

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS MATEMÁTICAS

Departamento de Estadística e Investigación Operativa I



**INFERENCIA ESTADÍSTICA EN PROBLEMAS DE
PROGRAMACIÓN LINEAL Y PROGRAMACIÓN DINÁMICA**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Tomás Prieto Rumeau

Bajo la dirección del doctor

Miguel Martín Díaz

Madrid, 2001

ISBN: 84-669-1806-X

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS MATEMÁTICAS

Departamento de Estadística e Investigación Operativa I

**INFERENCIA ESTADÍSTICA EN
PROBLEMAS DE PROGRAMACIÓN LINEAL
Y PROGRAMACIÓN DINÁMICA**

TOMÁS PRIETO RUMEAU

Madrid, noviembre de 2000

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS MATEMÁTICAS

Departamento de Estadística e Investigación Operativa I

**INFERENCIA ESTADÍSTICA EN
PROBLEMAS DE PROGRAMACIÓN LINEAL
Y PROGRAMACIÓN DINÁMICA**

TOMÁS PRIETO RUMEAU

Madrid, noviembre de 2000

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS MATEMÁTICAS
Departamento de Estadística e Investigación Operativa I

**Inferencia Estadística en Problemas de
Programación Lineal y Programación Dinámica**

TOMÁS PRIETO RUMEAU

Memoria para optar al grado de
Doctor en Ciencias Matemáticas
realizada bajo la dirección
del Dr. D. Miguel Martín Díaz.

**MIGUEL MARTÍN DÍAZ, PROFESOR DEL DEPARTAMENTO
DE ESTADÍSTICA E INVESTIGACIÓN OPERATIVA I DE LA
UNIVERSIDAD COMPLUTENSE DE MADRID**

CERTIFICA:

Que la presente memoria titulada:

**INFERENCIA ESTADÍSTICA EN PROBLEMAS DE
PROGRAMACIÓN LINEAL Y PROGRAMACIÓN DINÁMICA**

ha sido realizada bajo mi dirección por Tomás Prieto Rumeau, Licenciado en Ciencias Matemáticas, y constituye su Tesis para optar al grado de Doctor en Ciencias Matemáticas.

Y para que conste, en cumplimiento de la legislación vigente y a los efectos oportunos, firmo la presente en Madrid a veintiocho de noviembre de dos mil.

Agradezco encarecidamente a mi profesor, D. Miguel Martín Díaz, su dedicación, apoyo y consejos a lo largo de la realización de este trabajo.

También agradezco al Profesor D. Ricardo Vélez Ibarrola sus orientaciones y enseñanzas.

A mi abuelo Tomás Prieto Vilches

En esta memoria, titulada *Inferencia Estadística en Problemas de Programación Lineal y Programación Dinámica*, se estudian algoritmos estocásticos para la resolución de estos problemas. Se supone que algunos de los parámetros de los problemas considerados son desconocidos y que son observables variables aleatorias relacionadas con ellos, de ahí que los algoritmos propuestos sean aleatorios. Los problemas que se estudian son, en concreto, de programación lineal continua, de programación lineal entera y de parada óptima.

Pese a que el análisis de estos tres problemas se haga de forma independiente, los métodos utilizados responden a una filosofía común que es: hacer la mínima cantidad posible de estimaciones pero aprovechando al máximo la información que proporcionan estas estimaciones. Estos métodos se sitúan a medio camino entre dos procedimientos extremos: el primero que estima exhaustivamente el problema y después aproxima su solución; el otro consistente en estimar directamente la solución del problema. Se le puede reprochar al primer método que se produce una situación de sobre-estimación puesto que, presumiblemente, se están haciendo más estimaciones de las que son necesarias para cubrir los objetivos propuestos. El segundo procedimiento de estimación tiene el inconveniente de que se desaprovecha gran parte de la información que contiene la muestra. Este método de estimación se plantea además de forma prácticamente independiente del modelo considerado, de tal forma que no se tiene en cuenta la estructura del problema, obviando así sus propiedades específicas.

El método que se propone en este trabajo es intermedio en el sentido de que se estiman únicamente las cantidades de interés, adaptándose el algoritmo de estimación a la estructura particular del problema. Se trata pues de métodos directos de estimación desarrollados *ad hoc* y de forma específica para cada problema, superando en operatividad al primer método descrito

y en precisión al segundo. Esta filosofía de estimación es ampliable a todo problema estadístico que se plantee y ha sido aplicada a tres problemas de optimización concretos.

Otra problemática que ha sido recurrente a lo largo de este estudio ha sido la necesidad de alterar de forma infinitesimal los elementos del problema para poder llegar a resultados de convergencia. Estas alteraciones se han llevado a cabo bajo dos ópticas distintas, con afán de ilustrar varias posibilidades, que han sido: la modificación infinitesimal del algoritmo propuesto y la modificación infinitesimal del problema, técnicas que se han denominado de truncamiento y perturbación, respectivamente.

En el Capítulo 1 se estudia la estimación de la solución de un problema de programación lineal; en el Capítulo 2 se hace para un problema de programación lineal entera; el problema de parada óptima es el objeto de los capítulos 3 y 4. En el Capítulo 3 no se plantea un método directo de estimación sino que se analiza el método de máxima verosimilitud, basado en una estimación completa del problema de parada óptima. Es muy interesante la comparación de esta técnica con el método directo, tratado en el Capítulo 4, poniéndose en evidencia el tipo de resultados a los que ambas estimaciones llegan y comentándose las hipótesis que éstas necesitan. Se estudian simulaciones del problema de parada óptima en las que se vuelven a comparar, ahora desde un punto de vista empírico, los dos procedimientos descritos.

En todas las estimaciones se ha procurado proponer un algoritmo convergente casi seguramente, que identifica pues una solución del problema, y que además verifique un teorema central del límite de convergencia en distribución. Es preciso insistir en que el objetivo de esta memoria es menos el análisis concreto de los tres problemas de optimización tratados que la exposición de una metodología que se deriva de un enfoque general de la Inferencia Estadística, que es pues extrapolable a cualquier otro problema.

Notación

Definimos a continuación la notación que será utilizada en todos los capítulos del texto.

Por \mathbb{R} se designa el cuerpo de los números reales, por \mathbb{Z} los números enteros, por \mathbb{Z}^+ los números enteros mayores o iguales que cero, y por \mathbf{i} la unidad imaginaria. Se identifican los vectores reales con su representación en matriz columna, no haciéndose pues esa distinción. Las desigualdades:

$$\geq, \leq, > \text{ y } <$$

entre matrices reales de mismas dimensiones significan que se verifican esas mismas desigualdades término a término. Por $\mathbf{0}$ se denota una matriz cuyos términos sean todos nulos, sin especificar sus dimensiones, que quedarán aclaradas en el contexto. Si $A \geq \mathbf{0}$ se dirá que A es mayor o igual que $\mathbf{0}$, o que es no negativa; si $A > \mathbf{0}$ se dirá que A es positiva. La traspuesta de una matriz A se representa por A' . La matriz I es la matriz identidad.

Dado un conjunto A , la función I_A es la función indicatriz de A , que vale uno en A y cero fuera de A . Dados dos conjuntos A y B , la expresión $A \subseteq B$ significa que A está contenido en B , y $A \subset B$ que A está estrictamente contenido en B , es decir: $A \subseteq B$ y $A \neq B$.

En un espacio de probabilidad, el operador esperanza se representa por \mathbf{E} . Dada una familia \mathcal{X} de variables aleatorias, la σ -álgebra $\sigma(\mathcal{X})$ es la σ -álgebra generada por \mathcal{X} . Las convergencias: cs, p, d y L_p , designan las convergencias: casi segura, en probabilidad, en ley o distribución, y en media de orden p , respectivamente. Por \mathcal{N} se denota una distribución normal, ya sea ésta unidimensional o multi-dimensional. Se empleará la notación de Landau: o y O , con su significado habitual.

Índice

Introducción	1
Notación	3
1 Problema de Programación Lineal	9
1.1 Introducción	9
1.1.1 Objetivos del estudio	9
1.1.2 Notación	10
1.2 Algoritmo de estimación	12
1.2.1 Definiciones previas	12
1.2.2 Definición del algoritmo	14
1.3 Convergencia del algoritmo	21
1.3.1 Preliminares	21
1.3.2 Convergencia de los errores	31
1.3.3 Convergencia de la estimación	35
1.3.4 Velocidad de convergencia	38
1.4 Convergencia en distribución del algoritmo	44
1.4.1 Algoritmos γ -truncados	44
1.4.2 Convergencia en ley	49
1.4.3 Convergencia en distribución de los errores	65
1.4.4 Conclusiones	68
2 Problema de Programación Entera	71
2.1 Introducción	71
2.1.1 Objetivos del estudio	71
2.1.2 Presentación del problema	72
2.2 Algoritmo ED	76
2.2.1 Un método de perturbación	76
2.2.2 Estimación de una solución	82

2.3	Algoritmo EE	87
2.3.1	Definición del algoritmo	87
2.3.2	Convergencia del algoritmo	92
2.4	Conclusiones	110
3	Problema de Parada Óptima	113
3.1	Introducción	113
3.1.1	Objetivos	113
3.1.2	Notación	114
3.2	Estimación del problema	117
3.2.1	Estimación de las probabilidades de transición	117
3.2.2	Estimación del valor de una regla	123
3.2.3	Estimación de una regla óptima	134
3.2.4	Estimación del valor del problema	135
3.3	Métodos de perturbación	142
3.3.1	Perturbación de g	143
3.3.2	Perturbación de P	148
3.3.3	Comentarios y conclusiones	154
4	Método Directo en Parada Óptima	157
4.1	Introducción	157
4.1.1	Objetivos del estudio	157
4.1.2	Notación y definición del modelo	159
4.2	Estimación del valor de una regla de decisión	160
4.2.1	Introducción y definiciones	160
4.2.2	Convergencia del estimador. Primer método	166
4.2.3	Convergencia del estimador. Segundo método	181
4.2.4	Cambio de espacio muestral	188
4.3	Estimación de una regla óptima	196
4.3.1	Estudio de un ejemplo	196
4.3.2	Convergencia de la estimación	199
4.4	Estimación del valor del problema	201
4.4.1	Definición de los estimadores	201
4.4.2	Convergencia de la estimación	202
4.5	Métodos de perturbación	204
4.5.1	Introducción	204
4.5.2	Primer método de perturbación	206
4.5.3	Segundo método de perturbación	208
4.5.4	Tercer método de perturbación	210

ÍNDICE

7

4.6	Conclusiones	215
4.7	Simulaciones	216
4.7.1	Introducción	216
4.7.2	Simulación del método máximo-verosímil	218
4.7.3	Simulación del método directo	220
4.7.4	Comparación de los dos métodos	221
4.7.5	Apéndice de figuras	223
	Referencias	231

CAPÍTULO 1

Problema de Programación Lineal

1.1 Introducción

1.1.1 Objetivos del estudio

En este primer capítulo se estudia la estimación de un vértice óptimo en un problema de programación lineal. El objetivo de este capítulo es proponer un método de estimación de un vértice óptimo que no necesite estimar previamente la función objetivo. Se trata pues de un procedimiento directo, en el sentido de que está ideado para estimar únicamente un vértice óptimo, en contraposición con el método indirecto de estimación de la función objetivo, a partir de la cual se puede estimar la solución del problema, pero de forma más costosa y empleando una cantidad de información excesiva.

En la segunda sección, se definirá el procedimiento de estimación de un vértice óptimo. Dicho procedimiento se basa en la idea del algoritmo del símplex de Dantzig. El algoritmo que se presenta se va desplazando de un vértice del poliedro a otro. Cuando se sitúa en un vértice, estima, en base a toda la información disponible hasta el momento, cuál es el mejor vértice, de entre los contiguos al actual, desplazándose a ése que se ha estimado es mejor, o bien, no cambiando de vértice si no ha estimado que ninguno adyacente sea mejor.

En la tercera sección se prueba la convergencia de este algoritmo, en

un número finito de etapas, al conjunto de vértices óptimos del poliedro. Este procedimiento de estimación es una adaptación conforme del algoritmo del simplex, en cuya tabla algunos de los elementos son aleatorios y se van actualizando etapa a etapa. También se analiza la velocidad de convergencia del algoritmo.

En la cuarta sección del capítulo, se define otra clase de algoritmos de estimación. Esta clase depende de un parámetro real y contiene como caso particular el algoritmo estudiado previamente. Gran parte de los resultados probados para los estimadores anteriores siguen verificándose para estos nuevos estimadores. Los algoritmos de esta nueva clase tienen un comportamiento similar al ya estudiado, salvo que son más conservadores: necesitan una mayor evidencia de que un vértice contiguo al actual es mejor para desplazarse a él. Para valores del parámetro dentro de un cierto intervalo se demuestra la convergencia, siempre en un número finito de iteraciones, de estas estimaciones. Una vez probada la convergencia casi segura de los errores de estimación, se aborda el problema de la distribución asintótica de los errores. Bajo ciertas hipótesis adicionales, se prueba que la distribución límite del error es una variable aleatoria normal o una mixtura de variables aleatorias normales. Para el algoritmo inicialmente presentado, las condiciones para obtener una distribución límite son más restrictivas que las que hay que exigir a la clase ulteriormente introducida, puesto que hay que suponer que el vértice óptimo es único, hipótesis que no es necesaria para los otros algoritmos. Esto justifica el haber tenido en cuenta una familia más amplia de estimadores que la considerada en un principio.

Se ha hecho la hipótesis de que los errores de observación son variables aleatorias independientes. Los errores de estimación no tienen, en cambio, ninguna estructura probabilística conocida, aunque sí se pueden determinar a partir de diferencias de martingalas. Por ello, gran parte de los resultados que se demuestran están basados en la Teoría de Martingalas.

Los resultados sobre Programación Lineal, así como la notación, están tomados de BAZARAA et al.; los teoremas de Teoría de la Probabilidad que se citan están recogidos en IBARROLA et al.; los resultados de Teoría de Martingalas son de HALL et al..

1.1.2 Notación

Sea A una matriz real $m \times n$ de rango m y sea $b \in \mathbb{R}^m$. Se considera el poliedro P de \mathbb{R}^n :

$$P = \{x \in \mathbb{R}^n : Ax = b, x \geq \mathbf{0}\}.$$

Supondremos que P así definido es no vacío y acotado. Sea $c \in \mathbb{R}^n$ el vector de costes. Se quiere resolver el siguiente problema de programación lineal:

$$\begin{aligned} \text{mín } z(x) &= c'x \\ Ax &= b \\ x &\geq \mathbf{0}, \end{aligned}$$

suponiendo que el vector de costes c es desconocido. Admitiremos que existe algún vértice que no es óptimo.

Dada una base B , es decir, una sub-matriz $m \times m$ de A que sea invertible, y siendo N la sub-matriz $(n - m) \times m$ complementaria, se distinguen las variables según sean básicas o no, y análogamente con el vector de costes:

$$x = \begin{pmatrix} x_B \\ x_N \end{pmatrix} \quad y \quad c = \begin{pmatrix} c_B \\ c_N \end{pmatrix}.$$

Se designarán por I los índices de $\{1, \dots, n\}$ correspondientes a variables básicas y por J los correspondientes a variables no básicas. Se define entonces el vector z como:

$$z' = c'_B B^{-1} A,$$

de tal forma que el problema de programación lineal se puede escribir, referido a la base B :

$$\begin{aligned} \text{mín } c'_B B^{-1} b - \sum_{j \in J} (z_j - c_j) x_j \\ x_B = B^{-1} b - B^{-1} N x_N \\ x_B \geq \mathbf{0}, \quad x_N \geq \mathbf{0}. \end{aligned}$$

Se denotarán por y_{ij} los términos de la matriz $B^{-1} A$. Si la base B es factible, es decir, si $B^{-1} b \geq \mathbf{0}$, entonces la solución básica será óptima si para todo índice j : $z_j - c_j \leq 0$. Supondremos que ninguna base factible es degenerada, o lo que es lo mismo:

$$B^{-1} b \geq \mathbf{0} \implies B^{-1} b > \mathbf{0}.$$

Con esta suposición adicional, la condición necesaria y suficiente para que una base factible sea óptima es que:

$$\forall j \in J, \quad z_j - c_j \leq 0.$$

Se denotará por \mathcal{B} el conjunto de bases factibles.

A lo largo del capítulo se trabajará sobre un espacio de probabilidad:

$$(\Omega, \mathcal{F}, \text{Pr}),$$

y se utilizará el concepto de esperanza condicionada. Sabemos que la esperanza condicionada está unívocamente definida salvo en un conjunto de probabilidad cero. No haremos mención explícita de la propiedad de igualdad casi segura, sin que esto suponga pérdida de rigor.

1.2 Algoritmo de estimación

1.2.1 Definiciones previas

Para aplicar el algoritmo de estimación de un vértice óptimo se parte de un vértice que será la solución básica factible asociada a una base B . Puesto que la base B no es degenerada, es posible aumentar el valor de una variable no básica desde el nivel cero en el que está en la solución básica hasta un nivel positivo sin violar la factibilidad de la solución. Se puede tomar un número positivo η que permita esta operación para cualquier base factible y para cualquier dirección no básica de esa base.

PROPOSICIÓN 1.1: *Sea η el número real definido por:*

$$\eta = \min_{B \in \mathcal{B}} \min_{j \in N} \min_{\{i: y_{ij} > 0\}} \left\{ \frac{(B^{-1}b)_i}{y_{ij}} \right\}.$$

Se verifica que $\eta > 0$ y que, para cualquier $B \in \mathcal{B}$, la solución de $Ax = b$ que vale η en una variable no básica y 0 en todas las demás es factible.

DEMOSTRACIÓN: La propiedad de ser η positivo es consecuencia de suponer que todas las bases factibles son no degeneradas.

Para la segunda parte del enunciado, dada una base factible, si tomamos x_N igual a cero en todas sus coordenadas salvo en una: en la coordenada j , la que hacemos igual a η , se cumple entonces, para $i \in I$, que:

$$x_i = (B^{-1}b)_i - \eta y_{ij},$$

y se deduce directamente el resultado.

Q.E.D.

A continuación describimos el método de estimación que se empleará. Se supone dado un vértice inicial asociado a una base B_1 . Para cada una de las direcciones no básicas de ese vértice se hace una estimación del incremento del objetivo a lo largo de esa dirección, se estima cuál de esas direcciones es la de máximo decrecimiento en el objetivo— recuérdese que se está resolviendo un problema de minimización— y se considera una nueva base B_2 , asociada al vértice contiguo al primero al que se ha llegado desplazándose en el sentido la dirección escogida. En este nuevo vértice se repite el procedimiento anterior observando los incrementos del objetivo en sus direcciones no básicas. Además, las estimaciones que se hicieron en el primer vértice se pueden adaptar a esta segunda base de suerte que dispondríamos, para el segundo vértice, de dos estimaciones para cada dirección no básica del incremento del objetivo, lo que nos permite afinar las estimaciones para las direcciones del segundo vértice. Así, pasaríamos después a un tercer vértice para el cual consideraríamos las observaciones tomadas directamente en ese vértice y, además, las observaciones que se tomaron en los dos primeros vértices actualizadas a esta nueva base considerada, etc..

DEFINICIONES:

- *Sea:*

$$(\mathbf{Y}_k)_{k \geq 1}$$

una sucesión de variables aleatorias independientes $(n - m + 1)$ -dimensionales de media $\mathbf{0}$ y de segundos momentos uniformemente acotados.

- *Dada:*

$$(B_k)_{k \geq 1}$$

una sucesión de bases de \mathcal{B} , sean I_k y J_k los índices básicos y no básicos, respectivamente, de la base B_k , y sea:

$$B_k^{-1}A = (y_{ij}^{B_k})_{1 \leq i \leq m, 1 \leq j \leq n},$$

y en notación por columnas:

$$B_k^{-1}A = (y_j^{B_k})_{1 \leq j \leq n}.$$

- *Se define, para cada base B_k de la sucesión anterior:*

$$(z_j^{B_k})_{1 \leq j \leq n} = c'_{B_k} B_k^{-1}A,$$

verificándose que $z_j^{B_k} = c_j$ si $j \in I_k$.

Obsérvese que no se exige independencia entre las componentes de las variables \mathbf{Y}_k y que no hay ninguna hipótesis de que sean idénticamente distribuidas.

1.2.2 Definición del algoritmo

Primeros pasos

Sea $B_1 \in \mathcal{B}$ una base factible prefijada y sea P_1 el vértice del poliedro determinado por B_1 . Escribiendo en primer lugar las variables asociadas a los índices de I_1 y en segundo lugar las asociadas a los de J_1 , se tiene que:

$$P_1 = \begin{pmatrix} B_1^{-1}b \\ \mathbf{0} \end{pmatrix}. \quad (1.1)$$

En virtud de la Proposición 1.1, para cada $j \in J_1$, el punto:

$$P_{1,j} = \begin{pmatrix} B_1^{-1}b - \eta y_j^{B_1} \\ \eta e_j \end{pmatrix}, \quad (1.2)$$

donde e_j es un vector $(n-m)$ -dimensional cuyas componentes son todas nulas salvo la correspondiente a $j \in J_1$ que vale uno, es un punto del poliedro.

