir, regular positiva.



## b) Duración media de un vuelo

Para saber la duración media de un vuelo, hay que calcular la  $g$ . Puesto que no interesa saber los transitorios, calculemos la  $P^*$ , que tendrá todas las filas iguales:

$$a + b + c + d = 1 \quad a = 173/588$$

$$b = 0,4 a + 0,3 c + 0,3 d \quad b = 149/588$$

$$c = 0,4 a + 0,4 b + 0,4 d \Rightarrow c = 2/7$$

$$d = 0,2 a + 0,2 b + 0,2 \quad d = 1/6$$

Así pues:

$$g = P^* \cdot q = [173/588 \ 149/588 \ 2/7 \ 1/6] \cdot [5 \ 3 \ 4 \ 5]^T = 1.237 / 294 \approx 4,2075 \text{ horas}$$

## c) ¿Dónde estableceremos el control?

Los valores de  $P^*$  son, aproximadamente: 0,2942; 0,2534; 0,2857; 0,1667

Por tanto, estableceremos el control en A, que es adonde el avión viaja más a menudo (aunque por poco).

Entre dos controles, el número medio de vuelos será la inversa de la probabilidad:  $588/173 \approx 3,3988$  vuelos

## d) ¿Cuántas horas de vuelo habrá realizado de media entre dos controles?

Puesto que la duración media de cada vuelo es de 4,2075 horas, habrá realizado de media:  $3,3988 \cdot 4,2075 = 14,30$  horas de vuelo.

Tabla 7.6.  
Duración  
acumulada de  
los vuelos según  
la ciudad de  
origen

|   | v(0) | v(1) | v(2) | v(3)  | v(4)  | v(5)  | v(6)  | v(7)  |
|---|------|------|------|-------|-------|-------|-------|-------|
| A | 0    | 5    | 8,8  | 13,2  | 17,32 | 21,57 | 25,76 | 29,97 |
| B | 0    | 3    | 7,6  | 11,68 | 15,93 | 20,12 | 24,33 | 28,54 |
| C | 0    | 4    | 8,4  | 12,48 | 16,76 | 20,93 | 25,16 | 29,35 |
| D | 0    | 5    | 9    | 13,28 | 17,46 | 21,68 | 25,88 | 30,09 |

Nota: En realidad, el tiempo entre dos controles puede considerarse una evolución a corto plazo partiendo del estado A; sin embargo, dado que el número de vuelos entre dos controles es aleatorio, no sabemos qué N debe tomarse y, por tanto, no podemos calcular la remuneración a corto plazo, partiendo de A, lo cual nos daría una distribución de las horas esperadas según el número de vuelos, pero no el valor esperado total.



El resultado es que, si realiza tres vuelos, el valor esperado será de 13,2 horas, y de 17,32 si hace 4. Por supuesto, cuantos más vuelos realice el avión antes de volver a A, más horas habrá volado: los valores se indican en la tabla 7.6 (fila A).

Podemos conocer la probabilidad del número de vuelos antes de volver a A calculando las sucesivas potencias de P:

$$\begin{aligned}
 P^2 &= \begin{bmatrix} \mathbf{0,420} & 0,180 & 0,240 & 0,160 \\ 0,260 & 0,340 & 0,240 & 0,160 \\ 0,180 & 0,260 & 0,400 & 0,160 \\ 0,320 & 0,240 & 0,240 & 0,200 \end{bmatrix} & P^3 &= \begin{bmatrix} \mathbf{0,240} & 0,288 & 0,304 & 0,168 \\ 0,304 & 0,224 & 0,304 & 0,168 \\ 0,352 & 0,240 & 0,240 & 0,168 \\ 0,276 & 0,260 & 0,304 & 0,160 \end{bmatrix} \\
 P^4 &= \begin{bmatrix} \mathbf{0,318} & 0,238 & 0,278 & 0,166 \\ 0,292 & 0,263 & 0,278 & 0,166 \\ 0,266 & 0,263 & 0,304 & 0,166 \\ 0,304 & 0,250 & 0,278 & 0,168 \end{bmatrix} & P^5 &= \begin{bmatrix} \mathbf{0,284} & 0,260 & 0,289 & 0,167 \\ 0,294 & 0,250 & 0,289 & 0,167 \\ 0,307 & 0,248 & 0,278 & 0,167 \\ 0,289 & 0,256 & 0,289 & 0,166 \end{bmatrix} \\
 P^6 &= \begin{bmatrix} \mathbf{0,299} & 0,250 & 0,285 & 0,167 \\ 0,294 & 0,254 & 0,285 & 0,167 \\ 0,288 & 0,256 & 0,289 & 0,167 \\ 0,296 & 0,252 & 0,285 & 0,167 \end{bmatrix} & P^7 &= \begin{bmatrix} \mathbf{0,292} & 0,255 & 0,286 & 0,167 \\ 0,294 & 0,253 & 0,286 & 0,167 \\ 0,297 & 0,252 & 0,285 & 0,167 \\ 0,293 & 0,254 & 0,286 & 0,167 \end{bmatrix}
 \end{aligned}$$

Como puede observarse, los valores tienden muy rápidamente hacia los de  $P^*$ , aunque existe una cierta oscilación entre las transiciones pares e impares, debido a que la transición A-A no es posible.

## 7.4 El servicio en el tenis

Como es bien conocido, en el deporte de tenis, el jugador que sirve tiene dos oportunidades para que la pelota entre en el cuadro 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 que la pelota entre en el cuadro es de 0,8. Ahora bien, si entra el servicio fuerte, la probabilidad de ganar el punto es de 0,7 y, si entra el servicio flojo, la probabilidad de ganar el punto es de 0,4. Se desea conocer la estrategia a seguir para maximizar la ganancia esperada por servicio a largo plazo.

- Formula 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, identifica una posible política de servicio óptima.