Consideramos la variable aleatoria $(n-m+1)$ -dimensional \mathbf{Y}_1 y la escribimos:

$$\mathbf{Y}_1 = (Y_1, (Y_{1,j})_{j \in J_1}).$$

Se define Y_1 como el error de observación del valor de la función objetivo en el vértice P_1 y, para cada $j \in J_1$, $Y_{1,j}$ es el error de observación del objetivo en el punto $P_{1,j}$. Se definen así las variables Z_1 y $Z_{1,j}$, para $j \in J_1$, como:

$$Z_1 = z(P_1) + Y_1 \quad \text{y} \quad Z_{1,j} = z(P_{1,j}) + Y_{1,j}. \quad (1.3)$$

Recordemos la interpretación del coeficiente $c_j - z_j$, llamado usualmente coste reducido: cuando se aumenta el valor de una variable no básica del nivel cero al nivel uno, el coste asociado al aumento de esa variable es c_j , pero esta variable no varía libremente puesto que al aumentar de valor, las demás variables han debido reajustar sus valores para seguir manteniendo la factibilidad de la solución. Así pues, el coste total de aumentar en uno el valor de la variable x_j es de $c_j - z_j$. Por tanto, si se trata de estimar el incremento que se produce en el vértice P_1 al desplazarnos por las direcciones no básicas, queremos estimar los coeficientes $z_j^{B_1} - c_j$, también llamados costes reducidos.

PROPOSICIÓN 1.2: Para cada $j \in J_1$, la variable aleatoria:

$$\frac{Z_1 - Z_{1,j}}{\eta}$$

es un estimador insesgado de $z_j^{B_1} - c_j$ de varianza uniformemente acotada en j .

DEMOSTRACIÓN: Se tiene, según (1.1) y (1.3), que:

$$Z_1 = z(P_1) + Y_1 = c'_{B_1} B_1^{-1} b + Y_1,$$

y que, para cada $j \in J_1$, recordando (1.2) y (1.3):

$$Z_{1,j} = z(P_{1,j}) + Y_{1,j} = c'_{B_1} B_1^{-1} b - \eta c'_{B_1} y_j^{B_1} + \eta c_j + Y_{1,j},$$

es decir que:

$$\begin{aligned} Z_{1,j} &= z(P_1) - \eta(z_j^{B_1} - c_j) + Y_{1,j} \\ &= Z_1 - \eta(z_j^{B_1} - c_j) + Y_{1,j} - Y_1, \end{aligned}$$

y por tanto:

$$\frac{Z_1 - Z_{1,j}}{\eta} = z_j^{B_1} - c_j + \frac{Y_1 - Y_{1,j}}{\eta},$$

de donde se deduce el resultado buscado, sin más que tener en cuenta las propiedades con que se definió \mathbf{Y}_1 . Q.E.D.

Obsérvese que las variables $Z_{1,j}$, para distintos valores de j , no son independientes, luego las estimaciones de las variaciones del objetivo en las distintas direcciones no son independientes.

En este problema no tratamos de estimar la función objetivo, sino que se pretende dar una estimación directa del vértice óptimo sin pasar por la estimación de la función de coste. Por ello, supondremos que no son observables las variables Z_1 y $Z_{1,j}$ ni las variables Y_1 y $Y_{1,j}$, para $j \in J_1$, y que únicamente son observables las variables $Z_1 - Z_{1,j}$, para $j \in J_1$.

Se define, para $j \in J_1$:

$$x_{1,j}^1 = \frac{Z_1 - Z_{1,j}}{\eta} = z_j^{B_1} - c_j + \frac{Y_1 - Y_{1,j}}{\eta}, \quad (1.4)$$

que es la estimación de $z_j^{B_1} - c_j$ hecha a partir de la primera observación muestral. El primer sub-índice de $x_{1,j}^1$ indica que estamos haciendo estimaciones para el primer vértice, el segundo sub-índice hace referencia a la j -ésima dirección y el super-índice indica que la estimación está basada en las primeras observaciones muestrales, es decir, en \mathbf{Y}_1 . Para tener una notación uniforme, se define, para $j \in I_1$:

$$x_{1,j}^1 = 0,$$

puesto que sabemos que los costes reducidos de las variables básicas son siempre iguales a cero. Esto sería equivalente a considerar que \mathbf{Y}_1 es una variable aleatoria $(n+1)$ -dimensional con $m+1$ de sus componentes iguales, es decir:

$$Y_1 = Y_{1,j}, \text{ para } j \in I_1,$$

y entonces $x_{1,j}^1$ vendría definido, para todo $j \in \{1, \dots, n\}$, por (1.4). Se define también, para $1 \leq j \leq n$:

$$T_{1,j} = x_{1,j}^1.$$

Por tanto, $T_{1,j}$ es la estimación, en el primer vértice, del incremento de la función objetivo al aumentar en una unidad el valor de la variable no básica x_j .

Se describe a continuación el segundo paso del algoritmo. Una vez conocidas las estimaciones de los incrementos del objetivo, se estima cuál es la dirección según la cual más decrece el objetivo, y siguiendo esa dirección se busca un nuevo vértice del poliedro. Para ello se considera la cantidad:

$$\max_{1 \leq j \leq n} T_{1,j}. \quad (1.5)$$

- Si $\max_{1 \leq j \leq n} T_{1,j} > 0$, y nótese que:

$$\arg \max_{1 \leq j \leq n} T_{1,j} \subseteq J_1,$$

tomamos:

$$j_1 \in \arg \max_{1 \leq j \leq n} T_{1,j}.$$

Se determina un índice $i_1 \in I_1$ que verifique:

$$i_1 \in \arg \max \left\{ \frac{(B_1^{-1}b)_i}{y_{ij_1}^{B_1}} \mid i \in I_1 : y_{ij_1}^{B_1} > 0 \right\},$$

donde $\{i \in I_1 : y_{ij_1}^{B_1} > 0\} \neq \emptyset$ puesto que el poliedro P es acotado.

Se define la segunda base $B_2 \in \mathcal{B}$, que es la misma que la base B_1 salvo que: la variable x_{j_1} pertenece a la base B_2 y la variable x_{i_1} sale de la base B_1 . Es decir:

$$I_2 = I_1 - \{i_1\} \cup \{j_1\} \quad \text{y} \quad J_2 = J_1 \cup \{i_1\} - \{j_1\}.$$

- Si $\max_{1 \leq j \leq n} T_{1,j} = 0$ entonces no se ha estimado que haya ninguna dirección a lo largo de la cual se mejore el objetivo. La base B_2 se toma igual a la base B_1 . También se puede decir que se toma $j_1 \in I_1$ tal que:

$$j_1 \in \arg \max_{1 \leq j \leq n} T_{1,j},$$

y, según el procedimiento descrito en el punto anterior, se tendría que i_1 sería igual a j_1 , de lo cual ya se deduciría que $B_2 = B_1$.

Obsérvese que, independientemente de que sea la expresión (1.5) positiva o nula, se podría definir j_1 como un índice cualquiera de $\{1, \dots, n\}$ tal que:

$$j_1 \in \arg \max_{1 \leq j \leq n} T_{1,j},$$

ya que se tiene, salvo casos de dependencia lineal de las componentes de \mathbf{Y}_1 , que:

$$\Pr \left\{ \arg \max_{1 \leq j \leq n} T_{1,j} \subseteq I_1 \mid \max_{1 \leq j \leq n} T_{1,j} = 0 \right\} = 1,$$

siempre que esta probabilidad condicionada esté bien definida. Hay que hacer notar que los índices i_1 y j_1 , así como la base B_2 , son aleatorios y son variables aleatorias medibles respecto de la sigma-álgebra generada por \mathbf{Y}_1 .

La base B_2 tiene asociado el vértice P_2 del poliedro. Ahora trataremos de estimar la dirección de J_2 de mayor descenso del objetivo. En primer lugar se actualizan las estimaciones de $z_j^{B_1} - c_j$ para convertirlas en estimaciones de $z_j^{B_2} - c_j$. Las fórmulas del cambio de base son:

$$z_j^{B_2} - c_j = z_j^{B_1} - c_j - \left(z_{j_1}^{B_1} - c_{j_1} \right) \frac{y_{i_1 j}^{B_1}}{y_{i_1 j_1}^{B_1}},$$

y aplicando estas mismas fórmulas a las estimaciones $x_{1,j}^1$ obtenemos las estimaciones $x_{2,j}^1$. La notación $x_{2,j}^1$ quiere decir que estamos haciendo estimaciones en la j -ésima dirección del segundo vértice basadas en las primeras

observaciones muestrales. Para todo $j \in \{1, \dots, n\}$, se define:

$$x_{2,j}^1 = x_{1,j}^1 - x_{1,j_1}^1 \frac{y_{i_1 j}^{B_1}}{y_{i_1 j_1}^{B_1}}.$$

Nótese que cuando $j \in I_2$ se cumple que $x_{2,j}^1 = 0$, y esta definición de la estimación adaptada al segundo vértice es coherente con el hecho de que los costes reducidos de las variables básicas de B_2 son iguales a cero.

Para hacer nuevas observaciones en el vértice P_2 escribimos:

$$\mathbf{Y}_2 = (Y_2, (Y_{2,j})_{j \in J_2}),$$

y consideramos, suponiendo que las variables de I_2 son las m primeras y las de J_2 las $n - m$ últimas, los puntos del poliedro:

$$P_2 = \begin{pmatrix} B_2^{-1}b \\ \mathbf{0} \end{pmatrix} \quad \text{y} \quad P_{2,j} = \begin{pmatrix} B_2^{-1}b - \eta y_j^{B_2} \\ \eta e_j \end{pmatrix}, \quad (1.6)$$

para $j \in J_2$. Las variables de \mathbf{Y}_2 son entonces errores de observación del valor de la función objetivo en los puntos P_2 y $P_{2,j}$, para $j \in J_2$, y se define:

$$Z_2 = z(P_2) + Y_2 \quad \text{y} \quad Z_{2,j} = z(P_{2,j}) + Y_{2,j}. \quad (1.7)$$

Obsérvese que, a diferencia de las observaciones para la primera base, los términos $z(P_2)$ y $z(P_{2,j})$ son aleatorios puesto que el vértice P_2 también lo es. Probaremos un resultado del tipo del de la Proposición 1.2, a saber, que las variables:

$$\frac{Z_2 - Z_{2,j}}{\eta},$$

para $j \in J_2$, son estimadores condicionalmente insesgados de los costes reducidos no básicos de la segunda base.

Se define, siguiendo la notación ya explicada, la estimación de la variación del objetivo en la j -ésima dirección del segundo vértice basada en las segundas observaciones muestrales: $x_{2,j}^2$, como:

$$x_{2,j}^2 = \frac{Z_2 - Z_{2,j}}{\eta}, \quad \text{si } j \in J_2;$$

$$x_{2,j}^2 = 0, \quad \text{si } j \in I_2.$$

De nuevo, esto sería equivalente a definir $Y_{2,j} = Y_2$ para $j \in I_2$, teniendo así una única definición para $x_{2,j}^2$.

PROPOSICIÓN 1.3: *Se verifica, para $1 \leq j \leq n$, que:*

$$\mathbf{E} \left(\frac{Z_2 - Z_{2,j}}{\eta} \middle| \sigma(\mathbf{Y}_1) \right) = z_j^{B_2} - c_j.$$

DEMOSTRACIÓN: La integrabilidad de las variables $Z_2 - Z_{2,j}$ se deduce de (1.7) por ser la función z acotada. Se tiene, según (1.7), que:

$$\begin{aligned} \mathbf{E} \left(\frac{Z_2 - Z_{2,j}}{\eta} \middle| \sigma(\mathbf{Y}_1) \right) &= \mathbf{E} \left(\frac{z(P_2) - z(P_{2,j})}{\eta} \middle| \sigma(\mathbf{Y}_1) \right) + \\ &+ \mathbf{E} \left(\frac{Y_2 - Y_{2,j}}{\eta} \middle| \sigma(\mathbf{Y}_1) \right). \end{aligned}$$

El primer sumando es igual a $(z(P_2) - z(P_{2,j}))/\eta$ puesto que los puntos aleatorios P_2 y $P_{2,j}$ son $\sigma(\mathbf{Y}_1)$ -medibles. El segundo sumando es igual a cero ya que las σ -álgebras generadas por \mathbf{Y}_1 y \mathbf{Y}_2 son independientes, según la definición de $(\mathbf{Y}_k)_{k \geq 1}$, y las variables de \mathbf{Y}_2 tienen media cero.

El resultado se deduce ya de las expresiones de (1.6). Q.E.D.

Disponemos ahora de dos estimaciones de las variaciones del objetivo en el vértice P_2 : unas obtenidas a partir de la observación \mathbf{Y}_1 y otras a partir de \mathbf{Y}_2 . Haciendo la media aritmética de estas dos estimaciones, definimos, para cualquier $j \in \{1, \dots, n\}$:

$$T_{2,j} = \frac{x_{2,j}^1 + x_{2,j}^2}{2}.$$

Para obtener el vértice P_3 se analiza, como en la iteración anterior, la cantidad:

$$\max_{1 \leq j \leq n} T_{2,j}.$$

Planteamiento general

Una vez explicada la mecánica del algoritmo para las dos primeras iteraciones, podemos plantear el algoritmo en su forma general. En la etapa k , se parte de una base aleatoria $B_k \in \mathcal{B}$ que tiene asociado un vértice P_k , y para ese vértice, basadas en las k primeras observaciones muestrales, disponemos de las estimaciones de las variaciones del objetivo a lo largo de las distintas direcciones:

$$x_{k,j}^1, x_{k,j}^2, \dots, x_{k,j}^k,$$

para cada $j \in \{1, \dots, n\}$, siendo estas estimaciones todas nulas cuando el índice j es tal que $j \in I_k$. Se define:

$$T_{k,j} = \frac{1}{k} \sum_{r=1}^k x_{k,j}^r,$$

y se cambia de vértice, según el algoritmo del símplex, haciendo salir de la base una variable j_k tal que:

$$j_k \in \arg \max_{1 \leq j \leq n} T_{k,j} \quad (1.8)$$

si $\max_{1 \leq j \leq n} T_{k,j} > 0$, o bien tal que:

$$j_k \in I_k \quad (1.9)$$

si $\max_{1 \leq j \leq n} T_{k,j} = 0$, y entrando en la base una variable i_k . Se obtiene así una base $B_{k+1} \in \mathcal{B}$ a la cual se asocia un vértice P_{k+1} . Para cada $r \in \{1, \dots, k\}$ se define:

$$x_{k+1,j}^r = x_{k,j}^r - x_{k,j_k}^r \frac{y_{i_k,j}^{B_k}}{y_{i_k,j_k}^{B_k}}. \quad (1.10)$$

En base a las observaciones \mathbf{Y}_{k+1} , se define, para $j \in J_{k+1}$:

$$x_{k+1,j}^{k+1} = \frac{Z_{k+1} - Z_{k+1,j}}{\eta}, \quad (1.11)$$

siendo:

$$Z_{k+1} = z(P_{k+1}) + Y_{k+1},$$

y:

$$Z_{k+1,j} = z(P_{k+1,j}) + Y_{k+1,j},$$

donde $P_{k+1,j}$ tiene una definición análoga a la dada en (1.6), cf. p.18. Para $j \in I_{k+1}$, se toma:

$$x_{k+1,j}^{k+1} = 0,$$

lo que equivale a definir $Y_{k+1,j} = Y_j$ cuando $j \in I_{k+1}$. Disponemos ya de todos los elementos necesarios para llevar a cabo una iteración más del algoritmo.

1.3 Convergencia del algoritmo

1.3.1 Preliminares

Resultados de medibilidad

A continuación probamos algunas propiedades de las variables definidas.

PROPOSICIÓN 1.4: *Para cada j tal que $1 \leq j \leq n$ se verifica:*

1. *para $k \geq 1$, la variable aleatoria $x_{k,j}^k$ es $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)$ -medible;*
2. *para $k \geq 2$, las variables aleatorias: $x_{k,j}^1, x_{k,j}^2, \dots, x_{k,j}^{k-1}$, son medibles respecto de $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})$.*

DEMOSTRACIÓN: El caso $k = 1$, para el primer apartado, es inmediato. Se hará la demostración por inducción, para $k \geq 2$, para los dos apartados a la vez.

Para $k = 2$, ya se vio esta propiedad cuando se expuso en detalle cuál era el segundo paso del algoritmo. Supongamos la propiedad verdadera para un cierto k .

Tomada la observación muestral k -ésima, se definió:

$$T_{k,j} = \frac{1}{k} \sum_{r=1}^k x_{k,j}^r.$$

Por la hipótesis de inducción, la variable $T_{k,j}$ es medible respecto de la σ -álgebra generada por las k primeras observaciones muestrales. La variable j_k , según (1.8) y (1.9), es pues $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)$ -medible y de ello se deduce que también son medibles respecto de la misma σ -álgebra: i_k y B_{k+1} .

Teniendo en cuenta lo anterior y que Y_{k+1} y $Y_{k+1,j}$ son los errores de observación del objetivo en los puntos:

$$\begin{pmatrix} B_{k+1}^{-1}b \\ \mathbf{0} \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} B_{k+1}^{-1}b - \eta y_j^{B_{k+1}} \\ \eta e_j \end{pmatrix},$$

se tiene que las variables:

$$Z_{k+1} = c'_{B_{k+1}} B_{k+1}^{-1}b + Y_{k+1}, \quad (1.12)$$

y:

$$Z_{k+1,j} = c'_{B_{k+1}} B_{k+1}^{-1} b - \eta \left(c'_{B_{k+1}} y_j^{B_{k+1}} - c_j \right) + Y_{k+1,j}, \quad (1.13)$$

son $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k+1})$ -medibles, y por tanto, según se definió en (1.11), también es medible respecto de esa σ -álgebra la variable $x_{k+1,j}^{k+1}$.

De las fórmulas (1.10) del cambio de base y de la hipótesis de inducción se deduce que son medibles respecto de $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)$ las variables:

$$x_{k+1,j}^1, x_{k+1,j}^2, \dots, x_{k+1,j}^k,$$

como queríamos demostrar.

Q.E.D.

Hagamos algún comentario sobre esta proposición.

Fijemos r con $1 \leq r < k$. Pese a que la variable $x_{k,j}^r$ proviene de las observaciones muestrales \mathbf{Y}_r , se trata de una actualización a la base B_k de la estimación hecha en la r -ésima etapa. Puesto que para definir la base B_k se utilizan todas las observaciones muestrales: $\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1}$, ya que esta base se determina a partir de los valores $T_{k-1,j}$ — véase (1.8)—, en dicha actualización de las r -ésimas observaciones han intervenido todas las demás observaciones muestrales, y de ahí el resultado al que se llega. Para la variable $x_{k,j}^k$, se trata de una estimación de $z_j^{B_k} - c_j$ con errores determinados por \mathbf{Y}_k y puesto que $z_j^{B_k} - c_j$ es $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})$ -medible, resulta que $x_{k,j}^k$ depende de todas las observaciones muestrales anteriores.

Definamos, en cada etapa k , los errores de estimación.

DEFINICIÓN: Para cada $j \in \{1, \dots, n\}$ y para cada $k \geq 1$, considerando $1 \leq r \leq k$, se definen las variables $w_{k,j}^r$ como:

$$w_{k,j}^r = x_{k,j}^r - (z_j^{B_k} - c_j).$$

En cada etapa k queremos estimar $z_j^{B_k} - c_j$ y la dificultad es que en cada iteración queremos estimar una cantidad aleatoria distinta. La variable $w_{k,j}^r$ representa el error de estimación.

Es interesante insistir en la siguiente:

OBSERVACIÓN 1.5: Para cada $j \in \{1, \dots, n\}$ y cada $k > 1$ se cumple que:

1. las variables: $w_{k,j}^1, \dots, w_{k,j}^{k-1}$, son $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})$ -medibles;
2. la variable $w_{k,j}^k$ es $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)$ -medible,

no pudiéndose debilitar ninguna de estas propiedades de medibilidad.

DEMOSTRACIÓN: Se sigue de forma inmediata de la demostración de la Proposición 1.4. Lo interesante es notar que no se puede debilitar el resultado enunciado.

En cuanto al primer apartado, los errores de estimación: $w_{k,j}^1, \dots, w_{k,j}^{k-1}$, dependen efectivamente de las $k - 1$ primeras observaciones muestrales sin poderse prescindir de ninguna de ellas. Para el segundo apartado, a partir de las ecuaciones (1.12) y (1.13), se tiene que:

$$x_{k,j}^k = \frac{Y_k - Y_{k,j}}{\eta} + z_j^{B_k} - c_j,$$

de donde se concluye que:

$$w_{k,j}^k = \frac{Y_k - Y_{k,j}}{\eta}. \quad (1.14)$$

Aparentemente, a la vista de la ecuación (1.14), se tendría que $w_{k,j}^k$ es $\sigma(\mathbf{Y}_k)$ -medible y no $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)$ -medible como dice el enunciado. Esto, sin embargo, no es así. Hemos considerado que la variable aleatoria $(n + 1)$ -dimensional \mathbf{Y}_k tenía $m + 1$ de sus componentes iguales a la variable llamada Y_1 . Una de las $n + 1$ componentes corresponde a la observación del objetivo en el vértice P_k , otras m componentes— iguales a esta primera— se corresponden con las m variables básicas de B_k , y $n - m$ componentes de \mathbf{Y}_k con las variables no básicas. Lo que ocurre es que la forma de asignar las componentes de \mathbf{Y}_k a las variables es aleatoria, dependiendo, en cada etapa, de cuáles de las variables son básicas y cuáles no básicas. En otras palabras, fijado $j \in \{1, \dots, n\}$, la variable $Y_{k,j}$ que se le asocia no es la misma según sea $j \in J_k$ o sea $j \in I_k$, ya que en el primer caso será una variable “libre” mientras que en el segundo coincidirá necesariamente con Y_k . Lo que sí es cierto es que esa asignación aleatoria de componentes de \mathbf{Y}_k a variables no básicas es $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})$ -medible puesto que la base B_k es medible respecto de $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})$.

Se podría pensar que este inconveniente se puede resolver de la siguiente manera: se considera que las $n + 1$ componentes de \mathbf{Y}_k no tienen ninguna relación entre sí, como sí ocurría anteriormente, y sólo se observan en cada etapa las variables $Y_{k,j}$ correspondientes a los índices no básicos. En este caso, la aleatoriedad residiría en que la expresión (1.14) sólo sería válida para $j \in J_k$ — en lugar de para un índice j cualquiera—, y vuelve a aparecer, como en la asignación de componentes ya comentada, una aleatoriedad que es $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})$ -medible.