- c) Determina la ganancia esperada por servicio a largo plazo, si el tenista opta siempre por el servicio fuerte.
- d) Determina la política óptima a largo plazo y la ganancia correspondiente.

### Solución

Como siempre ocurre, la forma más adecuada de resolver el problema es definiendo sus parámetros; en este caso, podemos hacerlo del modo siguiente:

- *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 la bola 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 es si estamos en el primer servicio (estado 1) o en el segundo (estado 2). Por tanto, se trata de un sistema dinámico, con dos estados.
- *Variable de decisión*: el tenista puede controlar, en parte, la evolución del sistema con sus acciones. 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.
- La ganancia (remuneración) debe indicar si se gana o si se pierde el punto. Podemos establecer cualquier valoración que deseemos: en este caso, hemos optado por definir que la ganancia por punto ganado es de +10 y la ganancia por punto perdido es -10.

Por tanto, nos encontramos ante un sistema de Markov con remuneración y decisión (que también puede tratarse como un modelo de programación dinámica aleatoria homogénea en el tiempo).

Las reglas de 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 los estados del sistema y de las 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). Si la bola no entra, pasaremos al estado 2. Veamos qué probabilidades y qué remuneraciones obtenemos para cada una de las variables de decisión.



Si el tenista saca fuerte, la probabilidad de entrar el servicio es de 0,5. Esta será, entonces, la probabilidad de transición. Dado que 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. Esta 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, eso es, evolucionamos al estado 2. Al no disputarse el punto, la remuneración será 0. Las probabilidades de que esto suceda dependerán 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 tanto, en el estado 2, la probabilidad de transición al estado 1 es 1 y, en consecuencia, la probabilidad de transición al estado 2 es 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:

Si saca fuerte:  $0,5 \cdot 0,7 \cdot (+10) + \{0,5 \cdot 0,3 + 0,5\} \cdot (-10) = -3$

Si saca flojo:  $0,8 \cdot 0,4 \cdot (+10) + \{0,8 \cdot 0,6 + 0\} \cdot (-10) = -3,6$

### *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 del modo siguiente:

|          |    | Matriz P |          | Matriz R |          |      |
|----------|----|----------|----------|----------|----------|------|
|          |    | Estado 1 | Estado 2 | Estado 1 | Estado 2 | q    |
| 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 |

Tabla 7.7.  
Probabilidades,  
remuneraciones  
y ganancia  
esperada



Examinemos, primero, la política óptima a corto plazo (esto es, cuando el número de bolas que quedan para ganar el punto es pequeño):

Tabla 7.8  
Solución a  
corto plazo

|          | <b>v(0)</b> | FU | FL   | <b>v*(1)</b> | <b>v'*(1)</b> | <b>u*(1)</b> | FU  | FL  | <b>v*(2)</b> | <b>v'*(2)</b> | <b>u*(2)</b> |
|----------|-------------|----|------|--------------|---------------|--------------|-----|-----|--------------|---------------|--------------|
| <b>1</b> | <b>0</b>    | 2  | -1,6 | <b>2</b>     | <b>5</b>      | <b>FU</b>    | 4,5 | 2,4 | <b>4,5</b>   | <b>2,5</b>    | <b>FU</b>    |
| <b>2</b> | <b>0</b>    | -3 | -3,6 | <b>-3</b>    | <b>0</b>      | <b>FU</b>    | 2   | 1,4 | <b>2</b>     | <b>0</b>      | <b>FU</b>    |
| <b>g</b> |             |    |      | <b>-3</b>    |               |              |     |     | <b>2</b>     |               |              |

|          | FU   | FL   | <b>v*(3)</b> | <b>v'*(3)</b> | <b>u*(3)</b> | FU    | FL   | <b>v*(4)</b> | <b>v'*(4)</b> | <b>u*(4)</b> |
|----------|------|------|--------------|---------------|--------------|-------|------|--------------|---------------|--------------|
| <b>1</b> | 3,25 | 0,4  | <b>3,25</b>  | <b>3,75</b>   | <b>FU</b>    | 3,875 | 1,4  | <b>3,875</b> | <b>3,125</b>  | <b>FU</b>    |
| <b>2</b> | -0,5 | -1,1 | <b>-0,5</b>  | <b>0</b>      | <b>FU</b>    | 0,75  | 0,15 | <b>0,75</b>  | <b>0</b>      | <b>FU</b>    |
| <b>g</b> |      |      | <b>-0,5</b>  |               |              |       |      | <b>0,75</b>  |               |              |

|          | FU    | FL    | <b>v*(5)</b> | <b>v'*(5)</b> | <b>u*(5)</b> | FU    | FL    | <b>v*(6)</b> | <b>v'*(6)</b> | <b>u*(6)</b> |
|----------|-------|-------|--------------|---------------|--------------|-------|-------|--------------|---------------|--------------|
| <b>1</b> | 3,563 | 0,9   | <b>3,563</b> | <b>3,438</b>  | <b>FU</b>    | 3,719 | 1,15  | <b>3,719</b> | <b>3,281</b>  | <b>FU</b>    |
| <b>2</b> | 0,125 | -0,48 | <b>0,125</b> | <b>0</b>      | <b>FU</b>    | 0,438 | -0,16 | <b>0,438</b> | <b>0</b>      | <b>FU</b>    |
| <b>g</b> |       |       | <b>0,125</b> |               |              |       |       | <b>0,438</b> |               |              |