La única solución posible es considerar la variable \mathbf{Y}_k como en el párrafo anterior y tomar la expresión (1.14) como una definición. En ese caso, la variable $w_{k,j}^k$ sí que sería $\sigma(\mathbf{Y}_k)$ -medible. El inconveniente es que no se cumpliría:

$$T_{k,j} = 0$$

para $j \in I_k$, lo cual no sería, desde un punto de vista conceptual, coherente.

Nosotros mantendremos la definición de \mathbf{Y}_k como hasta ahora porque el desarrollo posterior es similar aun con la limitación de que $w_{k,j}^k$ no sea $\sigma(\mathbf{Y}_k)$ -medible. Q.E.D.

Procedimiento de aproximación estocástica

Para cada $k \geq 1$ y para cada $j \in \{1, \dots, n\}$ se define $W_{k,j}$ como el error que se comete al estimar $z_j^{B_k} - c_j$ utilizando el promedio $T_{k,j}$, es decir:

$$W_{k,j} = T_{k,j} - (z_j^{B_k} - c_j) = \frac{1}{k} \sum_{r=1}^k w_{k,j}^r.$$

De la Observación 1.5 se deduce que la variable $W_{k+1,j}$ es $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k+1})$ -medible. El siguiente resultado nos permite determinar el error medio esperado que se comete en la estimación en la etapa $(k + 1)$ -ésima, siendo conocidas las k primeras observaciones muestrales.

TEOREMA 1.6: *Para todo $k \geq 1$ y todo $j \in \{1, \dots, n\}$ se cumple que:*

$$W_{k+1,j} = \frac{k}{k+1} (W_{k,j} - \xi_{k,j} W_{k,j_k}) + \frac{1}{k+1} \frac{Y_{k+1} - Y_{k+1,j}}{\eta}, \quad (1.15)$$

siendo:

$$\xi_{k,j} = \frac{y_{i_k j}^{B_k}}{y_{i_k j_k}^{B_k}}.$$

DEMOSTRACIÓN: Se tiene que:

$$T_{k+1,j} = \frac{1}{k+1} \sum_{r=1}^{k+1} x_{k+1,j}^r.$$

Distinguiendo $r \leq k$ y $r = k+1$, y aplicando la fórmula (1.10), cf. p.20:

$$\begin{aligned} T_{k+1,j} &= \frac{1}{k+1} \sum_{r=1}^k (x_{k,j}^r - \xi_{k,j} x_{k,j_k}^r) + \\ &+ \frac{1}{k+1} \left(z_j^{B_{k+1}} - c_j + \frac{Y_{k+1} - Y_{k+1,j}}{\eta} \right) \\ &= \frac{k}{k+1} (T_{k,j} - \xi_{k,j} T_{k,j_k}) + \\ &+ \frac{1}{k+1} \left(z_j^{B_{k+1}} - c_j + \frac{Y_{k+1} - Y_{k+1,j}}{\eta} \right). \end{aligned}$$

Se cumple pues que:

$$\begin{aligned} W_{k+1,j} &= T_{k+1,j} - (z_j^{B_{k+1}} - c_j) \\ &= \frac{k}{k+1} (T_{k,j} - (z_j^{B_k} - c_j) - \\ &- \xi_{k,j} (T_{k,j_k} - (z_{j_k}^{B_k} - c_{j_k}))) + \frac{1}{k+1} \frac{Y_{k+1} - Y_{k+1,j}}{\eta}, \end{aligned}$$

de donde se llega a la expresión buscada.

Q.E.D.

COROLARIO 1.7: *Con la notación del teorema anterior, se verifica que el error medio esperado tras k observaciones es:*

$$\mathbf{E}(W_{k+1,j} | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)) = \frac{k}{k+1} (W_{k,j} - \xi_{k,j} W_{k,j_k}),$$

para $k \geq 1$ y $j \in \{1, \dots, n\}$.

DEMOSTRACIÓN: Se deduce de la Proposición 1.4 que las variables aleatorias: $W_{k,j}$, W_{k,j_k} y $\xi_{k,j}$ son medibles respecto de la σ -álgebra:

$$\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k),$$

de donde se concluye de forma inmediata el resultado, puesto que:

$$\mathbf{E}(Y_{k+1} - Y_{k+1,j} | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)) = 0,$$

por la hipótesis de independencia.

Q.E.D.

Si pretendiésemos utilizar alguna técnica de aproximación estocástica para demostrar la convergencia a cero de los errores $(W_{k,j})_{1 \leq j \leq n}$, llamando:

$$\varepsilon_k = \frac{1}{k+1},$$

el Teorema 1.6 diría que:

$$W_{k+1,j} = W_{k,j} - \xi_{k,j} W_{k,j_k} - \varepsilon_k \left(W_{k,j} - \xi_{k,j} W_{k,j_k} - \frac{Y_{k+1} - Y_{k+1,j}}{\eta} \right). \quad (1.16)$$

Obsérvese que las variables $T_{k,j}$ y $W_{k,j}$ se han definido como medias aritméticas de otras variables. La ecuación (1.16) sugiere que si en lugar de hacer la media aritmética se hubiera ponderado con pesos ε_k y $1 - \varepsilon_k$ se hubiese llegado a una expresión del tipo de (1.16).

Este esquema no se adapta a la ecuación tipo de la aproximación estocástica ya que intervienen otros términos que los $W_{k,j}$, como por ejemplo los de la forma W_{k,j_k} . Se puede lograr una fórmula parecida a:

$$\theta_{k+1} = \theta_k + \varepsilon_k \beta_{k+1}, \quad (1.17)$$

pero no para escalares sino que para vectores. Se define W_k como:

$$W_k = (W_{k,1} \cdots W_{k,n})'.$$

Por A_k se denota una matriz aleatoria $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)$ -medible de dimensiones $n \times n$ que tiene todos sus términos iguales a cero salvo los de la columna j_k que son iguales a $\xi_{k,j}$, para $1 \leq j \leq n$. Se define $Y^{(k+1)}$ como la matriz $n \times 1$ cuyos términos son $(Y_{k+1} - Y_{k+1,j})/\eta$, y I como la matriz identidad de orden n . Se tiene, al escribir para cada j la expresión (1.16), que:

$$W_{k+1} = (1 - \varepsilon_k)(I - A_k)W_k + \varepsilon_k Y^{(k+1)},$$

o lo que es lo mismo:

$$W_{k+1} = (I - A_k)W_k - \varepsilon_k (I - A_k)W_k + \varepsilon_k Y^{(k+1)}, \quad (1.18)$$

y se observa que (1.18) no es ni siquiera una versión vectorial de la ecuación (1.17) por lo que, en principio, no parece claro cómo abordar este problema como uno de aproximación estocástica.

Diferencias de martingalas

Nos encaminamos a probar el resultado principal de convergencia a cero de los errores $W_{k,j}$.

Partimos de la siguiente observación: si de una cierta base $B \in \mathcal{B}$ conocemos los costes reducidos, entonces éstos permiten determinar los costes reducidos de cualquier otra base. El número de vértices del poliedro P es finito. Numerémoslos de 1 hasta K , o lo que es lo mismo, supongamos que:

$$\mathcal{B} = \{B^{(1)}, \dots, B^{(K)}\}.$$

En cada etapa de la iteración, el algoritmo identifica una base de \mathcal{B} : la base $B^{(t_k)}$:

$$B_k = B^{(t_k)}. \quad (1.19)$$

Fijada una variable x_j , siendo $j \in \{1, \dots, n\}$, recordemos que los errores de estimación de los costes reducidos en la k -ésima etapa son:

$$\frac{Y_k - Y_{k,j}}{\eta}, \text{ para } k \geq 1.$$

Tomemos una base cualquiera $B \in \mathcal{B}$. Para cada etapa k de la estimación y dada una variable x_j , se tiene una estimación directa, a través de \mathbf{Y}_k , de $z_j^{B_k} - c_j$:

$$z_j^{B_k} - c_j + \frac{Y_k - Y_{k,j}}{\eta}.$$

Mediante un número finito de cambios de base del tipo de (1.10), cf. p.20, se construye una estimación, basada en las k -ésimas observaciones muestrales, de $z_j^B - c_j$, estimación que, según la observación hecha anteriormente, depende únicamente de los coeficientes $z_j^{B_k} - c_j + (Y_k - Y_{k,j})/\eta$, de la base B_k y de la base B . Así, consideramos la nueva variable:

$$\tilde{x}_{B,j}^k,$$

que será la estimación hecha de $z_j^B - c_j$ a partir de la k -ésima observación muestral:

$$\tilde{x}_{B,j}^k = z_j^B - c_j + \tilde{w}_{B,j}^k,$$

donde el término $\tilde{w}_{B,j}^k$ representa el error. Nótese que estas variables son medibles respecto de $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)$, pues en efecto, el error de estimación en la base B_k es $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)$ -medible.

La demostración de la convergencia a cero de $W_{k,j}$ pasa en primer lugar por probar que también convergen a cero las variables:

$$\frac{1}{k} \sum_{r=1}^k \tilde{w}_{B,j}^r.$$

Para ello utilizaremos el siguiente:

LEMA 1.8: *Para cada base $B \in \mathcal{B}$, cada $k \geq 1$ y cada $j \in \{1, \dots, n\}$, se cumple:*

1. $\mathbf{E}|\tilde{w}_{B,j}^k| < +\infty$;
2. $\mathbf{E}(\tilde{w}_{B,j}^k | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})) = 0$, para $k > 1$;
3. $\sup_{k \geq 1} \mathbf{E}(\tilde{w}_{B,j}^k)^2 < +\infty$.

DEMOSTRACIÓN:

1. La observación esencial es que el término de error $\tilde{w}_{B,j}^k$ se escribirá como suma de una cantidad finita— dependiendo del número de cambios de base que se hayan realizado para llegar de B_k a B — de variables del tipo:

$$(Y_k - Y_{k,\cdot})/\eta, \tag{1.20}$$

multiplicadas por unas constantes que dependerán de las distintas matrices $(y_{\cdot,\cdot})$ que aparezcan a lo largo de los cambios de base. La justificación de esta observación es muy tediosa pero no encierra dificultad. Para convencerse basta con escribir dos cambios de base.

Es importante hacer notar que:

- se puede hacer que el número de cambios de base necesarios para pasar de B_k a B sea uniformemente acotado— esa cota superior no depende más que del propio poliedro P —, y supondremos que los cambios de base que hacemos para pasar de una base a otra son siempre los mismos;
- por tanto, el número de sumandos de variables del tipo (1.20) está uniformemente acotado— sólo depende esta cota de P y de la cota para el número de cambios de base del punto anterior—;

- las constantes que pueden acompañar a los términos del tipo de (1.20) son sólo una cantidad finita y podemos considerar su número uniformemente acotado, dependiendo esa cota únicamente de P .

Por tanto, y dado que todas las componentes de la variable \mathbf{Y}_k tienen integral finita, se concluye de estas observaciones que $\tilde{w}_{B,j}^k$ es integrable, como queríamos demostrar.

2. Solamente una vez que sabemos que la variable $\tilde{w}_{B,j}^k$ tiene media finita podemos hablar de esperanza condicionada. Si recordamos, como se ha descrito en el apartado anterior, la estructura de suma que tiene $\tilde{w}_{B,j}^k$, observamos que esa suma tiene un número aleatorio de sumandos— la aleatoriedad depende sólo de B_k —, siendo estos sumandos productos de cantidades aleatorias— que son función de B_k — multiplicadas por componentes aleatorias de la variable \mathbf{Y}_k , siendo esta aleatoriedad en las componentes de \mathbf{Y}_k dependiente de B_k . Podemos pues escribir que:

$$\tilde{w}_{B,j}^k = \sum_{C \in \mathcal{B}} I_{\{B_k=C\}} \tilde{w}_{B,j}^k,$$

y por tanto se puede expresar cada sumando de la forma:

$$I_{\{B_k=C\}} \tilde{w}_{B,j}^k = I_{\{B_k=C\}} \tilde{w}_{B,C,j}^k,$$

donde $\tilde{w}_{B,C,j}^k$ resulta de sustituir, al ser conocida la base B_k , las cantidades aleatorias de $w_{B,j}^k$ por las constantes que se obtienen cuando es $B_k = C$, y por tanto $\tilde{w}_{B,C,j}^k$ es una variable aleatoria $\sigma(\mathbf{Y}_k)$ -medible que es combinación lineal— no aleatoria— de componentes de \mathbf{Y}_k que, dado que todas ellas tienen media cero, también tiene esperanza igual a cero:

$$\mathbf{E} \tilde{w}_{B,C,j}^k = 0. \quad (1.21)$$

Se cumple pues que:

$$\mathbf{E}(\tilde{w}_{B,j}^k | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1}))$$

es igual a:

$$\sum_{C \in \mathcal{B}} \mathbf{E} \left(I_{\{B_k=C\}} \tilde{w}_{B,C,j}^k | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1}) \right),$$

que a su vez es igual a:

$$\sum_{C \in \mathcal{B}} I_{\{B_k=C\}} \mathbf{E}(\tilde{w}_{B,C,j}^k | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})),$$

ya que B_k es $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})$ -medible. En virtud de la independencia de las σ -álgebras $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})$ y $\sigma(\mathbf{Y}_k)$ se tiene, a partir de (1.21), que:

$$\mathbf{E}(\tilde{w}_{B,C,j}^k | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})) = \mathbf{E}\tilde{w}_{B,C,j}^k = 0,$$

deduciéndose entonces que:

$$\mathbf{E}(\tilde{w}_{B,j}^k | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})) = 0,$$

como queríamos demostrar. Supondremos que la asignación aleatoria de componentes de \mathbf{Y}_k es siempre la misma en cada base.

3. Para probar el último resultado, escribimos, según la notación anterior:

$$(\tilde{w}_{B,j}^k)^2 = \sum_{C \in \mathcal{B}} I_{\{B_k=C\}} (\tilde{w}_{B,j}^k)^2 = \sum_{C \in \mathcal{B}} I_{\{B_k=C\}} (\tilde{w}_{B,C,j}^k)^2. \quad (1.22)$$

Se tiene que $w_{B,C,j}^k$ es suma de una cantidad uniformemente acotada en k de constantes uniformemente acotadas en k de componentes de \mathbf{Y}_k . Por hipótesis, los segundos momentos de las componentes de \mathbf{Y}_k están uniformemente acotados, deduciéndose de (1.22) que:

$$\sup_{k \geq 1} \mathbf{E}(\tilde{w}_{B,j}^k)^2 < +\infty,$$

ya que el conjunto \mathcal{B} es finito.

Q.E.D.

Utilizamos a continuación una ley fuerte de los grandes números para diferencias de martingalas para probar la convergencia a cero de los promedios de $\tilde{w}_{B,j}^k$.

LEMA 1.9: *Para cada base $B \in \mathcal{B}$ y cada $j \in \{1, \dots, n\}$ se cumple que:*

$$\frac{1}{k} \sum_{r=1}^k \tilde{w}_{B,j}^r \xrightarrow{\text{cs}} 0 \quad \text{y} \quad \frac{1}{k} \sum_{r=1}^k \tilde{w}_{B,j}^r \xrightarrow{L_1} 0,$$

cuando k tiende a $+\infty$.

DEMOSTRACIÓN: El resultado de convergencia está basado en el siguiente teorema (cf. IBAROLA et al., p.407):

Dada una sucesión $(X_k)_{k \geq 1}$ de variables aleatorias tales que:

$$\mathbf{E}(X_k | \sigma(X_1, \dots, X_{k-1})) = 0, \quad (1.23)$$

para todo $k > 1$, si existe una sucesión $(C_k)_{k \geq 1}$ de números positivos que crezca a $+\infty$ tal que:

$$\sum_{k=1}^{\infty} \mathbf{E}X_k^2 / C_k^2 < +\infty, \quad (1.24)$$

entonces:

$$\frac{X_1 + \dots + X_k}{C_k}$$

converge a cero casi seguramente y en media de orden uno.

Fijemos una base $B \in \mathcal{B}$ y un índice j . Haciendo $\tilde{w}_{B,j}^k = X_k$, aplicamos este teorema. Puesto que:

$$\sigma(\tilde{w}_{B,j}^1, \dots, \tilde{w}_{B,j}^{k-1}) \subseteq \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1}),$$

se tiene que:

$$\mathbf{E} \left(\mathbf{E}(\tilde{w}_{B,j}^k | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})) | \sigma(\tilde{w}_{B,j}^1, \dots, \tilde{w}_{B,j}^{k-1}) \right)$$

es igual a:

$$\mathbf{E}(\tilde{w}_{B,j}^k | \sigma(\tilde{w}_{B,j}^1, \dots, \tilde{w}_{B,j}^{k-1})),$$

que resulta ser igual a cero, en virtud del apartado 2 del Lema 1.8. Se cumple pues la condición (1.23).

Para verificar la condición (1.24), tomamos $C_k = k$ y de la acotación probada en el apartado 3 del Lema 1.8 se concluye que se cumple dicha propiedad.

Se verifican las hipótesis del citado teorema, deduciéndose el resultado buscado. Q.E.D.

1.3.2 Convergencia de los errores

Estamos ya en condiciones de probar el resultado de convergencia de los errores $W_{k,j}$.

TEOREMA 1.10: *Fijado $j \in \{1, \dots, n\}$, se verifica que:*

$$W_{k,j} \xrightarrow{\text{cs}} 0 \quad \text{y} \quad W_{k,j} \xrightarrow{L_1} 0,$$

cuando $k \rightarrow +\infty$.

DEMOSTRACIÓN: Tanto en el Lema 1.8 como en el Lema 1.9, habíamos fijado una base B y para cada observación muestral \mathbf{Y}_k hecha en la base B_k , determinábamos las estimaciones de los costes reducidos para la base B , obteniendo así, para las k primeras observaciones muestrales, unos errores de estimación de los coeficientes $z_j^B - c_j$ iguales a:

$$\tilde{w}_{B,j}^1, \tilde{w}_{B,j}^2, \dots, \tilde{w}_{B,j}^k.$$

Los errores de estimación que se cometían en la base aleatoria B_k al estimar el coste reducido aleatorio $z_j^{B_k} - c_j$:

$$w_{k,j}^1, w_{k,j}^2, \dots, w_{k,j}^k,$$

no eran más que una sucesiva actualización a la base B_k de los errores cometidos en cada base B_r al tomar la observación muestral \mathbf{Y}_r , siendo el índice r : $1 \leq r \leq k$. Todos estos errores de estimación coinciden con los errores que hubiesen resultado de tomar, en el procedimiento descrito en los dos lemas anteriores: $B = B_k$. En otras palabras, y recordando la notación de (1.19), cf. p.27, lo que se cumple es que para todo $k \geq 1$ y para todo r que verifique que $1 \leq r \leq k$:

$$w_{k,j}^r = \tilde{w}_{B_k,j}^r = \tilde{w}_{B^{(t_k)},j}^r, \quad (1.25)$$

de donde:

$$W_{k,j} = \frac{1}{k} \sum_{r=1}^k w_{k,j}^r = \frac{1}{k} \sum_{r=1}^k \tilde{w}_{B^{(t_k)},j}^r.$$

Recapitulando, en el lema anterior se ha probado que para cada $B \in \mathcal{B}$, manteniendo fijo el índice $j \in \{1, \dots, n\}$, tenemos una sucesión $(\tilde{w}_{B,j}^r)_{r \geq 1}$ cuyos promedios tienden casi seguro y en media de orden uno a cero. Lo que establece (1.25) es que $W_{k,j}$ es igual, para cada k , a un promedio hasta el término k de alguna de las K sucesiones — K es el número de elementos de \mathcal{B} — del lema anterior. La idea de la demostración es que la sucesión $(W_{k,j})_{k \geq 1}$ es la unión de una cantidad finita de sub-sucesiones de:

$$\frac{1}{k} \sum_{r=1}^k \tilde{w}_{B,j}^r,$$

para las cuales ya tenemos un resultado de convergencia.

- Probemos la convergencia casi segura.

Para cada $B \in \mathcal{B}$, sea A_B un conjunto de probabilidad uno tal que para todas las trayectorias de observaciones muestrales $(\mathbf{Y}_k)_{k \geq 1}$ de A_B se cumpla que:

$$\lim_{k \rightarrow +\infty} \frac{1}{k} \sum_{r=1}^k \tilde{w}_{B,j}^r = 0.$$

El conjunto:

$$A = \bigcap_{B \in \mathcal{B}} A_B,$$

tiene probabilidad uno. Fijada una trayectoria cualquiera de A y un número $\varepsilon > 0$, para cada B existe un natural N_B tal que si $k > N_B$ entonces:

$$\left| \frac{1}{k} \sum_{r=1}^k \tilde{w}_{B,j}^r \right| < \varepsilon.$$

Tomando $k > \max_{B \in \mathcal{B}} N_B$, y siendo para la trayectoria fijada: $B_k = B^{(t_k)}$, se cumple que:

$$|W_{k,j}| = \left| \frac{1}{k} \sum_{r=1}^k w_{k,j}^r \right| = \left| \frac{1}{k} \sum_{r=1}^k \tilde{w}_{B^{(t_k)},j}^r \right| < \varepsilon,$$

quedando probada la convergencia casi segura a cero de $(W_{k,j})_{k \geq 1}$.

- Demostremos la convergencia en L_1 .

Tenemos que probar que:

$$\lim_{k \rightarrow +\infty} \mathbf{E}|W_{k,j}| = 0.$$

Basándonos en (1.25), podemos escribir que:

$$W_{k,j} = \sum_{B \in \mathcal{B}} I_{\{B=B_k\}} \left(\frac{1}{k} \sum_{r=1}^k \tilde{w}_{B,j}^r \right),$$

y por tanto:

$$|W_{k,j}| \leq \sum_{B \in \mathcal{B}} \left| \frac{1}{k} \sum_{r=1}^k \tilde{w}_{B,j}^r \right|.$$

Sabemos que cada sumando del segundo miembro de esta última desigualdad es tal que su esperanza tiende a cero, como se estableció en el Lema 1.9. Puesto que el número de sumandos es finito, también $\mathbf{E}|W_{k,j}|$ tiende a cero cuando $k \rightarrow +\infty$. Q.E.D.