La política {FU, FU}, eso es, sacar siempre fuerte, resulta óptima a corto plazo y, por lo que parece, también a largo. Además, la situación es mejor en el primer servicio que en el segundo (obvio) y esta diferencia se puede cuantificar alrededor de 3,3 puntos en la escala elegida. Sin embargo, el valor de  $g$  que nos indica la frecuencia con que se ganará el punto todavía oscila demasiado para poder estimarla correctamente a largo plazo. Comprobemos si la política es también óptima a largo plazo y veamos cuál es su  $g$ :

$$w_1 + g = 2 + 0,5 w_1 + 0,5 w_2 \quad g = 1/3 = 0,333;$$

$$w_2 + g = -3 + w_1 \quad \Rightarrow \quad w_1 = 3 + 1/3 = 3,333; w_2 = 0$$

Comprobemos si es óptima:

$$w_1: \max \{2 + 0,5 w_1 + 0,5 w_2; -1,6 + 0,8 w_1 + 0,2 w_2\} - g =$$

$$= \max \{3,333; -3,6\} = w_1 \Rightarrow \text{OK}$$

$$w_2: \max \{-3 + w_1; -3,6 + w_1\} - g = \max \{0; -2,6\} = w_1 \Rightarrow \text{OK}$$

Por tanto, el tenista debe sacar siempre fuerte. Podemos preguntarnos con qué frecuencia ganará el punto si sigue esta estrategia. Sea  $p$  dicha probabilidad. Tal como hemos definido la remuneración, deberá cumplirse:

$$10 p - 10 (1-p) = g = 1/3$$

de donde  $p = 0,5167$

Es decir, con esta política, el tenista ganará, a largo plazo, el 51,67 % de los puntos.



## 7.5 Tres elementos delicados

La rentabilidad de un sistema productivo viene determinada por el estado de tres elementos muy delicados (A, B y C), que necesitan un mantenimiento constante. El rendimiento total del sistema en unas unidades arbitrarias viene dado por la suma de los rendimientos de cada uno de los tres elementos, según la tabla siguiente:

| Elemento     | A   | B  | C   |
|--------------|-----|----|-----|
| <b>Día 1</b> | 100 | 75 | 125 |
| <b>Resto</b> | 40  | 50 | 60  |

Tabla 7.9.  
Rendimientos de  
los tres  
elementos

Hay un único equipo de mantenimiento que cada día laborable puede hacer el mantenimiento de solo uno de los tres elementos; durante ese día, el elemento que está siendo cambiado tiene rendimiento 0 y, al día siguiente, tiene el rendimiento indicado como "Primer día"; a partir del segundo día después de la intervención de mantenimiento, y durante un tiempo indefinido, el rendimiento del elemento será el que se indica como "Resto".

- Analiza las características del caso y modelízalo con la técnica que consideres más apropiada. Justifica por qué has elegido esta técnica e indica con detalle qué significa cada uno de los elementos del modelo.
- Históricamente, la empresa sigue la regla de hacer el mantenimiento de los tres elementos de manera rotativa: el primer día se cambia el elemento A; al día siguiente, el B, y, al siguiente, el C, y se vuelve a empezar el ciclo. Determina el rendimiento diario con esta regla e indica si es óptimo; si no lo es, halla el óptimo. ¿Hay algo que llame la atención en esta política?

### Solución

a) Se trata de un sistema homogéneo en el tiempo, con horizonte ilimitado y determinista, con remuneración y decisión. Por tanto, se puede tratar por Markov o por programación dinámica, indistintamente.

Una transición o etapa será un día laborable. La decisión es atender A, B, C o no hacer nada ( $\emptyset$ ). La remuneración es el rendimiento, que es MAX.

El estado del sistema viene dado por el estado de los tres elementos, que pueden estar en el primer día (P) o en el resto (R). Hay, por tanto,  $2^3 = 8$  estados pero, como solo se puede hacer el mantenimiento de un equipo al día, solo son accesibles los estados en que hay uno o ningún elemento en estado P. Además, no tiene sentido cambiar los elementos que están en el estado P. La tabla 7.10 ilustra la cadena de Markov.



Tabla 7.10.  
Cadena de  
Markov con  
remuneración y  
decisión

| Estado | Decisión | Nuevo estado | Rendimiento |
|--------|----------|--------------|-------------|
| 1.PRR  | ∅        | 4.RRR        | 210         |
|        | B        | 2.RPR        | 160         |
|        | C        | 3.RRP        | 150         |
| 2.RPR  | ∅        | 4.RRR        | 175         |
|        | A        | 1.PRR        | 135         |
|        | C        | 3.RRP        | 115         |
| 3.RRP  | ∅        | 4.RRR        | 215         |
|        | A        | 1.PRR        | 175         |
|        | B        | 2.RPR        | 165         |
| 4.RRR  | ∅        | 4.RRR        | 150         |
|        | A        | 1.PRR        | 110         |
|        | B        | 2.RPR        | 100         |
|        | C        | 3.RRP        | 90          |

b) La regla indicada corresponde a la política {B, C, A, A}

(En realidad, no nos indican qué debe hacerse en el estado 4, que será transitorio con esta regla; hemos elegido A porque es la que da un rendimiento mayor en la primera iteración.)

Planteemos el sistema:

$$\begin{aligned}
 w_1 + g &= 160 + w_2 & g &= 150; \\
 w_2 + g &= 115 + w_3 & w_2 &= 0; \quad w_1 = 10 \\
 w_3 + g &= 175 + w_1 & w_3 &= 35; \quad w_4 = -30 \\
 w_4 + g &= 110 + w_1
 \end{aligned}
 \Rightarrow$$

Comprobemos:

$$\text{MAX } \{210 - 30; 160 + 0; \mathbf{150 + 35}\} - 150 = 35 \Rightarrow C \quad \text{Cambiar}$$

$$\text{MAX } \{175 - 30; 135 + 10; \mathbf{115 + 35}\} - 150 = 0 \Rightarrow C \quad \text{OK}$$

$$\text{MAX } \{\mathbf{215 - 30}; \mathbf{175 + 10}; 165 + 0\} - 150 = 35 \Rightarrow \emptyset \text{ o } A \quad \text{Puede cambiar}$$

$$\text{MAX } \{150 - 30; 110 + 10; 100 + 0; \mathbf{90 + 35}\} - 150 = -25 \Rightarrow C \quad \text{Cambiar}$$

Las políticas {C,C, ∅,C} y {C,C,A,C} son mejores que la dada.