Hagamos un comentario sobre esta demostración. Para probar la convergencia a cero de la sucesión $(W_{k,j})_{k \geq 1}$ hemos considerado esta sucesión dentro de un espacio más grande: hemos inmerso esta sucesión uni-dimensional en un espacio de dimensión igual al número de vértices del poliedro. En ese espacio de mayor dimensión hemos construido sucesiones que son diferencias de martingalas que verifican una ley de los grandes números. La sucesión original $(W_{k,j})_{k \geq 1}$ es tal que está formada por sub-sucesiones de sumas parciales de las diferencias de martingalas. De este hecho se ha concluido la convergencia de la propia sucesión $(W_{k,j})_{k \geq 1}$.

Se cumple el siguiente:

COROLARIO 1.11: *Bajo las hipótesis del Teorema 1.10:*

$$\max_{1 \leq j \leq n} |W_{k,j}| \xrightarrow{\text{cs}} 0 \quad \text{y} \quad \max_{1 \leq j \leq n} |W_{k,j}| \xrightarrow{L_1} 0,$$

cuando $k \rightarrow +\infty$.

DEMOSTRACIÓN: A partir del Teorema 1.10, es inmediata. Q.E.D.

La demostración del Teorema 1.10 ha pasado por probar que la sucesión cuyo k -ésimo término es:

$$\tilde{w}_{B,j}^1 + \cdots + \tilde{w}_{B,j}^k$$

es una martingala. No se ha podido hacer una demostración directa para la sucesión $(kW_{k,j})_{k \geq 1}$ ya que:

$$kW_{k,j} = w_{k,j}^1 + \cdots + w_{k,j}^k,$$

no es una martingala, como quedó probado en el Corolario 1.7. Se verifica que:

$$\mathbf{E}((k+1)W_{k+1,j} | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k))$$

es igual a:

$$\sum_{B \in \mathcal{B}} \mathbf{E}((k+1)W_{k+1,j} I_{\{B_{k+1}=B\}} | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)),$$

que es a su vez igual a:

$$\sum_{B \in \mathcal{B}} \mathbf{E}(I_{\{B_{k+1}=B\}} (\sum_{r=1}^{k+1} \tilde{w}_{B,j}^r) | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)) = \sum_{B \in \mathcal{B}} I_{\{B_{k+1}=B\}} (\sum_{r=1}^k \tilde{w}_{B,j}^r),$$

y para que esta expresión fuese igual a $kW_{k,j}$, tendría que ser:

$$\sum_{B \in \mathcal{B}} I_{\{B_{k+1}=B\}} \left(\sum_{r=1}^k \tilde{w}_{B,j}^r \right) = \sum_{B \in \mathcal{B}} I_{\{B_k=B\}} \left(\sum_{r=1}^k \tilde{w}_{B,j}^r \right),$$

lo que es falso. Esta es la razón por la que $(kW_{k,j})_{k \geq 1}$ no es una martingala. Tanto $\tilde{w}_{B,j}^r$ como $w_{k,j}^r$ son términos de error que están determinados por un cambio de base: de una base B_r a la base B en el primer caso y de una base B_r a la base B_k en el segundo. En el primer procedimiento, la base de llegada es fija— pudiéndose probar la propiedad de martingala— mientras que en el segundo es aleatoria y no es la misma en la etapa k -ésima que en la $(k+1)$ -ésima— no pudiéndose establecer la propiedad de martingala.

1.3.3 Convergencia de la estimación

Enunciamos el resultado principal de convergencia del algoritmo.

TEOREMA 1.12: *Para casi toda trayectoria de las observaciones existe un tamaño muestral a partir del cual el vértice del poliedro identificado por el algoritmo es óptimo.*

DEMOSTRACIÓN: Sea δ un número positivo definido mediante:

$$\delta < \min_{B \in \mathcal{B}} \min_{1 \leq j \leq n} \left\{ |z_j^B - c_j| : z_j^B - c_j \neq 0 \right\}.$$

Tenemos la garantía de que se puede tomar $\delta > 0$ puesto que existe algún vértice de P que no es óptimo. Para casi toda trayectoria de las observaciones, y aplicando el Corolario 1.11, existe un tamaño muestral N_0 tal que si $k > N_0$ entonces:

$$\max_{1 \leq j \leq n} |W_{k,j}| < \delta/2.$$

Tomando $k > N_0$, si para el vértice P_k se cumple que:

$$\max_{1 \leq j \leq n} T_{k,j} > \delta/2, \tag{1.26}$$

entonces el índice j_k es tal que $z_{j_k}^{B_k} - c_{j_k} > 0$. En efecto, como para todo índice $j \in \{1, \dots, n\}$:

$$z_j^{B_k} - c_j + W_{k,j} = T_{k,j},$$

si $T_{k,j_k} > \delta/2$, como $|W_{k,j_k}| < \delta/2$, resulta que $z_{j_k}^{B_k} - c_{j_k} > 0$. Al pasar, mediante el algoritmo del simplex, al siguiente vértice: P_{k+1} , resulta que:

$$z(P_{k+1}) < z(P_k),$$

puesto que las bases factibles son no degeneradas, es decir: ha habido una mejora en el valor de la función objetivo, pero no ya en una estimación del valor de la función objetivo, sino que en el verdadero valor de la función objetivo. Si, en cambio, para el vértice P_k , tomando $k > N_0$, se cumple que:

$$\max_{1 \leq j \leq n} T_{k,j} \leq \delta/2,$$

entonces $z(P_{k+1}) = z(P_k)$. Se concluye que, a partir del índice N_0 , la sucesión $(z(P_k))_{k \geq 1}$ es monótona no creciente. Puesto que hay una cantidad finita de vértices, la condición (1.26) sólo puede darse una cantidad finita de veces.

Existirá por tanto $N_1 > N_0$ tal que si $k > N_1$, entonces:

$$\max_{1 \leq j \leq n} T_{k,j} \leq \delta/2.$$

Eso significa que para todo $k > N_1$ y para todo $j \in \{1, \dots, n\}$:

$$z_j^{B_k} - c_j \leq 0, \quad (1.27)$$

ya que si fuese $z_j^{B_k} - c_j > 0$, se cumpliría, por la definición de δ , que:

$$z_j^{B_k} - c_j > \delta,$$

y puesto que $|W_{k,j}| < \delta/2$, necesariamente:

$$T_{k,j_k} = \max_{1 \leq j \leq n} T_{k,j} > \delta/2,$$

lo cual es imposible.

Al verificarse, para todo $k > N_1$ y para todo j , la condición (1.27), se deduce que los vértices P_k , para $k > N_1$, son óptimos. Q.E.D.

Comentemos esta demostración. Una base $B \in \mathcal{B}$ es óptima si y sólo si para todo $j \in \{1, \dots, n\}$ se cumple:

$$z_j^B - c_j \leq 0,$$

y no es óptima cuando existe $j \in \{1, \dots, n\}$ tal que:

$$z_j^B - c_j > 0.$$

Por tanto, el criterio de optimalidad depende únicamente de que haya algún coste reducido estrictamente mayor que cero, es decir: más que el valor concreto del coste reducido, nos interesa saber si es positivo, y esa propiedad se puede dilucidar, con probabilidad uno, en un número finito de etapas, hallando un vértice óptimo en un número finito de etapas. El algoritmo estudiado tiene la propiedad de que en cada etapa se estima un coste reducido distinto, en lugar de detenerse en un vértice hasta estimar con cierta consistencia los costes reducidos de ese vértice, y eso ha añadido complicaciones a las demostraciones.

La justificación de la convergencia finita del algoritmo tiene, aunque en un contexto distinto, la siguiente justificación intuitiva: si para una cierta función de coste, un vértice del poliedro es óptimo, entonces también lo será para funciones objetivo “cercanas” a la verdadera función objetivo. Por tanto, una estimación lo bastante fina de la función objetivo ya nos asegura que el vértice óptimo para la estimación es óptimo.

Vamos a definir un criterio de parada del algoritmo: fijado un número de etapas R , detenemos el algoritmo en cuanto se pase dos veces, en etapas posteriores a la R -ésima, por un mismo vértice. Dado que el número de vértices es finito, siguiendo esta regla, el algoritmo se detiene siempre. Se verifica la siguiente:

PROPOSICIÓN 1.13: *Para cada $R \geq 1$, se define $\Delta(R)$ como el conjunto de trayectorias tales que el vértice identificado por el algoritmo en el momento de parar es óptimo. Se cumple:*

$$\lim_{R \rightarrow +\infty} \Pr \{ \Delta(R) \} = 1.$$

DEMOSTRACIÓN: Retomamos la notación utilizada en la demostración del Teorema 1.12. Se probó que para casi toda trayectoria existía un tamaño N_1 a partir del cual los vértices identificados por el algoritmo eran todos óptimos. Tomando:

$$R > N_1$$

tenemos la garantía de que al detener el algoritmo, lo haremos en un vértice óptimo. Se deduce de esto que el conjunto:

$$\liminf_{R \rightarrow +\infty} \Delta(R)$$

contiene un conjunto de probabilidad uno, concluyéndose el resultado.

Q.E.D.

Más adelante, en la Proposición 1.18, se demostrará un resultado sobre la probabilidad de que el algoritmo estudiado repita vértice en dos iteraciones consecutivas. El resultado que se prueba es que la probabilidad de que los vértices identificados en las iteraciones k -ésima y $(k+1)$ -ésima sean iguales y no sean óptimos es un infinitésimo de orden:

$$O\left(\frac{1}{k}\right).$$

Este resultado tiene especial interés cuando exista un único vértice óptimo ya que, en este caso, el algoritmo, con probabilidad uno, a partir de un cierto tamaño muestral ya no se desplaza de vértice.

El resultado de la Proposición 1.18 a la que se hace referencia prueba este hecho para una clase más amplia de algoritmos. Esta nueva clase de algoritmos va a ser definida en la cuarta sección del capítulo.

1.3.4 Velocidad de convergencia

En este apartado se analiza la velocidad de convergencia del algoritmo desarrollado. Supondremos que la solución óptima del problema es única.

LEMA 1.14: *Fijados una base $B \in \mathcal{B}$, dos enteros: $s \geq 0$ y $k \geq 1$, un índice $j \in \{1, \dots, n\}$, y un número $\varepsilon > 0$, se cumple que:*

$$\Pr \left\{ \max_{0 \leq i \leq s} \frac{1}{k+i} \left| \sum_{r=1}^{k+i} \tilde{w}_{B,j}^r \right| \geq \varepsilon \right\} \leq \frac{A}{k},$$

donde la constante $A = A(B, s, j, \varepsilon)$ depende de B , s , j y ε , pero no de k .

DEMOSTRACIÓN: Recordemos el resultado del Lema 1.8. La sucesión:

$$(\tilde{w}_{B,j}^r)_{r \geq 1} \quad (1.28)$$

es una diferencia de martingalas de media cero para la filtración:

$$(\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_r))_{r \geq 1},$$

y es tal que las esperanzas de $(\tilde{w}_{B,j}^r)^2$ están uniformemente acotadas en r .

Observemos que:

$$\left\{ \max_{0 \leq i \leq s} \frac{1}{k+i} \left| \sum_{r=1}^{k+i} \tilde{w}_{B,j}^r \right| \geq \varepsilon \right\} \subseteq \left\{ \max_{0 \leq i \leq s} \left| \sum_{r=1}^{k+i} \tilde{w}_{B,j}^r \right| \geq k\varepsilon \right\}. \quad (1.29)$$

En efecto, si para algún $i \in \{0, \dots, s\}$ se tiene que:

$$\frac{1}{k+i} \left| \sum_{r=1}^{k+i} \tilde{w}_{B,j}^r \right| \geq \varepsilon,$$

entonces se cumple que:

$$\left| \sum_{r=1}^{k+i} \tilde{w}_{B,j}^r \right| \geq (k+i)\varepsilon \geq k\varepsilon,$$

ya que $i \geq 0$, quedando probado el contenido (1.29).

La sucesión finita:

$$\sum_{r=1}^k \tilde{w}_{B,j}^r, \sum_{r=1}^{k+1} \tilde{w}_{B,j}^r, \dots, \sum_{r=1}^{k+s} \tilde{w}_{B,j}^r$$

es una martingala para la filtración:

$$(\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k+i}))_{0 \leq i \leq s}.$$

Aplicamos la generalización a martingalas de la desigualdad de Kolmogorov, véase HALL et al., Corolario 2.1, p.14, para concluir que:

$$\Pr \left\{ \max_{0 \leq i \leq s} \left| \sum_{r=1}^{k+i} \tilde{w}_{B,j}^r \right| \geq k\varepsilon \right\} \leq \frac{1}{k^2 \varepsilon^2} \mathbf{E} \left(\sum_{r=1}^{k+s} \tilde{w}_{B,j}^r \right)^2.$$

Puesto que los términos de (1.28) son variables aleatorias ortogonales, es decir:

$$\mathbf{E}(\tilde{w}_{B,j}^{r_1} \tilde{w}_{B,j}^{r_2}) = 0,$$

para $r_1 \neq r_2$, por tratarse de una diferencia de martingalas, se cumple que:

$$\Pr \left\{ \max_{0 \leq i \leq s} \left| \sum_{r=1}^{k+i} \tilde{w}_{B,j}^r \right| \geq k\varepsilon \right\} \leq \frac{1}{k^2\varepsilon^2} \sum_{r=1}^{k+s} \mathbf{E}(\tilde{w}_{B,j}^r)^2,$$

y por la propiedad de acotación de los segundos momentos de las variables de (1.28), resulta que, para alguna constante C :

$$\frac{1}{k^2\varepsilon^2} \sum_{r=1}^{k+s} \mathbf{E}(\tilde{w}_{B,j}^r)^2 \leq \frac{k+s}{k^2\varepsilon^2} \cdot C \leq \frac{2s+1}{k\varepsilon^2} \cdot C = \frac{A}{k},$$

deduciéndose el resultado del lema.

Q.E.D.

Los errores de estimación que se cometen en el j -ésimo coste reducido de la base B_k : $W_{k,j}$, verifican el siguiente resultado.

LEMA 1.15: *Fijados dos enteros: $s \geq 0$ y $k \geq 1$, y un número $\varepsilon > 0$, se cumple:*

$$\Pr \left\{ \max_{0 \leq i \leq s} \max_{1 \leq j \leq n} |W_{k+i,j}| \geq \varepsilon \right\} \leq \frac{A}{k},$$

no dependiendo la constante $A = A(s, \varepsilon)$ más que de s y ε .

DEMOSTRACIÓN: Fijemos una trayectoria de:

$$\left\{ \max_{0 \leq i \leq s} \max_{1 \leq j \leq n} |W_{k+i,j}| \geq \varepsilon \right\}.$$

Existen entonces un entero i y un índice j tales que:

$$|W_{k+i,j}| \geq \varepsilon.$$

Si para la trayectoria fijada la base que el algoritmo identifica en la $(k+i)$ -ésima etapa del muestreo es: $B_{k+i} = B$, entonces:

$$W_{k+i,j} = \frac{1}{k+i} \sum_{r=1}^{k+i} \tilde{w}_{B,j}^r,$$

luego:

$$\left| \frac{1}{k+i} \sum_{r=1}^{k+i} \tilde{w}_{B,j}^r \right| \geq \varepsilon.$$

Hemos probado pues que:

$$\left\{ \max_{0 \leq i \leq s} \max_{1 \leq j \leq n} |W_{k+i,j}| \geq \varepsilon \right\} \subseteq \left\{ \max_{0 \leq i \leq s} \max_{B \in \mathcal{B}} \max_{1 \leq j \leq n} \left| \frac{1}{k+i} \sum_{r=1}^{k+i} \tilde{w}_{B,j}^r \right| \geq \varepsilon \right\}.$$

Por tanto, se cumple que:

$$\begin{aligned} \Pr \left\{ \max_{0 \leq i \leq s} \max_{1 \leq j \leq n} |W_{k+i,j}| \geq \varepsilon \right\} &\leq \sum_{B \in \mathcal{B}} \sum_{j=1}^n \Pr \left\{ \max_{0 \leq i \leq s} \left| \frac{1}{k+i} \sum_{r=1}^{k+i} \tilde{w}_{B,j}^r \right| \geq \varepsilon \right\} \\ &\leq \sum_{B \in \mathcal{B}} \sum_{j=1}^n \frac{A(B, s, j, \varepsilon)}{k} \\ &= \frac{A(s, \varepsilon)}{k}, \end{aligned}$$

donde, en la segunda desigualdad, se ha aplicado el Lema 1.14, quedando probado el resultado buscado. Q.E.D.

Se define el número $\delta > 0$ de la misma manera que en la demostración del Teorema 1.12, es decir:

$$\delta < \min_{B \in \mathcal{B}} \min_{1 \leq j \leq n} \left\{ |z_j^B - c_j| : z_j^B - c_j \neq 0 \right\}.$$

Vamos a aplicar el resultado del Lema 1.15 tomando:

$$\varepsilon < \frac{\delta}{2}. \tag{1.30}$$

Recordemos que $K \geq 2$ es el número de bases factibles del poliedro P . Probamos a continuación el resultado sobre la velocidad de convergencia del algoritmo.

TEOREMA 1.16: *La probabilidad de que la base identificada en la k -ésima etapa: B_k , no sea óptima es un infinitésimo de orden:*

$$O\left(\frac{1}{k}\right).$$

DEMOSTRACIÓN: Tomamos ε como se indicó en (1.30) y se define:

$$s = K - 1.$$

Sea $k \geq 1$ y fijemos una trayectoria de:

$$\left\{ \max_{0 \leq i < s} \max_{1 \leq j \leq n} |W_{k+i,j}| < \varepsilon \right\}. \quad (1.31)$$

Probemos que para la trayectoria fijada, la base B_{k+s} es óptima, es decir, si B^* es la base óptima de P , queremos demostrar que:

$$\left\{ \max_{0 \leq i < s} \max_{1 \leq j \leq n} |W_{k+i,j}| < \varepsilon \right\} \subseteq \{B_{k+s} = B^*\}.$$

Supongamos que alguna de las bases: $B_k, B_{k+1}, \dots, B_{k+s-1}$ es óptima, por ejemplo: $B_{k+i} = B^*$, para algún $i \in \{0, \dots, s-1\}$. Para los índices no básicos de dicha base se tiene:

$$z_j^{B^*} - c_j < -\delta < -2\varepsilon,$$

y puesto que:

$$|W_{k+i,j}| < \varepsilon,$$

se cumple que:

$$T_{k+i,j} < -\varepsilon < 0,$$

siendo esto cierto para cualquier índice no básico de B^* . Por tanto:

$$\max_{1 \leq j \leq n} T_{k+i,j} = 0,$$

y en consecuencia, el algoritmo repite vértice en la siguiente iteración, es decir:

$$B_{k+i+1} = B_{k+i} = B^*.$$

Procediendo así recursivamente, se tiene que la base B_{k+s} es óptima. En conclusión, si alguna de las bases anteriores a la $(k+s)$ -ésima es óptima, entonces también lo es B_{k+s} .

En el caso contrario, es decir, si ninguna de las bases:

$$B_k, B_{k+1}, \dots, B_{k+s-1}$$

es óptima, el análisis es el siguiente. Si un índice no básico j es tal que:

$$z_j^{B^k} - c_j \leq 0,$$

entonces se tiene que:

$$T_{k,j} < \varepsilon,$$

ya que $|W_{k,j}| < \varepsilon$. Si el índice no básico j verifica, en cambio, que:

$$z_j^{B_k} - c_j > 0,$$

entonces:

$$z_j^{B_k} - c_j > \delta > 2\varepsilon,$$

cumpléndose, por el mismo motivo que anteriormente, que:

$$T_{k,j} > \varepsilon.$$

Por ser la base B_k no óptima, se garantiza que existe un índice j en estas condiciones. En consecuencia:

$$\max_{1 \leq j \leq n} T_{k,j} > 0,$$

y cualquier índice j_k en el que se alcance el máximo es tal que:

$$z_{j_k}^{B_k} - c_{j_k} > 0,$$

y por tanto, en la siguiente base: B_{k+1} , el valor de la (verdadera) función objetivo es menor que en B_k . Como B_{k+1} tampoco es óptima, entonces la función objetivo en B_{k+2} es menor que en B_{k+1} , y así sucesivamente. Finalmente, la base B_{k+s-1} es estrictamente mejorada por la base B_{k+s} . Puesto que el número total de bases factibles de P es igual a:

$$K = s + 1,$$

y hemos encontrado K bases, cada una mejorando estrictamente la anterior, se cumple necesariamente que la base B_{k+s} es óptima.

En conclusión, hemos probado que para cualquier trayectoria de (1.31) la base B_{k+s} es óptima. Por tanto, se verifica que:

$$\Pr\{B_{k+s} \neq B^*\} \leq \Pr\left\{\max_{0 \leq i < s} \max_{1 \leq j \leq n} |W_{k+i,j}| \geq \varepsilon\right\} \leq \frac{A(s-1, \varepsilon)}{k},$$

como resulta de aplicar el Lema 1.15 para ε y $s-1$.

Para terminar con la demostración, tomemos $k > s$. Se tiene que:

$$\Pr\{B_k \neq B^*\} = \Pr\{B_{(k-s)+s} \neq B^*\} \leq \frac{A(s-1, \varepsilon)}{k-s} \leq \frac{(s+1)A(s-1, \varepsilon)}{k},$$

quedando probado el resultado buscado, ya que esta desigualdad se cumple salvo quizá para:

$$k \leq s,$$

que es una cantidad finita de casos.

Q.E.D.

Por tanto, la probabilidad de que, tras observar k muestras, no hayamos identificado la base óptima tiende a cero al menos tan rápidamente como la sucesión $(1/k)$.

1.4 Convergencia en distribución del algoritmo

Todos los resultados anteriores se podían haber probado suponiendo que la sucesión $(\mathbf{Y}_k)_{k \geq 1}$ era una diferencia de martingalas, es decir:

$$\mathbf{E}(\mathbf{Y}_k | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})) = \mathbf{0},$$

para $k \geq 1$, tomándose, para $k = 1$, la esperanza sin condicionar. Para el estudio de la convergencia en ley es, en cambio, imprescindible la hipótesis de independencia.

1.4.1 Algoritmos γ -truncados

Definición

Recordemos cómo se llevaba a cabo el cambio de base en el algoritmo que hemos analizado previamente. En la iteración k disponíamos, para cada dirección j , de k estimaciones de $z_j^{B_k} - c_j$:

$$x_{k,j}^1, \dots, x_{k,j}^k,$$

de las cuales se hacía la media aritmética:

$$T_{k,j} = \frac{1}{k} \sum_{r=1}^k x_{k,j}^r.$$

Se consideraba la cantidad:

$$\max_{1 \leq j \leq n} T_{k,j}.$$

Si:

$$\max_{1 \leq j \leq n} T_{k,j} > 0,$$

se tomaba, según (1.8), cf. p.20, un índice $j_k \in \arg \max_{1 \leq j \leq n} T_{k,j}$, determinándose así, tras un cambio de base, una nueva base B_{k+1} . Si:

$$\max_{1 \leq j \leq n} T_{k,j} = 0,$$

véase (1.9), cf. p.20, se cogía $j_k \in I_k$, de tal forma que $B_{k+1} = B_k$.

Consideremos ahora una variante en la forma de realizarse el cambio de base. Fijamos un número real:

$$\gamma \geq 0.$$

Si en la iteración k :

$$\max_{1 \leq j \leq n} T_{k,j} > \gamma, \tag{1.32}$$

entonces se toma:

$$j_k \in \arg \max_{1 \leq j \leq n} T_{k,j},$$

cumpléndose que $j_k \in J_k$, y se determina, según el cambio de base del algoritmo del simplex, la nueva base B_{k+1} a la que pertenece la variable j_k ; si en cambio:

$$\max_{1 \leq j \leq n} T_{k,j} \leq \gamma, \tag{1.33}$$

entonces se fija $j_k \in I_k$, o lo que es lo mismo, hacemos: $B_{k+1} = B_k$. Llamaremos a este algoritmo de estimación de un vértice óptimo: *algoritmo γ -truncado*.

Nótese que el algoritmo que se ha estudiado hasta el momento, es decir, aquél que aplicaba el algoritmo del simplex a las cantidades $T_{k,j}$, es el que resulta de hacer $\gamma = 0$ en las expresiones (1.32) y (1.33): se trata por tanto del algoritmo 0-truncado.

Damos a continuación la lista los resultados que han sido demostrados para el algoritmo 0-truncado y que siguen siendo ciertos para los algoritmos γ -truncados, con $\gamma > 0$: Proposición 1.4, Observación 1.5, Teorema 1.6, Corolario 1.7, Lema 1.8, Lema 1.9, Teorema 1.10, Corolario 1.11. Estos resultados siguen siendo válidos ya que, cuando se utiliza (1.32) y (1.33) para un valor $\gamma \geq 0$, la elección de la variable j_k es, en todo caso, $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)$ -medible, y la forma precisa en que ésta se elige no interviene en las demostraciones, siempre que se haga de forma $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k)$ -medible. No es cierto,

en cambio, el Teorema 1.12, cuya demostración ya no es correcta para un algoritmo γ -truncado con $\gamma > 0$.

Comparemos el algoritmo 0-truncado con el γ -truncado, para $\gamma > 0$. El algoritmo γ -truncado, con $\gamma > 0$, no cambia de vértice a no ser que:

$$\max_{1 \leq j \leq n} T_{k,j} > \gamma > 0,$$

es decir, a no ser que: ya sea tenga una “gran” evidencia de que el vértice al que se va a desplazar vaya a ser mejor, ya sea vaya a conseguir una mejora en el objetivo, al hacer el cambio de vértice, lo bastante grande. Resulta, en consecuencia, que el algoritmo 0-truncado es más móvil, mientras que el algoritmo γ -truncado, con $\gamma > 0$, es más conservador.

Convergencia

Se probará la convergencia del algoritmo γ -truncado cuando el parámetro de truncamiento γ sea lo suficientemente pequeño.

En concreto, se cumple el siguiente:

TEOREMA 1.17: *Existe un número $\gamma_0 > 0$ tal que para todo γ que verifique que $0 < \gamma \leq \gamma_0$, aplicando el algoritmo γ -truncado existe, para casi todas las trayectorias de las observaciones, un tamaño muestral a partir del cual el vértice identificado por el algoritmo es óptimo.*

DEMOSTRACIÓN: En la demostración del Teorema 1.12, se definió δ como un número que verificase:

$$0 < \delta < \min_{B \in \mathcal{B}} \min_{1 \leq j \leq n} \left\{ |z_j^B - c_j| : z_j^B - c_j \neq 0 \right\}.$$

Tomemos γ_0 cumpliendo:

$$0 < \gamma_0 < \delta/2, \tag{1.34}$$

sea γ un número real tal que: $0 < \gamma \leq \gamma_0$, y apliquemos el algoritmo γ -truncado. Del Corolario 1.11 se deduce que, para casi toda trayectoria de las observaciones, existe un tamaño muestral N_0 tal que si $k > N_0$:

$$\max_{1 \leq j \leq n} |W_{k,j}| \leq \gamma. \tag{1.35}$$

Si existe $k > N_0$ con:

$$\max_{1 \leq j \leq n} T_{k,j} > \gamma, \tag{1.36}$$

entonces se hace un cambio de vértice tomando una variable j_k tal que:

$$z_{j_k}^{B_k} - c_{j_k} > 0,$$

por lo que $z(P_{k+1}) < z(P_k)$. Si se cumple, para algún $k > N_0$, que:

$$\max_{1 \leq j \leq n} T_{k,j} \leq \gamma,$$

entonces no se cambia de vértice y $z(P_{k+1}) = z(P_k)$. Por tanto, (1.36) sólo puede darse una cantidad finita de veces, por lo que el algoritmo acaba deteniéndose, para $k > N_1$, en un vértice: P_k . Dicho vértice es necesariamente óptimo porque, de no serlo, existiría j tal que:

$$z_j^{B_k} - c_j > \delta,$$

y puesto que $\gamma < \delta/2$, se tendría que: $W_{k,j} > \delta/2 > \gamma$, lo que es contradictorio, siendo el vértice en el que se detiene el algoritmo γ -truncado un vértice óptimo. Q.E.D.

Obsérvese que también es válida cualquier elección de γ_0 que cumpla:

$$0 < \gamma_0 < \frac{1}{2} \min_{B \in \mathcal{B}} \min_{1 \leq j \leq n} \{z_j^B - c_j : z_j^B - c_j > 0\}.$$

Esta demostración difiere de la del Teorema 1.12 en que la desigualdad (1.35) no se cumple necesariamente cuando $\gamma = 0$.

Hay una diferencia esencial en el comportamiento de los algoritmos 0-truncado y γ -truncado, siendo $0 < \gamma \leq \gamma_0$, con γ_0 como en (1.34). Para ambos algoritmos, con probabilidad uno, existe un tamaño muestral a partir del cual los vértices que identifican son óptimos. El algoritmo γ -truncado cambia de vértice cuando:

$$\max_{1 \leq j \leq n} T_{k,j} > \gamma,$$

y, con probabilidad uno, a partir de un tamaño muestral se detiene en un vértice óptimo. En cambio, el algoritmo 0-truncado, que cambia de vértice cuando:

$$\max_{1 \leq j \leq n} T_{k,j} > 0,$$

puede, en principio, desplazarse indefinidamente de vértice en vértice. El motivo es que, para el algoritmo 0-truncado, si el vértice óptimo no es único, la cantidad aleatoria:

$$\max_{1 \leq j \leq n} T_{k,j}$$

es un estimador consistente de 0 y, salvo casos extremos de la distribución del error, este estimador tomará, con probabilidad uno, infinitas veces valores positivos, de ahí que el algoritmo 0-truncado se desplace indefinidamente por distintos vértices óptimos. En cambio, para el algoritmo γ -truncado, siendo γ tal que: $0 < \gamma \leq \gamma_0$, $\max_{1 \leq j \leq n} T_{k,j}$ es un estimador consistente de 0 que, a partir de un cierto tamaño muestral, tomará un valor más pequeño que γ , deteniéndose el algoritmo en un vértice óptimo, independientemente de que otros vértices también puedan ser óptimos.

En el siguiente resultado, al que se hizo referencia tras la demostración de la Proposición 1.13, se acota uniformemente en γ la probabilidad de que el algoritmo γ -truncado repita un vértice que no es óptimo.

PROPOSICIÓN 1.18: *Dada una base $B \in \mathcal{B}$ que no sea óptima y dado γ tal que $0 \leq \gamma \leq \gamma_0$, se verifica que¹:*

$$\sup_{0 \leq \gamma \leq \gamma_0} \Pr_\gamma \{B_k = B, B_{k+1} = B\} = O\left(\frac{1}{k}\right).$$

DEMOSTRACIÓN: Si $B_{k+1} = B_k$ es porque al final de la k -ésima iteración del algoritmo no ha habido un cambio de base, o lo que es lo mismo:

$$\max_{1 \leq j \leq n} T_{k,j} \leq \gamma,$$

es decir que, para cada índice j :

$$\frac{\tilde{x}_{B,j}^1 + \cdots + \tilde{x}_{B,j}^k}{k} \leq \gamma.$$

Si la base B no es óptima, existe entonces un índice j tal que $z_j^B - c_j > \delta$, y para esa variable j :

$$\begin{aligned} \tilde{w}_{B,j}^1 + \cdots + \tilde{w}_{B,j}^k &= \tilde{x}_{B,j}^1 + \cdots + \tilde{x}_{B,j}^k - k(z_j^B - c_j) \\ &\leq k(\gamma - (z_j^B - c_j)) < k(\gamma - \delta) \\ &< -k\delta/2, \end{aligned}$$

y por tanto:

$$|\tilde{w}_{B,j}^1 + \cdots + \tilde{w}_{B,j}^k| > k\delta/2.$$

¹Por \Pr_γ se designa la probabilidad cuando se aplica el algoritmo γ -truncado.

A partir de lo anterior y de la desigualdad de Tchebychev, se tiene que:

$$\begin{aligned} \Pr_\gamma\{B_k = B, B_{k+1} = B\} &\leq \sum_{j=1}^n \Pr_\gamma\{|\tilde{w}_{B,j}^1 + \cdots + \tilde{w}_{B,j}^k| > k\delta/2\} \\ &\leq \sum_{j=1}^n \frac{4}{k^2\delta^2} \mathbf{E}_\gamma|\tilde{w}_{B,j}^1 + \cdots + \tilde{w}_{B,j}^k|^2. \end{aligned}$$

Puesto que la sucesión $(\tilde{w}_{B,j}^r)_{r \geq 1}$ es, según se probó en el Lema 1.8, una diferencia de martingalas de cuadrado integrable, resulta que, para $r_1 \neq r_2$, se cumple: $\mathbf{E}_\gamma(\tilde{w}_{B,j}^{r_1}\tilde{w}_{B,j}^{r_2}) = 0$, y por tanto:

$$\Pr_\gamma\{B_k = B, B_{k+1} = B\} \leq \frac{4}{k^2\delta^2} \sum_{j=1}^n \left(\mathbf{E}_\gamma(\tilde{w}_{B,j}^1)^2 + \cdots + \mathbf{E}_\gamma(\tilde{w}_{B,j}^k)^2 \right). \quad (1.37)$$

Se demostró en el apartado 3 del Lema 1.8 que estaban uniformemente acotadas en r las esperanzas de $(\tilde{w}_{B,j}^r)^2$ y, repasando la demostración, se puede ver que también están uniformemente acotadas en γ . Sea:

$$C = \sup\{\mathbf{E}_\gamma(\tilde{w}_{B,j}^r)^2 : 1 \leq j \leq n, r \geq 1, 0 \leq \gamma \leq \gamma_0\} < +\infty.$$

Se deduce finalmente de (1.37) que, siempre que $0 \leq \gamma \leq \gamma_0$:

$$\Pr_\gamma\{B_k = B, B_{k+1} = B\} \leq \frac{4nC}{\delta^2} \cdot \frac{1}{k},$$

con lo que se concluye el resultado buscado. Q.E.D.

De esta proposición se deduce, asimismo, que el supremo en γ de la probabilidad de que el algoritmo γ -truncado repita una base que no es óptima—sin especificar qué base en concreto, como hace esta proposición— es un infinitésimo de orden $O(1/k)$. Hay que hacer notar que este resultado es más interesante en el contexto de un algoritmo γ -truncado, para el que $\gamma > 0$, ya que para ese algoritmo se sabe que, con probabilidad uno, a partir de un cierto tamaño muestral, todas las bases B_k son iguales.

1.4.2 Convergencia en ley

Resultados preliminares

Vamos a tratar el problema del Teorema Central del Límite para los algoritmos de búsqueda de un vértice óptimo que se han desarrollado. Los

resultados que se van a probar necesitan hipótesis más fuertes que las hechas hasta el momento. Las condiciones para la convergencia serán distintas para los algoritmos 0-truncados y γ -truncados, con $0 < \gamma \leq \gamma_0$. *Grosso modo*, el algoritmo 0-truncado requiere, para probar un resultado de convergencia débil, que el vértice óptimo sea único; no es en cambio necesaria esta hipótesis para los algoritmos γ -truncados, con $0 < \gamma \leq \gamma_0$. Ésta fue precisamente la motivación para introducir los algoritmos γ -truncados. La pauta de la demostración es similar a la de la convergencia casi segura del algoritmo: se establece en primer lugar un resultado de límite en distribución para:

$$\frac{1}{\sqrt{k}}(\tilde{w}_{B,j}^1 + \cdots + \tilde{w}_{B,j}^k),$$

basado en resultados de convergencia débil de martingalas, para luego extenderlo a:

$$\frac{1}{\sqrt{k}}(w_{k,j}^1 + \cdots + w_{k,j}^k) = \sqrt{k}W_{k,j}.$$

Nos basaremos en el siguiente teorema (cf. HALL et al., p.58):

TEOREMA 1.19: *Se considera $(S_{kr}, \mathcal{F}_{kr}, 1 \leq r \leq k, k \geq 1)$ un arreglo triangular que sea, para cada k fijo, una martingala de media cero y de cuadrado integrable, con diferencia de martingalas asociada:*

$$(X_{kr}, 1 \leq r \leq k, k \geq 1).$$

Supongamos que se verifican las condiciones:

- si $k \geq 1$ y $1 \leq r \leq k$:

$$\mathcal{F}_{kr} \subseteq \mathcal{F}_{k+1,r}; \quad (1.38)$$

- cuando k tiende a $+\infty$:

$$\max_{1 \leq r \leq k} |X_{kr}| \xrightarrow{\text{P}} 0; \quad (1.39)$$

- se cumple:

$$\sup_{k \geq 1} \mathbf{E}(\max_{1 \leq r \leq k} X_{kr}^2) < +\infty; \quad (1.40)$$

- existe una variable aleatoria finita casi seguramente: ξ^2 , tal que:

$$\sum_{r=1}^k X_{kr}^2 \xrightarrow{\text{P}} \xi^2, \quad (1.41)$$

cuando $k \rightarrow +\infty$.

Se cumple entonces que:

$$S_{kk} = \sum_{r=1}^k X_{kr},$$

converge en distribución, cuando $k \rightarrow +\infty$, a una variable aleatoria con función característica, en la variable $t \in \mathbb{R}$:

$$\mathbf{E} \exp\left\{-\frac{1}{2}\xi^2 t^2\right\}.$$

Fijemos una base $B \in \mathcal{B}$, un índice j , siendo $1 \leq j \leq n$, y un número γ , con $0 \leq \gamma \leq \gamma_0$, siendo γ_0 como en (1.34), cf. p.46. Consideramos, para aplicar el teorema enunciado:

$$X_{kr} = \frac{1}{\sqrt{k}} \tilde{w}_{B,j}^r,$$

y:

$$\mathcal{F}_{kr} = \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_r),$$

para $k \geq 1$ y $1 \leq r \leq k$.

Es obvio que la propiedad (1.38) se cumple.

Estudiemos bajo qué condiciones se verifica (1.39), es decir:

$$\frac{1}{\sqrt{k}} \max_{1 \leq r \leq k} |\tilde{w}_{B,j}^r| \xrightarrow{P} 0.$$

Necesitamos un resultado previo acerca de las variables \mathbf{Y}_k que, recordémoslo, eran independientes y de segundos momentos uniformemente acotados.

LEMA 1.20: Sea $\varepsilon > 0$. Se verifica que:

$$\lim_{k \rightarrow +\infty} \Pr \left\{ \max_{1 \leq r \leq k} \max_j |Y_r - Y_{r,j}| > \varepsilon \sqrt{k} \right\} = 0,$$

si y sólo si:

$$\lim_{k \rightarrow +\infty} \sum_{r=1}^k \Pr \{ \max_j |Y_r - Y_{r,j}| > \varepsilon \sqrt{k} \} = 0.$$

DEMOSTRACIÓN: Se basa en la observación siguiente. Si x es un número real tal que $0 \leq x \leq 1/2$, entonces:

$$e^{-2x} \leq 1 - x \leq e^{-x}. \quad (1.42)$$

Para todo número natural r se verifica que:

$$\Pr\{\max_j |Y_{r,j} - Y_r| > \varepsilon\sqrt{k}\} \leq \frac{1}{k\varepsilon^2} \mathbf{E} \max_j |Y_{r,j} - Y_r|^2,$$

y existe una constante A , que no depende de r , tal que, para todo r :

$$\mathbf{E} \max_j |Y_{r,j} - Y_r|^2 \leq A,$$

luego siempre que tomemos k mayor que un cierto k_0 , se cumplirá, para todo r :

$$\Pr\{\max_j |Y_{r,j} - Y_r| > \varepsilon\sqrt{k}\} \leq 1/2.$$

Puesto que las variables $\max_j |Y_{r,j} - Y_r|$ son, para distintos valores de r , independientes, se tiene que:

$$p_k = \Pr\left\{\max_{1 \leq r \leq k} \max_j |Y_r - Y_{r,j}| > \varepsilon\sqrt{k}\right\},$$

es tal que:

$$p_k = 1 - \prod_{r=1}^k (1 - \Pr\{\max_j |Y_{r,j} - Y_r| > \varepsilon\sqrt{k}\}),$$

y para $k > k_0$, a la vista de (1.42), se cumple:

$$0 \leq 1 - \exp\left\{-\sum_{r=1}^k \Pr\{\max_j |Y_r - Y_{r,j}| > \varepsilon\sqrt{k}\}\right\} \leq p_k,$$

y también:

$$0 \leq p_k \leq 1 - \exp\left\{-2 \sum_{r=1}^k \Pr\{\max_j |Y_r - Y_{r,j}| > \varepsilon\sqrt{k}\}\right\}.$$

De estas desigualdades se deduce inmediatamente que: $\lim_{k \rightarrow +\infty} p_k = 0$ si y sólo si:

$$\lim_{k \rightarrow +\infty} \sum_{r=1}^k \Pr\{\max_j |Y_r - Y_{r,j}| > \varepsilon\sqrt{k}\} = 0,$$

ya que estas desigualdades son ciertas para todo $k > k_0$.

Q.E.D.

La proposición siguiente enuncia dos condiciones suficientes para que se verifique la propiedad (1.39), cf. p.50. Se supondrán ciertas también todas las condiciones ya impuestas a $(\mathbf{Y}_k)_{k \geq 1}$.

PROPOSICIÓN 1.21: *Si se cumple alguna de las dos condiciones:*

1. *las variables \mathbf{Y}_k , para $k \geq 1$, tienen la misma distribución;*
2. *existe $\nu > 0$ tal que $\sup_{k \geq 1} \mathbf{E} \|\mathbf{Y}_k\|^{2+\nu} < +\infty$,*

entonces se verifica:

$$\frac{1}{\sqrt{k}} \max_{1 \leq r \leq k} |\tilde{w}_{B,j}^r| \xrightarrow{p} 0,$$

cuando $k \rightarrow +\infty$.

DEMOSTRACIÓN: Tenemos que probar que para todo $\varepsilon > 0$:

$$\lim_{k \rightarrow +\infty} \Pr\{ \max_{1 \leq r \leq k} |\tilde{w}_{B,j}^r| > \varepsilon \sqrt{k} \} = 0.$$

Se demostró en el Lema 1.8 que cada término $\tilde{w}_{B,j}^r$ era una combinación lineal de una cantidad acotada de sumandos de la forma $(Y_r - Y_{r,l})/\eta$, siendo las constantes que multiplicaban a estas variables también acotadas. Sea D el número que resulta de multiplicar el número máximo de sumandos por el máximo valor absoluto de los coeficientes de la combinación lineal. Si para algún r se tiene que: $|\tilde{w}_{B,j}^r| > \varepsilon \sqrt{k}$, es porque para algún l :

$$\frac{|Y_r - Y_{r,l}|}{\eta} > \frac{\varepsilon}{D} \sqrt{k}.$$

Por tanto:

$$\Pr\{ \max_{1 \leq r \leq k} |\tilde{w}_{B,j}^r| > \varepsilon \sqrt{k} \} \leq \Pr\left\{ \max_{1 \leq r \leq k} \max_l |Y_r - Y_{r,l}| > \frac{\eta \varepsilon}{D} \sqrt{k} \right\},$$

y para llegar al resultado deseado bastará con probar que el segundo miembro de esta desigualdad tiende a cero, y esto es equivalente, según el Lema 1.20, a probar que para todo $\varepsilon > 0$:

$$\lim_{k \rightarrow +\infty} \sum_{r=1}^k \Pr\{ \max_l |Y_r - Y_{r,l}| > \varepsilon \sqrt{k} \} = 0. \quad (1.43)$$

Demostremoslo para cada una de las dos hipótesis del enunciado.