La primera de estas políticas tampoco es óptima: su  $g$  es 152,5. La política óptima es la segunda (cambiar A y C), cuya  $g$  = 162,5.



Observamos que estas políticas son muy peculiares. La primera consiste en cambiar C un día y no hacer nada al siguiente. La segunda es alternar las intervenciones de A y C. En ningún caso se cambia el elemento B. Esto se debe a que el modelo supone que los elementos nunca se estropean ni pierden rendimiento aunque hayan estado funcionando un período indefinidamente largo. Lógicamente, esto es difícilmente aceptable en la práctica, por lo cual deberemos considerar si el modelo está realmente bien definido. Sin embargo, el enunciado no nos aporta datos suficientes para establecer un modelo mejor. Sin más datos que los aportados por el enunciado, la recomendación que haríamos sería cambiar un día A y, al siguiente, C, y que B siga funcionando hasta que se estropee. Solo cambiar B cuando se estropee y, al día siguiente, cambiar A, ya que el elemento C usado nos da un mejor rendimiento que A (60 contra 40).

## 7.6 Embarcaciones de lujo

En el embarcadero de un negocio de embarcaciones de lujo solamente caben 2 embarcaciones, pero esto no preocupa a su propietario, pues ha observado que, en una semana, no venderá nunca más de 2 embarcaciones; concretamente, la demanda es de 0 embarcaciones el 50 % de las semanas, de 1 embarcación el 40 %, y solo en el 10 % de las semanas la demanda es de 2 embarcaciones. Cada semana, antes de finalizar el viernes, puede pedir que le suministren tantas embarcaciones como desee, a un coste fijo de 10.000 € por el transporte, más un coste variable de 5.000 € por embarcación; como sus clientes son compradores compulsivos, es importante no quedarse sin *stock*, porque tan solo podrá vender las embarcaciones que tenga en el embarcadero cuando se produzca la demanda. Cada venta reporta unos ingresos de 15.000 € por embarcación y el propietario no considera ningún coste por tener las embarcaciones amarradas.

- a) El propietario del negocio sigue la política de pedir 2 embarcaciones cuando ya no le queda ninguna, y no pedir nada en caso contrario. ¿Cuál es la ganancia media esperada por semana?
- b) Determina si la política es óptima y, en caso contrario, halla la óptima.

### Solución

a) Las transiciones son lo que sucede cada semana, los estados (0, 1, 2) son las embarcaciones que hay el viernes por la tarde, la decisión es la cantidad de embarcaciones que el propietario encarga y la remuneración es la ganancia que obtiene (ingresos menos costes), en miles de euros.

Observemos que la transición va desde el viernes por la tarde hasta el viernes siguiente, con lo cual a lo largo de la transición han llegado las embarcaciones



pedidas y se han vendido las de la semana. Las matrices P y R y el vector q se ilustran en la tabla 7.11.

Tabla 7.11.  
Probabilidades,  
remuneraciones  
y ganancias  
esperadas

| Estado | Decisión | Matriz P |     |     | Matriz R |     |     | q    |
|--------|----------|----------|-----|-----|----------|-----|-----|------|
|        |          | 0        | 1   | 2   | 0        | 1   | 2   |      |
| 0      | 0        | 1        | 0   | 0   | 0        | -   | -   | 0    |
|        | 1        | 0,5      | 0,5 | 0   | 4        | -11 | -   | -3,5 |
|        | 2        | 0,1      | 0,4 | 0,5 | 18       | 3   | -12 | -3   |
| 1      | 0        | 0,5      | 0,5 | 0   | 15       | 0   | -   | 7,5  |
|        | 1        | 0,1      | 0,4 | 0,5 | 19       | 4   | -11 | -3   |
| 2      | 0        | 0,1      | 0,4 | 0,5 | 30       | 15  | 0   | 9    |

La política que sigue nuestro vendedor es (2, 0, 0) y lleva a la cadena de Markov de la tabla 7.12.

Tabla 7.12.  
Cadena de  
Markov con  
remuneración

| Estado | Decisión | P   |     |     | R  |    |     | q   |
|--------|----------|-----|-----|-----|----|----|-----|-----|
|        |          | 0   | 1   | 2   | 0  | 1  | 2   |     |
| 0      | 2        | 0,1 | 0,4 | 0,5 | 18 | 3  | -12 | -3  |
| 1      | 0        | 0,5 | 0,5 | 0   | 15 | 0  | -   | 7,5 |
| 2      | 0        | 0,1 | 0,4 | 0,5 | 30 | 15 | 0   | 9   |

de donde  $P^* = [5/18, 8/18, 5/18]$  y  $g = P^* q = 5$  K€/semana

Para saber si es óptima, hay que determinar las w con el sistema siguiente:

$$w_0 + g = -3 + 0,1 w_0 + 0,4 w_1 + 0,5 w_2$$

$$w_1 + g = 7,5 + 0,5 w_0 + 0,5 w_1$$

$$w_2 + g = 9 + 0,1 w_0 + 0,4 w_1 + 0,5 w_2$$

Haciendo  $w_0 = 0$  se obtiene  $g = 5$  (como era de esperar);  $w_1 = 5$   $w_2 = 12$ , y, comprobando si cumplen la condición de óptimo:

$$W_0 = \text{MAX} [0 + w_0; -3,5 + 0,5 w_0 + 0,5 w_1; -3 + 0,1 w_0 + 0,4 w_1 + 0,5 w_2] - g = 0 = w_0 \Rightarrow \text{OK}$$

$$W_1 = \text{MAX} [7,5 + 0,5 w_0 + 0,5 w_1; -3 + 0,1 w_0 + 0,4 w_1 + 0,5 w_2] - g = 5 = w_1 \Rightarrow \text{OK}$$

$$W_2 = \text{MAX} [9 + 0,1 w_0 + 0,4 w_1 + 0,5 w_2] - g = 12 = w_2 \Rightarrow \text{OK}$$

Por tanto, la política es óptima.



## 7.7 Mantenimiento

Una línea automática de mecanizado trabaja cada semana de lunes a viernes ininterrumpidamente, y el sábado se aprovecha para hacer diversas operaciones de mantenimiento, una de las cuales es revisar una herramienta sometida a un gran desgaste.

Para describir el desgaste de la herramienta, se han definido los siguientes niveles: Nueva (N), Seminueva (S), Usada (U), Desgastada (D) y Rota (R). Para evaluar las probabilidades de cambio de estado, se dispone de una estadística que indica el número de veces que se ha observado el cambio indicado en una semana:

| Estado el sábado →              |   | N  | S  | U  | D | R |
|---------------------------------|---|----|----|----|---|---|
| Estado el lunes por la mañana → | N | 20 | 15 | 15 |   |   |
|                                 | S |    | 16 | 20 | 4 |   |
|                                 | U |    |    | 3  | 5 | 2 |
|                                 | D |    |    |    | 1 | 1 |

Tabla 7.14a.  
Números de cambios de estado observados en una semana

Una herramienta rota (R) siempre se queda rota, naturalmente.

Si la herramienta no es nueva (N), hay unos costes por pérdida de producción que dependen del estado inicial y final de la misma; en concreto, los costes estimados si la herramienta acaba la semana igual que la ha empezado se indican en la tabla 7.14.

| Estado    | N | S   | U   | D     | R      |
|-----------|---|-----|-----|-------|--------|
| Coste (€) | 0 | 200 | 500 | 1.000 | 60.000 |

Tabla 7.14b.  
Costes semanales por pérdida de producción

Si el estado final no es igual al inicial, se considera un coste, que es la semisuma de los correspondientes a estos dos estados: por ejemplo si empieza N y acaba U, el coste es 250 €.

El problema que se debe estudiar es cuándo y cómo hay que cambiar la herramienta, porque una herramienta nueva es muy cara (2.500 €), si bien también cabe la posibilidad de poner una de segunda mano, en estado S, que tan solo cuesta 1.000 €.

- Desde tiempo inmemorial, existe en la empresa la polémica de si es mejor poner una herramienta nueva cuando se halla en estado S o peor, o bien es preferible esperar a que esté U o peor y entonces poner una seminueva. ¿Cuál de las dos propuestas es mejor?
- Determina, para cualquiera de las dos políticas indicadas en el apartado anterior, cada cuántas semanas se cambiará la herramienta.
- Comprueba, para la mejor política conocida, si es óptima o no y, si no lo es, aporta una política más económica.



## Solución

a) Una iteración es lo que ocurre en una semana y determinamos el momento de cambio de transición como la observación que se hace el sábado. Los estados del sistema son N, S, U, D, R, que son los de la herramienta. Las probabilidades de cambio de estado durante la semana se obtienen dividiendo el número de observaciones por el total de la fila, y la remuneración son los costes totales durante la semana, que incluyen la herramienta repuesta (N o S) y los costes de producción. Las acciones posibles son poner una herramienta nueva (N), poner una seminueva (S) o no cambiar (NC).

En estas condiciones, la cadena de Markov con remuneración y decisión se ilustra en la tabla 7.15 (se incluyen solo las opciones que tienen sentido).

Tabla 7.15.  
Probabilidades,  
remuneraciones y  
ganancias  
esperadas de la  
cadena de Markov

| Estado | Decisión | Probabilidades (P) |     |     |     |     | Costes (R) |       |       |       |        | q      |
|--------|----------|--------------------|-----|-----|-----|-----|------------|-------|-------|-------|--------|--------|
|        |          | N                  | S   | U   | D   | R   | N          | S     | U     | D     | R      |        |
| N      | NC       | 0,4                | 0,3 | 0,3 | 0   | 0   | 0          | 100   | 250   | 500   | 30.000 | 105    |
| S      | N        | 0,4                | 0,3 | 0,3 | 0   | 0   | 2.500      | 2.600 | 2.750 | 3.000 | 32.500 | 2.605  |
| S      | NC       | 0                  | 0,4 | 0,5 | 0,1 | 0   | 0          | 200   | 350   | 600   | 30.100 | 315    |
| U      | N        | 0,4                | 0,3 | 0,3 | 0   | 0   | 2.500      | 2.600 | 2.750 | 3.000 | 32.500 | 2.605  |
| U      | S        | 0                  | 0,4 | 0,5 | 0,1 | 0   | 1.000      | 1.200 | 1.350 | 1.600 | 31.100 | 1.315  |
| U      | NC       | 0                  | 0   | 0,3 | 0,5 | 0,2 | 0          | 0     | 500   | 750   | 30.250 | 6.575  |
| D      | N        | 0,4                | 0,3 | 0,3 | 0   | 0   | 2.500      | 2.600 | 2.750 | 3.000 | 32.500 | 2.605  |
| D      | S        | 0                  | 0,4 | 0,5 | 0,1 | 0   | 1.000      | 1.200 | 1.350 | 1.600 | 31.100 | 1.315  |
| D      | NC       | 0                  | 0   | 0   | 0,5 | 0,5 | 0          | 0     | 0     | 1.000 | 30.500 | 15.750 |
| R      | N        | 0,4                | 0,3 | 0,3 | 0   | 0   | 2.500      | 2.600 | 2.750 | 3.000 | 32.500 | 2.605  |
| R      | S        | 0                  | 0,4 | 0,5 | 0,1 | 0   | 1.000      | 1.200 | 1.350 | 1.600 | 31.100 | 1.315  |

La política de cambiar por N cuando está en estado S o peor es [NC, N, N, N, N] y lleva a:

Tabla 7.16.  
Política de  
reemplazar la  
herramienta por una  
nueva (N) cuando  
está seminueva (S)

| Est. | Dec | Probs. (P) |     |     |   |   | Costes (R) |       |       |       |        |
|------|-----|------------|-----|-----|---|---|------------|-------|-------|-------|--------|
|      |     | N          | S   | U   | D | R | N          | S     | U     | D     | R      |
| N    | NC  | 0,4        | 0,3 | 0,3 | 0 | 0 | 0          | 100   | 250   | 500   | 30.000 |
| S    | N   | 0,4        | 0,3 | 0,3 | 0 | 0 | 2.500      | 2.600 | 2.750 | 3.000 | 32.500 |
| U    | N   | 0,4        | 0,3 | 0,3 | 0 | 0 | 2.500      | 2.600 | 2.750 | 3.000 | 32.500 |
| D    | N   | 0,4        | 0,3 | 0,3 | 0 | 0 | 2.500      | 2.600 | 2.750 | 3.000 | 32.500 |
| R    | N   | 0,4        | 0,3 | 0,3 | 0 | 0 | 2.500      | 2.600 | 2.750 | 3.000 | 32.500 |

Hay una única CF que incluye N, S, U; la matriz P ya tiene todas las filas iguales y, por tanto, es directamente la P\*. Para calcular g, basta con hacer:

$$P^* \cdot q = [0,4 \ 0,3 \ 0,3 \ 0 \ 0] [105 \ 2.605 \ 2.605 \ 2.605 \ 2.605]^T = 1.605 \text{ €/semana}$$



Cambiarla por seminueva (S) cuando está usada (U) o peor es [NC, NC, S, S, S] e implica:

| Est. | Dec | Probs. (P) |     |     |     |   | Costes (R) |       |       |       |        |
|------|-----|------------|-----|-----|-----|---|------------|-------|-------|-------|--------|
|      |     | N          | S   | U   | D   | R | N          | S     | U     | D     | R      |
| N    | NC  | 0,4        | 0,3 | 0,3 | 0   | 0 | 0          | 100   | 250   | 500   | 30.000 |
| S    | NC  | 0          | 0,4 | 0,5 | 0,1 | 0 | 0          | 200   | 350   | 600   | 30.100 |
| U    | S   | 0          | 0,4 | 0,5 | 0,1 | 0 | 1.000      | 1.200 | 1.350 | 1.600 | 31.100 |
| D    | S   | 0          | 0,4 | 0,5 | 0,1 | 0 | 1.000      | 1.200 | 1.350 | 1.600 | 31.100 |
| R    | S   | 0          | 0,4 | 0,5 | 0,1 | 0 | 1.000      | 1.200 | 1.350 | 1.600 | 31.100 |

Tabla 7.17.  
Política de reemplazar la herramienta por una seminueva (S) cuando está usada (U)

Ahora, la CF incluye los estados S, U y D, y la  $P^*$  es también igual en las filas de la CF en P, de donde:

$$g = P^* q = [0 \ 0,4 \ 0,5 \ 0,1 \ 0] [105 \ 315 \ 1.315 \ 1.315 \ 1.315]^T = 915 \text{ €/semana}$$

Resulta, pues, que es preferible la política de esperar a que la herramienta esté usada (U) o peor, y entonces cambiarla por una seminueva (S).

b) Para saber cada cuántas semanas se cambiará la herramienta, hemos de definir una remuneración que cuente las transiciones en que efectivamente se realice el cambio. Para ello, basta con definir un vector  $q'_1 = [0 \ 1 \ 1 \ 1 \ 1]^T$  en el caso de la primera política, o bien  $q'_2 = [0 \ 0 \ 1 \ 1 \ 1]^T$  para la segunda. Multiplicando por las respectivas  $P^*$ , tenemos:

$$g'_1 = [0,4 \ 0,3 \ 0,3 \ 0 \ 0] [0 \ 1 \ 1 \ 1 \ 1]^T = 0,6: \text{ se cambiará el } 60 \% \text{ de las semanas, es decir, cada } 1/0,6 \approx 1,67 \text{ semanas de media.}$$

$$g'_2 = [0 \ 0,4 \ 0,5 \ 0,1 \ 0] [0 \ 0 \ 1 \ 1 \ 1]^T = 0,6: \text{ se cambiará el } 60 \% \text{ de las semanas, es decir, cada } 1/0,6 \approx 1,67 \text{ semanas de media.}$$

c) Partimos de la mejor política conocida [NC, NC, S, S, S] y veamos si es óptima:

Resolvemos el sistema (es más fácil porque ya sabemos que  $g = 915$ ):

$$w_n + 915 = 105 + 0,4 w_n + 0,3 w_s + 0,3 w_u$$

$$w_s + 915 = 315 + 0 w_n + 0,4 w_s + 0,5 w_u + 0,1 w_d$$

$$w_u + 915 = 1.315 + 0 w_n + 0,4 w_s + 0,5 w_u + 0,1 w_d$$

$$w_d + 915 = 1.315 + 0 w_n + 0,4 w_s + 0,5 w_u + 0,1 w_d$$

$$w_r + 915 = 1.315 + 0 w_n + 0,4 w_s + 0,5 w_u + 0,1 w_d$$

y, haciendo  $w_s = 0$ , se obtiene:  $w_u = w_d = w_r = 1.000$ ;  $w_n = -850$



Comprobemos ahora la optimidad:

$w_n$ : no hace falta comprobar, porque solo hay una opción.

$$w_s: \min\{2.605 + 0,4 w_n + 0,3 w_s + 0,3 w_u; \mathbf{1.315 + 0,4 w_s + 0,5 w_u + 0,1 w_d}\} - 915 = 0 = w_s \rightarrow \text{OK}$$

$$w_u: \min\{2.605 + 0,4 w_n + 0,3 w_s + 0,3 w_u; \mathbf{1.315 + 0,4 w_s + 0,5 w_u + 0,1 w_d}\} - 915 = 1.000 = w_u \rightarrow \text{OK}$$

$$w_d: \min\{2.605 + 0,4 w_n + 0,3 w_s + 0,3 w_u; \mathbf{1.315 + 0,4 w_s + 0,5 w_u + 0,1 w_d}\} - 915 = 1.000 = w_d \rightarrow \text{OK}$$

$$w_t: \min\{2.605 + 0,4 w_n + 0,3 w_s + 0,3 w_u; \mathbf{1.315 + 0,4 w_s + 0,5 w_u + 0,1 w_d}\} - 915 = 1.000 = w_t \rightarrow \text{OK}$$

La política es óptima.