1. Si las variables $(\mathbf{Y}_k)_{k \geq 1}$ son independientes e idénticamente distribuidas, entonces también lo son las variables $(\max_l |Y_k - Y_{k,l}|)_{k \geq 1}$. Si definimos, para $k \geq 1$ y $1 \leq r \leq k$:

$$Z_{kr} = \frac{1}{\sqrt{k}} \max_l |Y_r - Y_{r,l}|,$$

entonces (Z_{kr}) resulta ser un arreglo triangular de variables independientes para cada k fijo. Este arreglo triangular es infinitesimal—verifica la condición UAD— ya que, dado $\varepsilon > 0$:

$$\max_{1 \leq r \leq k} \Pr\{|Z_{kr}| \geq \varepsilon\} = \Pr\{\max_l |Y_1 - Y_{1,l}| \geq \varepsilon\sqrt{k}\},$$

por tener todas las variables \mathbf{Y}_r la misma distribución, y es obvio que se cumple, cuando $k \rightarrow +\infty$, que: $\Pr\{\max_l |Y_1 - Y_{1,l}| \geq \varepsilon\sqrt{k}\}$ tiende a cero, y por tanto:

$$\lim_{k \rightarrow +\infty} \max_{1 \leq r \leq k} \Pr\{|Z_{kr}| \geq \varepsilon\} = 0.$$

El arreglo triangular (Z_{kr}) es también asintóticamente normal en el sentido de que existe una sucesión $(a_k)_{k \geq 1}$ tal que:

$$\left(\sum_{r=1}^k Z_{kr} \right) - a_k$$

tiende en distribución, cuando $k \rightarrow +\infty$, a una variable normal, y este hecho es una consecuencia inmediata del Teorema Central del Límite de Lévy para variables aleatorias independientes e idénticamente distribuidas con segundos momentos finitos.

Bajo estas dos hipótesis: infinitesimalidad del arreglo triangular y normalidad asintótica, se verifica— cf. IBARROLA et al., p.511, Teorema 15.14— que, para todo $\varepsilon > 0$:

$$\lim_{k \rightarrow +\infty} \sum_{r=1}^k \Pr\{|Z_{kr}| > \varepsilon\} = 0,$$

o lo que es lo mismo:

$$\lim_{k \rightarrow +\infty} \sum_{r=1}^k \Pr\{\max_l |Y_r - Y_{r,l}| > \varepsilon\sqrt{k}\} = 0,$$

como queríamos demostrar— véase (1.43).

2. Supongamos que para algún $\nu > 0$:

$$\sup_{k \geq 1} \mathbf{E} \|\mathbf{Y}_k\|^{2+\nu} < +\infty.$$

Se tiene, por la desigualdad de Tchebychev, que existe una constante C tal que, para todo r y todo k :

$$\Pr\{\max_l |Y_r - Y_{r,l}| > \varepsilon\sqrt{k}\} \leq \frac{C}{\varepsilon^{2+\nu} k^{1+\frac{\nu}{2}}},$$

de donde:

$$\sum_{r=1}^k \Pr\{\max_l |Y_r - Y_{r,l}| > \varepsilon\sqrt{k}\} \leq \frac{C}{\varepsilon^{2+\nu} k^{\frac{\nu}{2}}},$$

que tiende a cero cuando $k \rightarrow +\infty$, verificándose (1.43). Q.E.D.

La condición (1.40), cf. p.50, es decir:

$$\sup_{k \geq 1} \frac{1}{k} \mathbf{E} \left(\max_{1 \leq r \leq k} (\tilde{w}_{B,j}^r)^2 \right) < +\infty,$$

se verifica siempre, ya que:

$$\frac{1}{k} \mathbf{E} \left(\max_{1 \leq r \leq k} (\tilde{w}_{B,j}^r)^2 \right) \leq \frac{1}{k} \sum_{r=1}^k \mathbf{E} (\tilde{w}_{B,j}^r)^2,$$

y el miembro de la derecha de esta desigualdad está uniformemente acotado en k puesto que:

$$\sup_{k \geq 1} \mathbf{E} (\tilde{w}_{B,j}^r)^2 < +\infty,$$

como se probó en el apartado 3 del Lema 1.8.

Para comprobar todas las hipótesis del Teorema 1.19 falta únicamente estudiar cuándo se cumple la condición (1.41), cf. p.50, es decir: existe una variable ξ^2 , finita casi seguramente, tal que, cuando $k \rightarrow +\infty$:

$$\frac{1}{k} \sum_{r=1}^k (\tilde{w}_{B,j}^r)^2 \xrightarrow{\text{p}} \xi^2.$$

LEMA 1.22: *Si las variables $(\mathbf{Y}_k)_{k \geq 1}$ están idénticamente distribuidas entonces se verifica la propiedad de Lindeberg condicional:*

$$\frac{1}{k} \sum_{r=1}^k \mathbf{E} \left((\tilde{w}_{B,j}^r)^2 I_{\{|\tilde{w}_{B,j}^r| > \varepsilon\sqrt{k}\}} \mid \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{r-1}) \right) \xrightarrow{\text{p}} 0,$$

cuando $k \rightarrow +\infty$, para todo $\varepsilon > 0$.

DEMOSTRACIÓN: Si la base B_r es la base $B^{(i)}$, para algún i tal que:

$$i \in \{1, \dots, K\},$$

entonces existen constantes $c_{l,i,B}$ tales que:

$$\tilde{w}_{B,j}^r = \sum_{l=1}^{n-m} c_{l,i,B}(Y_r - Y_{r,l}).$$

Por tanto:

$$Z_k = \frac{1}{k} \sum_{r=1}^k \mathbf{E} \left((\tilde{w}_{B,j}^r)^2 I_{\{|\tilde{w}_{B,j}^r| > \varepsilon \sqrt{k}\}} | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{r-1}) \right),$$

verifica que:

$$\begin{aligned} Z_k &= \frac{1}{k} \sum_{i=1}^K \sum_{r=1}^k \mathbf{E} \left(\left(\sum_{l=1}^{n-m} c_{l,i,B}(Y_r - Y_{r,l}) \right)^2 \right. \\ &\quad \cdot I_{\{|\sum_l c_{l,i,B}(Y_r - Y_{r,l})| > \varepsilon \sqrt{k}\}} I_{\{B_r = B^{(i)}\}} | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{r-1}) \Big). \end{aligned}$$

Se tiene que la variable $I_{\{B_r = B^{(i)}\}}$ es $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{r-1})$ -medible y que las variables:

$$\sum_{l=1}^{n-m} c_{l,i,B}(Y_r - Y_{r,l})$$

son independientes de $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{r-1})$, de donde se concluye que Z_k es igual a:

$$\frac{1}{k} \sum_{i=1}^K \sum_{r=1}^k I_{\{B_r = B^{(i)}\}} \mathbf{E} \left(\left(\sum_{l=1}^{n-m} c_{l,i,B}(Y_r - Y_{r,l}) \right)^2 I_{\{|\sum_l c_{l,i,B}(Y_r - Y_{r,l})| > \varepsilon \sqrt{k}\}} \right),$$

y por tanto:

$$Z_k \leq \frac{1}{k} \sum_{i=1}^K \sum_{r=1}^k \mathbf{E} \left(\left(\sum_{l=1}^{n-m} c_{l,i,B}(Y_r - Y_{r,l}) \right)^2 I_{\{|\sum_l c_{l,i,B}(Y_r - Y_{r,l})| > \varepsilon \sqrt{k}\}} \right). \quad (1.44)$$

Las variables aleatorias:

$$\sum_{l=1}^{n-m} c_{l,i,B}(Y_r - Y_{r,l}),$$

cuando está fijado i y varía el índice r , son variables aleatorias independientes e idénticamente distribuidas de media cero y varianza finita, y verifican por tanto la condición de Lindeberg:

$$\lim_{k \rightarrow +\infty} \frac{1}{k} \sum_{r=1}^k \mathbf{E} \left(\left(\sum_{l=1}^{n-m} c_{l,i,B}(Y_r - Y_{r,l}) \right)^2 I_{\{|\sum_l c_{l,i,B}(Y_r - Y_{r,l})| > \varepsilon \sqrt{k}\}} \right) = 0,$$

y se deduce de (1.44) que $(Z_k)_{k \geq 1}$ tiende en probabilidad a cero. Q.E.D.

Siguiendo bajo la hipótesis de que las variables \mathbf{Y}_k tienen la misma distribución, definamos una sucesión de variables aleatorias: $(V_k^2)_{k \geq 1}$, que nos será de ayuda en lo sucesivo. Se define:

$$V_k^2 = \frac{1}{k} \sum_{r=1}^k \mathbf{E} \left((\tilde{w}_{B,j}^r)^2 | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{r-1}) \right).$$

Se tiene que:

$$\begin{aligned} V_k^2 &= \frac{1}{k} \sum_{i=1}^K \sum_{r=1}^k \mathbf{E} \left(\left(\sum_l c_{l,i,B}(Y_r - Y_{r,l}) \right)^2 I_{\{B_r=B(i)\}} | \sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{r-1}) \right) \\ &= \frac{1}{k} \sum_{i=1}^K \sum_{r=1}^k I_{\{B_r=B(i)\}} \mathbf{E} \left(\sum_l c_{l,i,B}(Y_r - Y_{r,l}) \right)^2 \\ &= \sum_{i=1}^K c_{i,B}^2 \left(\frac{1}{k} \sum_{r=1}^k I_{\{B_r=B(i)\}} \right), \end{aligned}$$

donde se ha utilizado la notación:

$$c_{i,B}^2 = \mathbf{E} \left(\sum_{l=1}^{n-m} c_{l,i,B}(Y_r - Y_{r,l}) \right)^2, \quad (1.45)$$

y obsérvese que esta esperanza no depende de r . Así pues, la variable V_k^2 , que representa el error cuadrático medio en las k primeras iteraciones, condicionado cada sumando por las observaciones anteriores, se puede escribir de la siguiente forma:

$$V_k^2 = \sum_{i=1}^K c_{i,B}^2 S_k(i), \quad (1.46)$$

donde:

$$S_k(i) = \frac{1}{k} \sum_{r=1}^k I_{\{B_r=B(i)\}},$$

es igual a la proporción de tiempo que ha pasado el algoritmo en la base $B^{(i)}$ durante las k primeras iteraciones y, en consecuencia, V_k^2 tiene la siguiente interpretación: V_k^2 es el promedio de los errores cuadráticos esperados en la estimación del coste reducido j -ésimo de la base B , ponderados estos errores con arreglo a la proporción de tiempo de ocupación de cada vértice.

Una primera convergencia en ley

En el siguiente teorema se prueba, bajo las hipótesis utilizadas, que si las proporciones de tiempo que el algoritmo pasa en cada vértice son convergentes en probabilidad, entonces:

$$\frac{1}{\sqrt{k}} \sum_{r=1}^k \tilde{w}_{B,j}^r$$

tiene distribución límite. Hay que mencionar que dicha convergencia de las proporciones de tiempo de estancia en los vértices es, en esencia, la clave de todos los resultados que serán probados posteriormente.

TEOREMA 1.23: *Si las variables $(\mathbf{Y}_k)_{k \geq 1}$ están idénticamente distribuidas y para cada $i \in \{1, \dots, K\}$ existe una variable aleatoria $S(i)$ tal que, cuando $k \rightarrow +\infty$:*

$$S_k(i) \xrightarrow{P} S(i),$$

entonces:

$$\frac{1}{\sqrt{k}} \sum_{r=1}^k \tilde{w}_{B,j}^r$$

converge en distribución a una variable aleatoria con función característica, en la variable t :

$$\mathbf{E} \exp \left\{ -\frac{1}{2} \sum_{i=1}^K c_{i,B}^2 S(i) t^2 \right\}.$$

DEMOSTRACIÓN: Para poder aplicar el Teorema 1.19 estamos a falta de probar que existe una variable aleatoria ξ^2 , finita casi seguramente, tal que:

$$\frac{1}{k} \sum_{r=1}^k (\tilde{w}_{B,j}^r)^2 \xrightarrow{P} \xi^2.$$

Vamos a utilizar el Teorema 2.23 de HALL et al., p.44. En el Lema 1.22 ya se comprobó que se verificaba la propiedad de Lindeberg condicional; por

otra parte:

$$\lim_{\lambda \rightarrow +\infty} \sup_{k \geq 1} \Pr\{V_k^2 > \lambda\} = 0,$$

ya que la sucesión $(V_k^2)_{k \geq 1}$ está uniformemente acotada, como se deduce de (1.46). De aplicar el citado Teorema 2.23 se concluye que se cumple, cuando $k \rightarrow +\infty$:

$$V_k^2 - \frac{1}{k} \sum_{r=1}^k (\tilde{w}_{B,j}^r)^2 \xrightarrow{P} 0.$$

Como, por hipótesis, $(V_k^2)_{k \geq 1}$ converge en probabilidad a:

$$\xi^2 = \sum_{i=1}^K c_{i,B}^2 S(i),$$

que es finita casi seguramente, resulta que:

$$\frac{1}{k} \sum_{r=1}^k (\tilde{w}_{B,j}^r)^2 \xrightarrow{P} \xi^2,$$

que es justamente la condición (1.41), cf. p.50. El resto de la demostración se deduce directamente del Teorema 1.19. Q.E.D.

El resultado del Teorema 1.23 es de tipo teórico y debemos encontrar alguna condición suficiente para que las variables $S_k(i)$ converjan en probabilidad. Estas condiciones suficientes serán función del algoritmo γ -truncado que utilizemos. En concreto:

TEOREMA 1.24: *Si las variables $(\mathbf{Y}_k)_{k \geq 1}$ son independientes e idénticamente distribuidas, entonces:*

$$\frac{1}{\sqrt{k}} \sum_{r=1}^k \tilde{w}_{B,j}^r$$

1. *converge en distribución a una variable aleatoria normal si hay un único vértice óptimo y se aplica el algoritmo 0-truncado;*
2. *converge en distribución a una mixtura de variables aleatorias normales si se aplica el algoritmo γ -truncado, con $0 < \gamma \leq \gamma_0$.*

DEMOSTRACIÓN:

1. En el primer caso, puesto que hay un único vértice óptimo, se deduce del Teorema 1.12 que, para casi toda trayectoria de las observaciones, existe un tamaño muestral a partir del cual la base en la que está el algoritmo es óptima. Sea $B^{(i^*)}$ esa única base óptima. Por tanto, cuando $k \rightarrow +\infty$:

$$S_k(i) \xrightarrow{\text{cs}} 0, \quad \text{si } i \neq i^*, \text{ y:}$$

$$S_k(i^*) \xrightarrow{\text{cs}} 1.$$

Podemos aplicar ahora el Teorema 1.23, siendo la variable ξ^2 constante, haciendo:

$$\xi^2 = c_{i^*,B}^2,$$

de donde resulta que el límite en distribución que se obtiene es una variable aleatoria de función característica, en la variable $t \in \mathbb{R}$:

$$\mathbf{E} \exp\left\{-\frac{1}{2}\xi^2 t^2\right\},$$

es decir:

$$\frac{1}{\sqrt{k}} \left(\tilde{w}_{B,j}^1 + \cdots + \tilde{w}_{B,j}^k \right) \xrightarrow{\text{d}} \mathcal{N}(0, c_{i^*,B}^2),$$

cuando $k \rightarrow +\infty$.

2. Supongamos que para algún valor γ tal que $0 < \gamma \leq \gamma_0$ hemos aplicado el algoritmo γ -truncado. Para esta demostración no necesitaremos la hipótesis de unicidad del óptimo. Del Teorema 1.17 se dedujo que para casi todas las trayectorias de las observaciones, a partir de un cierto tamaño muestral, el algoritmo se detenía en un vértice óptimo. Para cada $i \in \{1, \dots, K\}$ se define Γ_i como el conjunto de trayectorias para las cuales el algoritmo acaba quedándose en la base $B^{(i)}$. Los conjuntos Γ_i son disjuntos, su unión tiene probabilidad uno, aunque no necesariamente se cumple que:

$$\bigcup_{i=1}^K \Gamma_i = \Omega.$$

Las bases $B^{(i)}$ para las cuales:

$$\Pr\{\Gamma_i\} > 0$$

son óptimas. Resulta entonces que si $\omega \in \Gamma_{i_0}$ para algún:

$$i_0 \in \{1, \dots, K\},$$

entonces:

$$\begin{aligned} \lim_{k \rightarrow +\infty} S_k(i)(\omega) &= 0, \text{ si } i \neq i_0, \text{ y:} \\ \lim_{k \rightarrow +\infty} S_k(i_0)(\omega) &= 1, \end{aligned}$$

cumpléndose las hipótesis del Teorema 1.23. Se deduce de esto que la sucesión $(V_k^2)_{k \geq 1}$ converge casi seguramente a una variable aleatoria que vale, en el conjunto Γ_i : $c_{i,B}^2$. Concretamente:

$$\xi^2(\omega) = \sum_{i=1}^K c_{i,B}^2 I_{\{\omega \in \Gamma_i\}}.$$

Del Teorema 1.23 se concluye que:

$$\frac{1}{\sqrt{k}} \sum_{r=1}^k \tilde{w}_{B,j}^r$$

converge en distribución a una variable aleatoria con función característica, en la variable t :

$$\mathbf{E} \exp \left\{ -\frac{1}{2} \xi^2 t^2 \right\},$$

función característica que resulta ser igual, para cada $t \in \mathbb{R}$, a:

$$\sum_{i=1}^K \Pr\{\Gamma_i\} \exp \left\{ -\frac{1}{2} c_{i,B}^2 t^2 \right\}, \quad (1.47)$$

que es la función característica de una mixtura de variables normales. De forma más precisa, si: para cada $i \in \{1, \dots, K\}$ la variable X_i es $\mathcal{N}(0, c_{i,B}^2)$; la variable T toma los valores $1, \dots, K$ con probabilidades $\Pr\{\Gamma_i\}$, y T es independiente de (X_1, \dots, X_K) , entonces la función característica de la variable aleatoria X_T viene dada por la expresión (1.47). Q.E.D.

La interpretación del teorema anterior es la siguiente. La suma de los errores cometidos en la estimación del coste reducido $z_j^B - c_j$ de la base B —

multiplicada esa suma por $1/\sqrt{k}$ — tiende: en el primer caso, a una variable normal de media cero y de varianza el error cuadrático medio de los errores de estimación del coste reducido de la base B evaluado al pasar de la única base óptima a la base B ; en el segundo caso, a una mixtura de normales de media cero y de varianzas los errores cuadráticos medios de los errores de estimación al hacer un cambio de base de las distintas bases óptimas en las que se detiene el algoritmo a la base B , siendo esa mixtura con arreglo a la probabilidad de detenerse en cada una de esas bases óptimas.

Es importante hacer notar, para la demostración anterior, que si la variable j -ésima pertenece a la base B entonces $c_{i,B}^2 = 0$, y las distribuciones normales que han aparecido pueden ser degeneradas.