Nota: Este problema, como todos los problemas de cadenas de Markov con remuneración y decisión, puede resolverse mediante programación dinámica. En el volumen de esta misma colección dedicado a la programación dinámica, se halla la resolución de este mismo problema a través de la programación dinámica.

## 7.8 Cultivos rotativos

Una ONG que regenta una explotación agraria en un país del tercer mundo se encuentra cada año con el problema de decidir qué cultivo plantar, entre tres posibles: bambú, café o tabaco. Los resultados que se obtengan dependerán del tipo de cultivo que se decida, del cultivo anterior y de si las circunstancias climatológicas y económicas son favorables o desfavorables. En concreto, los beneficios esperados (en miles de €) son los que se indican en la tabla 7.18.

Tabla 7.18.  
Beneficios  
esperados por tipo  
de cultivo

| Sí el año anterior<br>se ha plantado... | ...y este año<br>se planta... | en año favorable,<br>se obtendrá | y si es desfavorable |
|-----------------------------------------|-------------------------------|----------------------------------|----------------------|
| Bambú                                   | Bambú                         | 20                               | 5                    |
| Bambú                                   | Café                          | 60                               | 20                   |
| Bambú                                   | Tabaco                        | 100                              | 30                   |
| Café                                    | Bambú                         | 85                               | 30                   |
| Café                                    | Café                          | 10                               | 10                   |
| Café                                    | Tabaco                        | 50                               | 40                   |
| Tabaco                                  | Bambú                         | 55                               | 20                   |
| Tabaco                                  | Café                          | 90                               | 15                   |
| Tabaco                                  | Tabaco                        | 30                               | 10                   |



Se ha observado que, estadísticamente, la alternancia de años favorables y desfavorables sigue una regularidad notable: después de un año favorable, solo el 30 % de las veces habrá otro también favorable; en cambio, después de un año desfavorable, el siguiente es favorable un 60 % de las veces.

- A largo plazo, ¿qué porcentaje de años serán favorables y desfavorables?
- Halla la política óptima a largo plazo.

### Solución

a) Para calcular la frecuencia de años favorables, el enunciado lo define como una cadena de Markov de 2 estados, cuya  $P^*$  es:

$$P = \begin{array}{c} \begin{array}{cc} & \begin{array}{cc} F & D \end{array} \\ \begin{array}{c} F \\ D \end{array} & \begin{bmatrix} 0,3 & 0,7 \\ 0,6 & 0,4 \end{bmatrix} \end{array} \Rightarrow P^* = \begin{array}{c} \begin{array}{cc} & \begin{array}{cc} F & D \end{array} \\ \begin{array}{c} F \\ D \end{array} & \begin{bmatrix} 6/13 & 7/13 \\ 6/13 & 7/13 \end{bmatrix} \end{array}$$

Es decir, serán favorables 6 de cada 13 años ( $\approx 46\%$ ) y el resto ( $\approx 54\%$ ), desfavorables.

b) El problema es aleatorio, de horizonte indeterminado y homogéneo en el tiempo, de modo que se puede modelizar como una cadena de Markov con decisión (v. tabla 7.19).

| Estado | Decisión | Matriz P |     |     |     |     |     | Matriz R |    |    |    |     |    | $q^T$ |
|--------|----------|----------|-----|-----|-----|-----|-----|----------|----|----|----|-----|----|-------|
|        |          | BF       | BD  | CF  | CD  | TF  | TD  | BF       | BD | CF | CD | TF  | TD |       |
| BF     | B        | 0,3      | 0,7 |     |     |     |     | 20       | 5  |    |    |     |    | 9,5   |
|        | C        |          |     | 0,3 | 0,7 |     |     |          |    | 60 | 20 |     |    | 32    |
|        | T        |          |     |     |     | 0,3 | 0,7 |          |    |    |    | 100 | 30 | 51    |
| BD     | B        | 0,6      | 0,4 |     |     |     |     | 20       | 5  |    |    |     |    | 14    |
|        | C        |          |     | 0,6 | 0,4 |     |     |          |    | 60 | 20 |     |    | 44    |
|        | T        |          |     |     |     | 0,6 | 0,4 |          |    |    |    | 100 | 30 | 72    |
| CF     | B        | 0,3      | 0,7 |     |     |     |     | 85       | 30 |    |    |     |    | 46,5  |
|        | C        |          |     | 0,3 | 0,7 |     |     |          |    | 10 | 10 |     |    | 10    |
|        | T        |          |     |     |     | 0,3 | 0,7 |          |    |    |    | 50  | 40 | 43    |
| CD     | B        | 0,6      | 0,4 |     |     |     |     | 85       | 30 |    |    |     |    | 63    |
|        | C        |          |     | 0,6 | 0,4 |     |     |          |    | 10 | 10 |     |    | 10    |
|        | T        |          |     |     |     | 0,6 | 0,4 |          |    |    |    | 50  | 40 | 46    |
| TF     | B        | 0,3      | 0,7 |     |     |     |     | 55       | 20 |    |    |     |    | 31    |
|        | C        |          |     | 0,3 | 0,7 |     |     |          |    | 90 | 15 |     |    | 38    |
|        | T        |          |     |     |     | 0,3 | 0,7 |          |    |    |    | 30  | 10 | 16    |
| TD     | B        | 0,6      | 0,4 |     |     |     |     | 55       | 20 |    |    |     |    | 41    |
|        | C        |          |     | 0,6 | 0,4 |     |     |          |    | 90 | 15 |     |    | 60    |
|        | T        |          |     |     |     | 0,6 | 0,4 |          |    |    |    | 30  | 10 | 22    |