Este último teorema se ha obtenido en relación con la primera hipótesis que se había hecho en la Proposición 1.21. Siguiendo la línea de la segunda hipótesis, se cumple el siguiente:

TEOREMA 1.25: *Si las variables \mathbf{Y}_k tienen la misma matriz de covarianzas y:*

$$\sup_{k \geq 1} \mathbf{E} \|\mathbf{Y}_k\|^4 < +\infty, \quad (1.48)$$

entonces:

$$\frac{1}{\sqrt{k}} (\tilde{w}_{B,j}^1 + \cdots + \tilde{w}_{B,j}^k)$$

1. *tiende en distribución a una variable aleatoria normal si existe un único vértice óptimo y aplicamos el algoritmo 0-truncado;*
2. *tiende en distribución a una mixtura de variables aleatorias normales si aplicamos el algoritmo γ -truncado, siendo $0 < \gamma \leq \gamma_0$.*

DEMOSTRACIÓN:

1. Queremos probar que:

$$\frac{1}{k} \left((\tilde{w}_{B,j}^1)^2 + \cdots + (\tilde{w}_{B,j}^k)^2 \right) \quad (1.49)$$

converge en probabilidad. Si la base óptima es única: $B^{(i^*)}$, entonces para casi todo $\omega \in \Omega$ existe un tamaño muestral k_0 a partir del cual:

$$\tilde{w}_{B,j}^r = \sum_{l=1}^{n-m} c_{l,i^*,B} (Y_r - Y_{r,l}).$$

Dado $k \geq k_0$, y para el punto ω fijado, se tiene que:

$$\frac{1}{k} \sum_{r=1}^k (\tilde{w}_{B,j}^r)^2 - \frac{1}{k} \sum_{r=1}^k \left(\sum_{l=1}^{n-m} c_{l,i^*,B}(Y_r - Y_{r,l}) \right)^2$$

es igual a:

$$\frac{1}{k} \sum_{r=1}^{k_0} (\tilde{w}_{B,j}^r)^2 - \frac{1}{k} \sum_{r=1}^{k_0} \left(\sum_{l=1}^{n-m} c_{l,i^*,B}(Y_r - Y_{r,l}) \right)^2,$$

que tiende, cuando $k \rightarrow +\infty$, a cero. Por tanto, para estudiar la convergencia en probabilidad de (1.49) será suficiente con estudiar la de:

$$\frac{1}{k} \sum_{r=1}^k \left(\sum_{l=1}^{n-m} c_{l,i^*,B}(Y_r - Y_{r,l}) \right)^2.$$

Recordando la definición dada en (1.45), cf. p.57, se tiene que:

$$\frac{1}{k} \sum_{r=1}^k \left(\sum_{l=1}^{n-m} c_{l,i^*,B}(Y_r - Y_{r,l}) \right)^2 \xrightarrow{\text{cs}} c_{i^*,B}^2,$$

cuando $k \rightarrow +\infty$, como se deduce de aplicar la Ley Fuerte de los Grandes Números de Kolmogorov. En efecto, las variables:

$$\left(\sum_{l=1}^{n-m} c_{l,i^*,B}(Y_r - Y_{r,l}) \right)^2$$

son independientes, de varianzas uniformemente acotadas— en virtud de (1.48)— y tienen todas la misma media— por la igualdad de las matrices de covarianzas de las variables \mathbf{Y}_r — que es igual a $c_{i^*,B}^2$. La sucesión de término general (1.49) converge casi seguramente a la constante $c_{i^*,B}^2$, de donde, según el Teorema 1.19, cuando $k \rightarrow +\infty$:

$$\frac{1}{\sqrt{k}} (\tilde{w}_{B,j}^1 + \cdots + \tilde{w}_{B,j}^k) \xrightarrow{\text{d}} \mathcal{N}(0, c_{i^*,B}^2).$$

2. La demostración de este apartado es similar a la anterior. Para cada índice $i \in \{1, \dots, K\}$ se cumple, cuando $k \rightarrow +\infty$, que:

$$\frac{1}{k} \sum_{r=1}^k \left(\sum_{l=1}^{n-m} c_{l,i,B}(Y_r - Y_{r,l}) \right)^2 \xrightarrow{\text{cs}} c_{i,B}^2.$$

Dado un punto ω del espacio muestral tal que:

$$\omega \in \Gamma_i,$$

existe un tamaño muestral k_0 tal que si $r \geq k_0$ entonces:

$$\tilde{w}_{B,j}^r = \sum_{l=1}^{n-m} c_{l,i,B}(Y_r - Y_{r,l}),$$

y puesto que $\Pr\{\cup \Gamma_i\} = 1$, se tiene que, cuando $k \rightarrow +\infty$:

$$\frac{1}{k} \sum_{r=1}^k (\tilde{w}_{B,j}^r)^2 - \frac{1}{k} \sum_{i=1}^K \sum_{r=1}^k I_{\Gamma_i} \left(\sum_{l=1}^{n-m} c_{l,i,B}(Y_r - Y_{r,l}) \right)^2 \xrightarrow{\text{cs}} 0,$$

de donde se deduce que:

$$\frac{1}{k} \sum_{r=1}^k (\tilde{w}_{B,j}^r)^2 \xrightarrow{\text{cs}} \sum_{i=1}^K c_{i,B}^2 I_{\Gamma_i}.$$

Una vez probada esta convergencia, la demostración es análoga a la del Teorema 1.24, y la distribución límite a la que se llega es la misma.

Q.E.D.

Obsérvese que esta misma demostración hubiese servido para probar el Teorema 1.24. Sin embargo, el método utilizado en esta demostración no permite llegar al resultado del Teorema 1.23. Por eso, nos ha parecido interesante hacer una demostración del Teorema 1.24 en la línea de las técnicas empleadas por el Teorema 1.23.

Resumamos los resultados de distribuciones límite que hemos obtenido hasta el momento. Bajo dos tipos de hipótesis, por un lado: observaciones independientes e idénticamente distribuidas con segundos momentos finitos, y por otro lado: observaciones independientes con mismas covarianzas y cuartos momentos acotados, hemos llegado a probar un teorema central del límite para la sucesión:

$$\frac{1}{\sqrt{k}}(\tilde{w}_{B,j}^1 + \cdots + \tilde{w}_{B,j}^k),$$

con dos tipos de algoritmos: el algoritmo 0-truncado, con óptimo único, y el algoritmo γ -truncado, con $0 < \gamma \leq \gamma_0$. En ambos casos, las condiciones para el teorema central del límite presuponen algún conocimiento

sobre el propio problema. El algoritmo 0-truncado no depende— en sí— del problema, pero necesitamos saber que el óptimo es único para obtener una distribución límite; para el algoritmo γ -truncado, el teorema central del límite tiene hipótesis mucho menos restrictivas— ya no es necesaria la unicidad del óptimo— pero, a cambio, se paga el precio de que el parámetro de truncamiento γ debe verificar: $0 < \gamma \leq \gamma_0$, donde γ_0 depende de parámetros desconocidos, aunque se tiene la ventaja de que es válida la convergencia para valores de γ suficientemente pequeños.

1.4.3 Convergencia en distribución de los errores

Aún no hemos abordado el resultado fundamental, que es el de la convergencia en ley, fijado un índice j , de la sucesión, en la variable k :

$$\sqrt{k}W_{k,j} = \frac{1}{\sqrt{k}}(w_{k,j}^1 + \cdots + w_{k,j}^k).$$

Hemos analizado la distribución límite de:

$$\frac{1}{\sqrt{k}}(\tilde{w}_{B,j}^1 + \tilde{w}_{B,j}^2 + \cdots + \tilde{w}_{B,j}^k) \tag{1.50}$$

y queremos estudiar la de $\sqrt{k}W_{k,j}$, es decir:

$$\frac{1}{\sqrt{k}}(\tilde{w}_{B_k,j}^1 + \tilde{w}_{B_k,j}^2 + \cdots + \tilde{w}_{B_k,j}^k). \tag{1.51}$$

A la vista de estas dos expresiones, queda claro que las sucesiones (1.50) y (1.51) son de naturaleza diferente. En (1.50), la base de llegada es fija; en (1.51), la base de llegada es aleatoria, pudiendo ser distinta de una etapa de la estimación a otra. La demostración de la convergencia en ley de (1.51) se hará a través de la convergencia de (1.50), pero sin poderse aplicar los métodos de distribuciones límite para diferencias de martingalas ya que:

$$\frac{1}{\sqrt{k}} \tilde{w}_{B_k,j}^r$$

no es un arreglo triangular de diferencias de martingalas.

TEOREMA 1.26: Si se verifica alguna de las dos condiciones siguientes:

- las variables \mathbf{Y}_k son independientes, idénticamente distribuidas y con segundos momentos finitos;
- las variables \mathbf{Y}_k son independientes, tienen la misma matriz de covarianzas y sus cuartos momentos son acotados,

entonces se verifica, cuando $k \rightarrow +\infty$, que:

$$\sqrt{k}W_{k,j} = \frac{1}{\sqrt{k}}(w_{k,j}^1 + \cdots + w_{k,j}^k),$$

1. tiende en distribución a una variable aleatoria normal si existe un único vértice óptimo y aplicamos el algoritmo 0-truncado;
2. tiende en distribución a una mixtura de variables aleatorias normales si aplicamos el algoritmo γ -truncado, siendo $0 < \gamma \leq \gamma_0$.

DEMOSTRACIÓN:

1. Para el primer apartado, se verifican las hipótesis del Teorema 1.24 y del Teorema 1.25, por lo que, si la única base óptima es $B^{(i^*)}$, haciendo $B = B^{(i^*)}$, se verifica, cuando $k \rightarrow +\infty$, que:

$$\frac{1}{\sqrt{k}}(\tilde{w}_{B^{(i^*)},j}^1 + \cdots + \tilde{w}_{B^{(i^*)},j}^k) \xrightarrow{d} \mathcal{N}(0, c_{i^*,B^{(i^*)}}^2).$$

Puesto que para casi toda trayectoria de las observaciones existe un tamaño muestral a partir del cual $B_k = B^{(i^*)}$, resulta que:

$$\frac{1}{\sqrt{k}}(\tilde{w}_{B^{(i^*)},j}^1 + \cdots + \tilde{w}_{B^{(i^*)},j}^k) - \frac{1}{\sqrt{k}}(w_{k,j}^1 + \cdots + w_{k,j}^k) \quad (1.52)$$

tiende a cero cuando $k \rightarrow +\infty$, por lo que ambos términos tienen la misma distribución asintótica, es decir:

$$\frac{1}{\sqrt{k}}(w_{k,j}^1 + \cdots + w_{k,j}^k) \xrightarrow{d} \mathcal{N}(0, c_{i^*,B^{(i^*)}}^2).$$

2. Para el algoritmo γ -truncado, con $0 < \gamma \leq \gamma_0$, utilizando un razonamiento similar al empleado para estudiar la diferencia (1.52), dado que:

$$\sqrt{k}W_{k,j} = \sum_{i=1}^K \frac{1}{\sqrt{k}}(\tilde{w}_{B^{(i)},j}^1 + \cdots + \tilde{w}_{B^{(i)},j}^k)I_{\{B_k=B^{(i)}\}},$$

se tiene que:

$$\sqrt{k}W_{k,j} - \sum_{i=1}^K \frac{1}{\sqrt{k}} (\tilde{w}_{B^{(i)},j}^1 + \cdots + \tilde{w}_{B^{(i)},j}^k) I_{\Gamma_i}$$

tiende, cuando $k \rightarrow +\infty$, casi seguramente a cero. Por tanto, de tener:

$$\sum_{i=1}^K \frac{1}{\sqrt{k}} (\tilde{w}_{B^{(i)},j}^1 + \cdots + \tilde{w}_{B^{(i)},j}^k) I_{\Gamma_i} \quad (1.53)$$

distribución límite, también la tendrá $\sqrt{k}W_{k,j}$ y éstas coincidirán.

Sabemos— Teorema 1.24 y Teorema 1.25— que para cada índice:

$$i \in \{1, \dots, K\},$$

se tiene que:

$$\frac{1}{\sqrt{k}} (\tilde{w}_{B^{(i)},j}^1 + \cdots + \tilde{w}_{B^{(i)},j}^k) \xrightarrow{d} Z_i,$$

cuando $k \rightarrow +\infty$, donde Z_i es una variable aleatoria con función característica, en la variable t :

$$\mathbf{E} \exp\{itZ_i\} = \mathbf{E} \exp\{-\frac{1}{2}\xi_i^2 t^2\},$$

para una cierta variable ξ_i^2 finita casi seguramente, por lo que:

$$\lim_{k \rightarrow +\infty} \mathbf{E} \exp\{it \frac{1}{\sqrt{k}} (\tilde{w}_{B^{(i)},j}^1 + \cdots + \tilde{w}_{B^{(i)},j}^k)\} = \mathbf{E} \exp\{-\frac{1}{2}\xi_i^2 t^2\},$$

pero, es más, de la demostración del Teorema 1.19—véase HALL et al., p.63— se deduce que para todo conjunto medible F se cumple:

$$\lim_{k \rightarrow +\infty} \mathbf{E}(I_F \exp\{it \frac{1}{\sqrt{k}} (\tilde{w}_{B^{(i)},j}^1 + \cdots + \tilde{w}_{B^{(i)},j}^k)\}) = \mathbf{E}(I_F \exp\{-\frac{1}{2}\xi_i^2 t^2\}). \quad (1.54)$$

La función característica, en t , de la variable (1.53) es:

$$\begin{aligned} \varphi_k(t) &= \mathbf{E} \exp\left\{it \left(\sum_{i=1}^K \frac{1}{\sqrt{k}} (\tilde{w}_{B^{(i)},j}^1 + \cdots + \tilde{w}_{B^{(i)},j}^k) I_{\Gamma_i} \right)\right\} \\ &= \sum_{i=1}^K \mathbf{E}(I_{\Gamma_i} \exp\{it \frac{1}{\sqrt{k}} (\tilde{w}_{B^{(i)},j}^1 + \cdots + \tilde{w}_{B^{(i)},j}^k)\}), \end{aligned}$$

y aplicando sucesivamente (1.54), tomando $F = \Gamma_i$, resulta que:

$$\lim_{k \rightarrow +\infty} \varphi_k(t) = \sum_{i=1}^K \mathbf{E}(I_{\Gamma_i} \exp\{-\frac{1}{2}\xi_i^2 t^2\}). \quad (1.55)$$

Recordando las demostraciones de los teoremas 1.24 y 1.25, se había probado que:

$$\xi_i^2 = \sum_{l=1}^K c_{l,B^{(i)}}^2 I_{\Gamma_l},$$

de donde:

$$I_{\Gamma_i} \exp\{-\frac{1}{2}\xi_i^2 t^2\} = I_{\Gamma_i} \exp\{-\frac{1}{2}c_{i,B^{(i)}}^2 t^2\}, \quad (1.56)$$

y por tanto:

$$\mathbf{E}(I_{\Gamma_i} \exp\{-\frac{1}{2}\xi_i^2 t^2\}) = \Pr\{\Gamma_i\} \exp\{-\frac{1}{2}c_{i,B^{(i)}}^2 t^2\},$$

y de (1.55) se concluye que:

$$\lim_{k \rightarrow +\infty} \varphi_k(t) = \sum_{i=1}^K \Pr\{\Gamma_i\} \exp\{-\frac{1}{2}c_{i,B^{(i)}}^2 t^2\},$$

que es la función característica de una mixtura de variables aleatorias normales. Se concluye que $\sqrt{k}W_{k,j}$ converge en distribución a una variable aleatoria con la citada función característica.

Obsérvese que $c_{i,B^{(i)}}^2$ tiene una expresión sencilla:

$$c_{i,B^{(i)}}^2 = \frac{1}{\eta^2} \mathbf{E}(Y_1 - Y_{1,j_i})^2,$$

donde j_i es la componente de \mathbf{Y}_1 que se asigna, en la base $B^{(i)}$, a la variable j , tomando esta última expresión el valor cero si la variable j está en la base $B^{(i)}$. Q.E.D.

1.4.4 Conclusiones

Hemos demostrado dos tipos de teoremas de distribución límite. Por un lado, en los teoremas 1.23, 1.24 y 1.25 se han probado resultados para la sucesión $(\tilde{w}_{B,j}^r)$; por otro lado, en el Teorema 1.26, se ha estudiado un resultado

para la sucesión $(\tilde{w}_{B_k, j}^r)$. La forma de demostrar estos dos tipos de teoremas parece similar, pero, en realidad, hay una diferencia sutil.

En todo este procedimiento, para los términos de error analizados, subyace, implícitamente, una sucesión de índice doble. Dichos términos de error representan un cambio de base: de una base de partida a una base de llegada. El primer índice de ese índice doble sería la base de partida, y el segundo de los dos índices sería la base de llegada. Los teoremas centrales del límite 1.23, 1.24 y 1.25 se corresponden con un paso al límite en el primer índice, fijado el segundo índice, ya que se fija la base de llegada y se analiza el comportamiento asintótico de la base de partida. El Teorema 1.26 resulta, en cambio, de una toma de límites en ambos índices: primero con respecto al primer índice y después con respecto al segundo. Puesto que estas dos formas de tomar límites da, en los dos casos, como resultado una misma base aleatoria óptima, al enfrentar, en el Teorema 1.26, los dos pasos al límite, se produce un efecto de “diagonal principal”, que es lo que se ha observado en la igualdad (1.56).

Lo esencial de este comentario es percatarse de que lo que se ha hecho es simplemente una toma de límites en una sucesión doble: primeramente para un índice— teoremas 1.23, 1.24 y 1.25— y después para el otro índice— Teorema 1.26.

CAPÍTULO 2

Problema de Programación Entera

2.1 Introducción

Tomando como base el estudio que se ha hecho del problema de programación lineal continua en el primer capítulo, se aborda la resolución de un problema de programación lineal entera haciendo la hipótesis de que la función objetivo es desconocida.

De las numerosas formas que han surgido para tratar este problema de estimación muchas han sido descartadas. Así, en este capítulo, no sólo se propone una solución para el problema estadístico, sino que también se trata de justificar con detalle por qué no se han tenido en cuenta otras alternativas, poniendo en evidencia el tipo de problemática que plantea la resolución del programa entero.

2.1.1 Objetivos del estudio

Para poder sacar partido de los resultados obtenidos para el problema de programación lineal continua, se tratará el problema de programación lineal entera mediante la técnica de los planos de corte, que reduce la resolución del problema entero a la resolución de problemas lineales continuos. Para dar un panorama más amplio de las posibles formas de abordar el problema, se estudian dos métodos de estimación, aunque el que presenta un mayor

interés, desde el punto de vista estadístico, es el segundo.

Describimos brevemente la primera técnica desarrollada. La manera de dar solución al problema de programación entera es resolver, en primer lugar, un problema lineal continuo, construyendo, a partir de éste, una sucesión de problemas lineales encajados. El primer algoritmo estudiado es estocástico en la fase de resolución del primer programa lineal, y es determinista en la fase de resolución de los problemas lineales encajados; lo llamaremos por tanto algoritmo ED. Se prueba la convergencia casi segura del algoritmo ED a una solución del problema entero y se analiza la velocidad de convergencia.

El segundo algoritmo es, en cambio, totalmente estocástico en el sentido de que la resolución del primer problema es estocástica, como en el primer método, pero la resolución de los problemas encajados es también estocástica, y lo llamaremos algoritmo EE. Bajo hipótesis bastante generales se prueba que el algoritmo EE converge casi seguramente a una solución del problema entero, demostrándose además un teorema central del límite para los errores de estimación.

Estos dos métodos están basados en muestreos diferentes y aunque la información de las muestras se podría utilizar para estimar la función objetivo, la estimación de la solución del problema entero se hará de forma directa, y no a través de la resolución de los diferentes problemas aproximados, siendo estos dos algoritmos del tipo “estimar sin resolver”, en contraposición con los métodos de “resolver y estimar”.

2.1.2 Presentación del problema

Seguiremos utilizando la misma notación que en el estudio del primer capítulo del problema de programación lineal continua.

Sea A una matriz real $m \times n$ de rango m y sea $b \in \mathbb{R}^m$. Se considera el poliedro P de \mathbb{R}^n :

$$P = \{x \in \mathbb{R}^n : Ax = b, x \geq \mathbf{0}\}.$$

Suponemos que P es acotado y que existe alguna solución entera factible, es decir, existe un vector:

$$x \in \mathbb{Z}^n$$

tal que $Ax = b$ y $x \geq \mathbf{0}$. Sea $c \in \mathbb{R}^n$ el vector de costes, que son los coeficientes de la función objetivo. Se pretende resolver el siguiente problema

de programación lineal entera:

$$\begin{aligned} \text{mín } z(x) &= c'x \\ Ax &= b \\ x &\geq \mathbf{0}, x \in \mathbb{Z}^n, \end{aligned}$$

suponiendo que el vector de costes c es desconocido. Este problema entero tiene asociado, de forma natural, un problema de programación lineal continua:

$$\begin{aligned} \text{mín } z(x) &= c'x \\ Ax &= b \\ x &\geq \mathbf{0}. \end{aligned}$$

Dada una base B , se designarán por I los índices de $\{1, \dots, n\}$ correspondientes a variables básicas y por J los correspondientes a variables no básicas. El problema de programación lineal continua se puede escribir, referido a la base B :

$$\begin{aligned} \text{mín } c'_B B^{-1}b - \sum_{j \in J} (z_j - c_j)x_j \\ x_B + B^{-1}N x_N = B^{-1}b \\ x_B \geq \mathbf{0}, x_N \geq \mathbf{0}. \end{aligned} \tag{2.1}$$

Indiquemos brevemente cómo se aplica el algoritmo de planos de corte de Gomory para dar solución al problema entero. Pueden encontrarse más detalles en: PARDO et al. y WOLSEY. Se resuelve, en una primera fase, el problema continuo asociado al problema entero, obteniéndose una base B factible y óptima, es decir, tal que:

$$B^{-1}b \geq \mathbf{0} \quad \text{y} \quad \forall j \in J, z_j - c_j \leq 0.$$

Si el vector $B^{-1}b$ tiene todas sus componentes enteras, entonces el vector que vale cero en las variables de J y $B^{-1}b$ en las variables de I es una solución del problema de programación entera, terminándose el algoritmo. Si la solución óptima no es entera es porque existe alguna componente, digamos \bar{b}_k , de $B^{-1}b$ que no es entera. Si la k -ésima ecuación de (2.1) es:

$$x_k + \sum_{j \in J} y_{kj} x_j = \bar{b}_k,$$

expresando los coeficientes de dicha ecuación como suma de sus partes entera y fraccionaria, es decir:

$$y_{kj} = [y_{kj}] + f_{kj} \quad \text{y} \quad \bar{b}_k = [\bar{b}_k] + f_k,$$

siendo $[y_{kj}]$ y $[\bar{b}_k]$ enteros, f_{kj} y f_k números de $[0, 1)$, se tiene que la k -ésima ecuación se escribe:

$$x_k + \sum_{j \in J} [y_{kj}] x_j - [\bar{b}_k] = f_k - \sum_{j \in J} f_{kj} x_j. \quad (2.2)$$

Esta k -ésima fila se llamará fila fuente. Restringidas las variables a ser enteras, el primer miembro de (2.2) es entero, siéndolo también el segundo. Por la no negatividad de las variables y dado que $f_k < 1$, resulta que el segundo miembro de (2.2) verifica:

$$f_k - \sum_{j \in J} f_{kj} x_j < 1,$$

pero como se ha justificado que dicho segundo miembro es un número entero, en realidad se cumple:

$$f_k - \sum_{j \in J} f_{kj} x_j \leq 0. \quad (2.3)$$

Esta ecuación es la ecuación del plano de corte. Al añadir la restricción dada por (2.3) se elimina la solución óptima del problema continuo (2.1) pero no se ha eliminado ninguna solución factible entera, con lo cual, imponer esta nueva restricción no modifica el problema de programación entera. Se incorpora una variable de holgura: s_1 , que es no negativa y entera, de tal forma que el problema entero se ha convertido en:

$$\begin{aligned} \text{mín } c'_B B^{-1} b - \sum_{j \in J} (z_j - c_j) x_j \\ x_B + B^{-1} N x_N = B^{-1} b \\ - \sum_{j \in J} f_{kj} x_j + s_1 = -f_k \\ x \in (\mathbb{Z}^+)^n, s_1 \in \mathbb{Z}^+. \end{aligned} \quad (2.4)$$

El efecto que tiene considerar la ecuación del plano de corte es que el problema ha pasado de tener n variables y m restricciones a tener $n + 1$ variables y $m + 1$ restricciones.