Tabla 7.19.  
Cadena de Markov  
con decisión



Puesto que nos interesa la solución óptima a largo plazo, debemos establecer una política inicial para iterar en el espacio de las políticas. A falta de mejor información, tomaremos la que mejor ganancia dé el primer año, es decir [T, T, B, B, C, C], que implica un cultivo rotativo entre las tres posibilidades:  $B \rightarrow T \rightarrow C \rightarrow B$ , del cual resulta:

$$P^* = [6/39 \ 7/39 \ 6/39 \ 7/39 \ 6/39 \ 7/39]$$

Obsérvese que la suma de probabilidades de cada par de estados suma  $1/3$ , como corresponde a una cadena de Markov cíclica con  $p = 3$  (hacemos un mismo cultivo cada 3 años) y, dentro de cada subclase, la probabilidad de los dos estados es  $6/13$  y  $7/13$ , como corresponde a los años favorables y desfavorables, según se ha visto en el apartado *a*.

Para ver si la política es óptima, planteamos el sistema correspondiente y, haciendo  $w_{TD} = 0$ , resulta:

$$g = 55,8; w_{BF} = -10,45; w_{BD} = 4,87; w_{CF} = -8,99; w_{CD} = 2,91; w_{TF} = -18,93$$

política que el lector podrá comprobar que es óptima.

Nota final: Este problema, como todos los de cadenas de Markov con remuneración y decisión, también se puede resolver mediante la programación dinámica. En el volumen dedicado a ella dentro de esta misma colección se puede hallar la resolución correspondiente.







## Bibliografía

Avriel, M.; Golany, B. (ed.) (1996). *Mathematical Programming for Industrial Engineers*. Nueva York: Mordecai.

Bharucha-Reid, A. T. (1960). *Elements of the Theory of Markov Processes and Their Applications*. McGraw Hill Series in Probability and Statistics.

Cox, D. R.; Miller, A. T. (1977). *The Theory of Stochastic Processes*. Nueva York: CRC Press.

Ethier, S. N.; Kurtz, T. G. (1986). *Markov Processes: Characterization and convergence*. Nueva York: John Wiley & Sons.

Hillier, F.; Lieberman, G. (2010). *Introducción a la investigación de operaciones*. 9ª ed. McGraw Hill México.

Kijima, M. (1997). *Markov Processes for Stochastic Modelling*. Cambridge: Chapman & Hall.

Meyn, S. P.; Tweedie, R. L. (1993). *Markov Chains and Stochastic Stability*. Londres: Springer-Verlag.

Thierauf, R. J.; Grosse, R. A. (1980). *Toma de decisiones por medio de la investigación operativa*. Limusa.

Wagner, H. M. (1969). *Principles of Operations Research with Applications to Managerial Decisions*. Prentice-Hall.





## Glosario

### 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 semiergódica cíclica) o no tenerlas (cadena ergódica cíclica).

### Cadena de Markov ergódica

Cadena de Markov en que todas las filas de la matriz de probabilidades estacionarias de  $P^*$  son iguales, de manera que ninguno de los elementos de  $P^*$  es igual a 0. Se trata de una cadena con una sola clase final y sin clases de paso.

### Cadena de Markov mixta

Cadena de Markov no ergódica, con clases finales cíclicas y acíclicas.



### Cadena de Markov no ergódica

Cadena de Markov cuyas probabilidades estacionarias dependen del estado inicial, por lo cual las filas de  $P^*$  pueden ser diferentes entre sí. 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 que todas las filas de la matriz de probabilidades estacionarias de  $P^*$  son iguales, de modo que alguno de los elementos de  $P^*$  es igual a 0. Se trata de una cadena con una sola clase final y al menos una clase 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  $j$  es descendiente de  $i$ .

### Ciclo

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

**Clase final**

Conjunto de estados comunicados entre ellos de modo 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 período igual a 1.

**Clase final cíclica (o periódica)**

Clase final de una cadena de Markov de período superior a 1.

**Clase de paso**

Conjunto de estados comunicados entre ellos de modo 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 a remuneraciones para ese estado.

**Estado**

Variable descriptiva de la situación en que se encuentra el proceso en un momento dado. Puede ser una variable cuantitativa o cualitativa. También puede ser continua o discreta, aunque en este módulo solo 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  $j$  es descendiente de  $i$  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

Algoritmo para la obtención de la remuneración esperada a largo plazo para una política determinada (cadenas sin decisión).

Algoritmo para la obtención de la política óptima a largo plazo y la correspondiente remuneración esperada (cadenas con decisión).

### Iteración en el espacio de los estados

Algoritmo para la obtención de la remuneración esperada cuando faltan  $N$  etapas para finalizar el proceso, operando transición a transición. Utilizando la normalización, podemos llegar a obtener la remuneración esperada por transición a largo plazo después de operar para un número de transiciones bastante elevado. Si se trata de una cadena con decisión, permite hallar la política óptima a corto plazo.

### Longitud de un ciclo

Número mínimo de transiciones necesarias para regresar al estado  $i$  en un ciclo, si se ha iniciado el proceso en ese estado.

### Matriz de probabilidades de transición de un paso (P)

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

**Matriz de probabilidades estacionarias ( $P^*$ )**

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

**Matriz de remuneraciones ( $R$ )**

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 a largo plazo. Véase la sección 5.2.2 para una definición en su contexto.

**Período de una clase final**

Máximo común divisor de las longitudes de ciclo que pueden hallarse 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 la política, el 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 ese 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 transición siguiente, si en ese momento se encuentra en el estado  $i$ . En una cadena de Markov de orden 1, las probabilidades de transición de un paso no varían con el tiempo y no dependen de la evolución que se haya seguido en etapas anteriores.

### Probabilidades estacionarias ( $\pi_{ij}$ )

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 ha comenzado 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 posible. 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 0. 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.