El problema, con la nueva restricción, referido a la base formada por las variables de I y por la nueva variable: s_1 , se escribe como (2.4). Dicha base es dual-factible puesto que todos los costes reducidos son menores o iguales que cero, pero no es primal-factible ya que:

$$-f_k < 0.$$

El problema ampliado se resuelve por el algoritmo dual del símplex hasta restablecer la factibilidad primal. Si la nueva solución óptima es entera, entonces lo será para el problema entero originalmente planteado; si no, se busca otra fila fuente y se repite el procedimiento anteriormente descrito, añadiendo una nueva variable de holgura entera: s_2 .

Para la convergencia de este algoritmo, se requiere utilizar la ecuación de la función objetivo:

$$z(x) + \sum_{j \in J} (z_j - c_j)x_j = c'_B B^{-1}b,$$

como una restricción más, pudiéndose escoger como fila fuente. Nosotros no lo haremos porque es precisamente esta ecuación la que desconocemos, al no saber cuáles son los coeficientes de la función objetivo.

Como se ha observado, en cada paso del algoritmo, el problema de programación lineal continua que se considera tiene una restricción más y una variable más. Por tanto, después de k etapas hay $m + k$ restricciones y $n + k$ variables, y entonces las bases son de tamaño $m + k$. Este aumento del tamaño de las bases es el mayor problema, desde el punto de vista teórico, al que se enfrenta el algoritmo estocástico EE que se va a proponer. Contamos en cambio con la ventaja de que el número de variables no básicas es un invariante: es siempre igual a $n - m$.

Hagamos una observación importante. En el problema inicial, con n variables, desconocíamos los n costes de esas variables. En los problemas ampliados, con $n + k$ variables, seguimos teniendo n parámetros desconocidos puesto que los costes de las k variables de holgura son conocidos e iguales a cero.

Asumimos las siguientes hipótesis. Suponemos que las bases factibles de los programas lineales que consideraremos no son degeneradas. Asimismo, admitimos que los programas enteros que se obtengan tras la incorporación de planos de corte son tales que se pueden resolver mediante la introducción de planos de corte y posterior resolución del problema continuo, tomando como fila fuente, utilizando uno cualquiera de los criterios habituales, una de las correspondientes a las restricciones del problema lineal.

Consideramos los siguientes problemas:

$$\begin{array}{ll} \text{mín } c'x & \text{mín } c'x \\ Ax = b & \text{y} \quad Ax = b \\ x \geq \mathbf{0} & x \geq \mathbf{0}, x \in \mathbb{Z}^n, \end{array}$$

que denotaremos por (PC)— problema continuo— y (PE)— problema entero—, respectivamente. Para evitar situaciones triviales supondremos que existe algún vértice que no es óptimo para el problema (PC).

2.2 Algoritmo ED

El algoritmo ED resuelve de forma estocástica, siguiendo la técnica desarrollada en el capítulo anterior, el problema (PC) de programación lineal continua:

$$\begin{aligned} \text{mín } c'x \\ Ax = b \\ x \geq \mathbf{0}, \end{aligned}$$

y una vez estimada su solución, aplica de forma determinista el resto del algoritmo de planos de corte de Gomory. Se trata pues de un algoritmo que es estocástico en su primera fase y determinista en la segunda, de ahí que lo llamemos ED.

2.2.1 Un método de perturbación

Motivación

Supongamos que el problema (PC) tiene una única solución óptima. El algoritmo desarrollado en el capítulo anterior, con probabilidad uno a partir de un cierto tamaño muestral, se detiene en la única base óptima de (PC), según se deduce del Teorema 1.12, cf. p.35. Dicha base verifica que:

$$\forall j \in J, z_j - c_j < 0,$$

y puesto que los errores $W_{k,j}$, para $1 \leq j \leq n$, convergen casi seguramente a cero, véase el Teorema 1.10, cf. p.31, resulta que para casi todo $\omega \in \Omega$ existe un tamaño muestral N tal que:

$$\max_{1 \leq j \leq n} T_{k,j} \leq 0,$$

para $k \geq N$. Cuando en el transcurso de la aplicación del algoritmo de resolución del problema (PC) llegamos a unas estimaciones de los costes reducidos $T_{k,j}$ que son todas menores o iguales que cero, lo que hemos obtenido

es una base óptima para un cierto problema de programación lineal continua, en concreto, uno que tenga como costes:

$$\tilde{c}_j = -T_{k,j}, \text{ para } 1 \leq j \leq n.$$

La base k -ésima es, para el problema con costes \tilde{c}_j , tanto dual-factible como primal-factible, aunque no es necesariamente óptima para el problema con los verdaderos costes c_j . Si se añade un plano de corte a la tabla con costes reducidos $T_{k,j}$, entonces la base a la que se ha incorporado la variable de holgura es dual-factible pero no primal-factible, pudiéndosele por tanto aplicar el algoritmo dual del símplex para proseguir con la resolución del problema entero. En conclusión, un problema (PC) con una única solución óptima parece prestarse bien al tipo de método de resolución del algoritmo ED.

Si el problema (PC) no tiene una única solución óptima entonces lo explicado anteriormente no puede aplicarse. Sabemos que, con probabilidad uno y a partir de un cierto tamaño muestral, el algoritmo identifica una base óptima. Lo que ocurre es que esa base es tal que existe algún índice no básico que verifica:

$$z_j - c_j = 0,$$

con lo cual puede cumplirse, para tamaños muestrales arbitrariamente grandes, que:

$$T_{k,j} > 0.$$

La consecuencia de este hecho es la siguiente. A pesar de que el algoritmo haya identificado una base óptima, es decir, una base tal que:

$$\max_{1 \leq j \leq n} z_j - c_j = 0,$$

al no conocer estos coeficientes y estimarlos, puede cumplirse:

$$\max_{1 \leq j \leq n} T_{k,j} > 0,$$

siendo pues la base considerada por el algoritmo una base primal-factible pero no dual-factible— para los costes $\tilde{c}_j = -T_{k,j}$ — no pudiéndosele aplicar el algoritmo dual del símplex. Por tanto, un problema (PC) con varias soluciones óptimas no se presta a ser resuelto empleando la técnica del algoritmo ED.

Es por este motivo que se propone un método de perturbación. El objeto de este método es introducir una pequeña alteración en los coeficientes de la función objetivo para que el problema perturbado tenga una única solución

óptima, y esté el problema perturbado en condiciones de ser resuelto, en primer lugar, de forma estocástica y, en segundo lugar, a través del algoritmo determinista de planos de corte, por el algoritmo dual del símplex.

Definición y propiedades

Se define el vector de costes perturbado: c_ε , de la siguiente manera:

$$c_{1,\varepsilon} = c_1 + \varepsilon, \quad c_{2,\varepsilon} = c_2 + \varepsilon^2, \dots, \quad c_{n,\varepsilon} = c_n + \varepsilon^n,$$

para $\varepsilon > 0$. En la perturbación se utiliza un único parámetro ε , pero las alteraciones que se hacen en los costes son de magnitud diferente para cada variable, siendo cada perturbación un infinitésimo de las anteriores. Los problemas perturbados continuo y entero:

$$\begin{array}{ll} \text{mín } c'_\varepsilon x & \text{mín } c'_\varepsilon x \\ Ax = b & \text{y } Ax = b \\ x \geq \mathbf{0} & x \geq \mathbf{0}, x \in \mathbb{Z}^n, \end{array}$$

se designan mediante $(PC)_\varepsilon$ y $(PE)_\varepsilon$, respectivamente. El siguiente resultado prueba que el problema continuo perturbado tiene una única solución óptima.

LEMA 2.1: *Existe $\eta > 0$ tal que si ε verifica: $0 < \varepsilon < \eta$, entonces $(PC)_\varepsilon$ tiene una única solución óptima.*

DEMOSTRACIÓN: Los problemas (PC) y $(PC)_\varepsilon$ están definidos sobre el mismo poliedro acotado: P . Por tanto, el problema perturbado tiene alguna solución óptima.

Supongamos, en primer lugar, dados dos vértices de P : x y y , que no tienen el mismo valor en la función objetivo del problema original, es decir, tales que:

$$c'x \neq c'y.$$

Tomando ε lo bastante pequeño, se verifica también:

$$c'_\varepsilon x \neq c'_\varepsilon y,$$

con lo cual, estos dos vértices tampoco tienen el mismo valor en la función objetivo del problema perturbado tomando:

$$\varepsilon \in (0, \eta_{x,y}),$$

para algún $\eta_{x,y} > 0$.

Supongamos ahora dados dos vértices de P distintos: x y y , que tienen el mismo valor en la función objetivo del problema original:

$$c'x = c'y.$$

Los vértices x y y están asociados a dos bases distintas: B y C , respectivamente, que son factibles para (PC), siendo sus índices básicos I y I' , respectivamente. Se cumple:

$$\sum_{i \in I} c_i x_i = \sum_{i \in I'} c_i y_i, \quad (2.5)$$

y obsérvese que los coeficientes x_i y y_i que aparecen en los sumatorios de (2.5) son todos positivos porque las bases factibles no son degeneradas. Analicemos la diferencia de los valores de la función objetivo perturbada en los vértices x y y . Esta diferencia es:

$$\sum_{i \in I} (c_i + \varepsilon^i) x_i - \sum_{i \in I'} (c_i + \varepsilon^i) y_i,$$

que es igual, observando (2.5), a:

$$\sum_{i \in I} \varepsilon^i x_i - \sum_{i \in I'} \varepsilon^i y_i. \quad (2.6)$$

Puesto que los conjuntos de índices I y I' son distintos y los coeficientes x_i , para $i \in I$, y y_i , para $i \in I'$, son todos positivos, resulta que el polinomio (2.6), en la variable ε , no es el polinomio nulo. Se puede pues tomar ε dentro de un intervalo, que depende de los vértices x y y , de la forma $(0, \eta_{x,y})$ tal que en ese intervalo:

$$\sum_{i \in I} \varepsilon^i x_i - \sum_{i \in I'} \varepsilon^i y_i \neq 0,$$

o lo que es lo mismo: para todo $\varepsilon \in (0, \eta_{x,y})$ se cumple:

$$c'_\varepsilon x \neq c'_\varepsilon y.$$

Estos dos razonamientos, según coincida el objetivo original en los dos vértices o no, se pueden repetir para cada par de vértices de P , que son una cantidad finita, resultando entonces que existe $\eta > 0$ tal que siempre que:

$$0 < \varepsilon < \eta,$$

no hay dos vértices de P que tengan el mismo valor en la función objetivo perturbada: $c'_\varepsilon x$, no pudiendo pues existir varios vértices óptimos, siendo, en conclusión, la solución del problema $(PC)_\varepsilon$ única. Q.E.D.

Una vez logrado, gracias a la perturbación por potencias de ε , un problema continuo con una única solución óptima, comprobemos que toda solución óptima del problema entero perturbado: $(PE)_\varepsilon$, es también óptima para el problema entero original: (PE) .

LEMA 2.2: *Existe $\eta > 0$ tal que si ε verifica: $0 < \varepsilon < \eta$, entonces toda solución óptima de $(PE)_\varepsilon$ es óptima para (PE) .*

DEMOSTRACIÓN: Sea $x \in \mathbb{Z}^n$ un punto de P que no sea una solución óptima de (PE) y sea $x^* \in \mathbb{Z}^n$ una solución óptima de (PE) . Se verifica por tanto:

$$c'x > c'x^*,$$

luego tomando $\varepsilon > 0$ lo bastante pequeño se cumple que:

$$c'_\varepsilon x > c'_\varepsilon x^*,$$

no pudiendo ser el punto x una solución óptima de $(PE)_\varepsilon$. Procediendo análogamente para todos los puntos factibles de (PE) que no sean óptimos, que son una cantidad finita, resulta que tomando $\varepsilon \in (0, \eta)$, para un cierto número $\eta > 0$, se tiene que toda solución óptima de $(PE)_\varepsilon$ lo es de (PE) .

Q.E.D.

Comentemos esta última demostración. No se ha probado que el punto x^* , que es óptimo para el problema (PE) , lo sea para $(PE)_\varepsilon$, sino que se ha justificado que el punto x no puede ser óptimo para $(PE)_\varepsilon$ puesto que existe otro punto factible: x^* , con menor valor de la función objetivo perturbada, sin ser necesariamente x^* la solución óptima.

Para probar el Lema 2.2 no ha sido determinante que la perturbación considerada sea por potencias de ε . Cualquier perturbación: c_ε , en la variable $\varepsilon \geq 0$ siendo $c_0 = c$, que fuera continua en $\varepsilon = 0$ hubiese permitido demostrar este resultado. En cambio, en el Lema 2.1 sí ha sido decisivo que la perturbación se haya hecho por potencias de ε .

En conclusión, tomando un parámetro de perturbación ε en un intervalo de la forma $(0, \eta)$, para algún $\eta > 0$, se obtiene un problema perturbado continuo con una única solución óptima, y tal que cualquier solución óptima

del problema perturbado entero es también óptima para el problema original entero. Por tanto, para encontrar una solución óptima del problema entero (PE) es suficiente resolver $(PE)_\varepsilon$.

Muestreo del problema perturbado

Recordemos el tipo de muestreo que se consideraba para resolver un problema de programación lineal continua. Se partía de una sucesión:

$$(\mathbf{Y}_k)_{k \geq 1}$$

de variables aleatorias $(n - m + 1)$ -dimensionales independientes, de media cero, y de segundos momentos uniformemente acotados. Las componentes de la variable \mathbf{Y}_k representaban errores en la observación de la función objetivo en unos ciertos puntos de P .

La dificultad que plantea el método de perturbación es cómo muestrearlo, es decir, cómo obtener las observaciones muestrales necesarias para resolver el problema $(PC)_\varepsilon$.

Sea $z(\cdot)$ la función objetivo original y $z_\varepsilon(\cdot)$ la función objetivo perturbada. Nosotros podemos observar, en la k -ésima etapa del muestreo, los valores:

$$z(P_k) \quad \text{y} \quad z(P_{k,j}), \quad \text{para } j \in J_k,$$

con errores, que no son observables:

$$Y_k \quad \text{y} \quad Y_{k,j}, \quad \text{para } j \in J_k,$$

para un cierto vértice P_k de P y para los puntos $P_{k,j}$ de P , siendo todos estos puntos medibles respecto de la σ -álgebra:

$$\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1}).$$

Nótese bien que si podemos observar $z(P_k) + Y_k$, entonces también podemos observar la variable $z_\varepsilon(P_k) + Y_k$, ya que:

$$z_\varepsilon(P_k) + Y_k = z(P_k) + \sum_{i \in I_k} \varepsilon^i (B_k^{-1}b)_i + Y_k,$$

siendo:

$$\sum_{i \in I_k} \varepsilon^i (B_k^{-1}b)_i$$

medible respecto de $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})$. Análogamente se procede con el término $z_\varepsilon(P_{k,j}) + Y_{k,j}$ y con la diferencia de ambos términos.

La conclusión es que para obtener el muestreo necesario para estimar la solución óptima del problema perturbado no es preciso hacer un nuevo muestreo, sino que el propio muestreo del problema original se puede adaptar al problema perturbado obteniendo los mismos errores de observación: \mathbf{Y}_k .

Vamos pues a suponer, a partir de ahora, que el problema (PC) tiene una única solución óptima ya que, de no ser así, consideraríamos en su lugar el problema perturbado, aplicando los lemas 2.1 y 2.2, sin que el muestreo de este problema plantee ninguna dificultad.

2.2.2 Estimación de una solución

A continuación, definimos el algoritmo ED de estimación de una solución del problema entero (PE).

Definición del algoritmo

Se aproxima la solución del problema (PC) por el algoritmo de estimación desarrollado en el capítulo anterior. En la k -ésima etapa del muestreo nos situamos en una base B_k , siendo sus costes reducidos:

$$z_j^{B_k} - c_j, \text{ para } 1 \leq j \leq n,$$

que se estiman mediante las cantidades $T_{k,j}$, cometiendo errores de estimación $W_{k,j}$, para $1 \leq j \leq n$. Si se verifica que:

$$\max_{1 \leq j \leq n} T_{k,j} = 0,$$

resulta que la tabla del simplex de la k -ésima etapa es tanto dual-factible como primal-factible. Se resuelve entonces el problema de programación lineal entera:

$$\text{mín } \sum_{j=1}^n -T_{k,j}x_j,$$

sujeto a las restricciones:

$$Ax = b, x \geq \mathbf{0}, x \in \mathbb{Z}^n,$$

llevándose a cabo esta resolución por el algoritmo de planos de corte de Gomory.

Se define una sucesión de variables aleatorias mediante:

- $K_0 = 0$;
- $K_r = \min\{k > K_{r-1} : \max_{1 \leq j \leq n} T_{k,j} = 0\}$, para $r \geq 1$.

Si $\{k > K_{r-1} : \max_{1 \leq j \leq n} T_{k,j} = 0\}$ es vacío, entonces K_r, K_{r+1}, \dots , toman el valor $+\infty$.

La sucesión aleatoria de instantes $(K_r)_{r \geq 1}$ representa los momentos en los que la tabla de la resolución del problema (PC) es dual-factible, y son por tanto las etapas en las que se resuelve un problema entero aproximado. Se verifica el siguiente resultado.

PROPOSICIÓN 2.3: *Los términos de la sucesión $(K_r)_{r \geq 1}$ son finitos casi seguramente.*

DEMOSTRACIÓN: Puesto que el problema (PC) tiene, por hipótesis, una única base óptima, que llamaremos B , siendo I y J sus índices básicos y no básicos, respectivamente, del Teorema 1.12, cf. p.35, se deduce que para casi todo $\omega \in \Omega$ existe un tamaño muestral N a partir del cual:

$$B_k = B.$$

Para $j \in I$ y $k \geq N$, se cumple:

$$T_{k,j} = 0.$$

Para las variables no básicas se verifica:

$$\max_{j \in J} z_j^B - c_j < 0,$$

y del Corolario 1.11, cf. p.34, se deduce que para casi todo punto del espacio muestral, para un tamaño suficientemente grande, se tiene:

$$\max_{j \in J} T_{k,j} < 0,$$

quedando probado que para casi todo $\omega \in \Omega$ existe una infinidad— de hecho, son todos menos una cantidad finita— de tamaños muestrales para los que $\max_{1 \leq j \leq n} T_{k,j} = 0$. Q.E.D.

En esta última demostración, es esencial el hecho de que (PC) tenga una única solución óptima. Para un tamaño muestral k tal que:

$$K_r \leq k < K_{r+1},$$

para algún $r \geq 1$, se define la estimación $\hat{x}^{(k)}$ de una solución óptima de (PE) como la solución óptima del problema entero con función objetivo:

$$\sum_{j=1}^n -T_{K_r, j} x_j.$$

Si $k < K_1$, entonces la definición de $\hat{x}^{(k)}$ es cualquiera.

Hay que hacer notar que la solución primal-factible identificada en la etapa K_r -ésima no es necesariamente óptima para el problema (PC), aunque la tabla del simplex sí sea óptima.

Convergencia del algoritmo

El algoritmo ED que se ha descrito encuentra, con probabilidad uno, una solución óptima del problema entero.

TEOREMA 2.4: *Para casi todo $\omega \in \Omega$ existe un tamaño muestral N tal que si $k > N$ entonces la estimación $\hat{x}^{(k)}$ es una solución óptima de (PE).*

DEMOSTRACIÓN: Retomando la notación de la demostración de la Proposición 2.3, con probabilidad uno y a partir de un tamaño muestral N_0 se tiene que:

$$B_k = B.$$

En etapas K_r posteriores a la N_0 -ésima, el algoritmo ED resuelve un problema entero, con la misma región factible que (PE), cuya función objetivo es:

$$\sum_{j=1}^n -T_{K_r, j} x_j, \quad (2.7)$$

siendo la función objetivo del problema original igual a:

$$-\sum_{j=1}^n (z_j^B - c_j) x_j \quad (2.8)$$

más una constante que solamente depende de B . Como los coeficientes de la función objetivo (2.7) convergen, dentro de un conjunto de probabilidad uno, a los coeficientes de (2.8), aplicando un razonamiento similar al de la demostración del Lema 2.2, resulta que para k mayor que un cierto tamaño muestral $N \geq N_0$, toda solución óptima del problema entero con objetivo (2.7) también es óptima para el problema entero con objetivo (2.8).

Q.E.D.

Hay que observar que esta estimación de la solución óptima de (PE) se hace sin estimar la función objetivo. Lo que se estima es una representación de la función objetivo: la representación en la base B , pero la base B no es un dato conocido sino que se va estimando etapa a etapa, siendo esta estimación de B consistente puesto que el problema continuo, tras quizá aplicarle una perturbación, tiene una única base óptima.

La velocidad de convergencia de este algoritmo al conjunto de soluciones óptimas de (PE) es la misma que la del algoritmo para el problema de programación lineal continua.

TEOREMA 2.5: *La probabilidad de que la estimación $\hat{x}^{(k)}$ no sea una solución óptima de (PE) es un infinitésimo de orden:*

$$O\left(\frac{1}{k}\right).$$

DEMOSTRACIÓN: Sea B la única base óptima del problema (PC). Tal y como se justificó en la demostración del Teorema 2.4, para la resolución del problema (PE) se puede sustituir la función objetivo:

$$c'x,$$

por la función:

$$-\sum_{j=1}^n (z_j^B - c_j)x_j,$$

dada en (2.8). También se ha probado que tomando unos números reales T_j , para $1 \leq j \leq n$, lo bastante cercanos a los coeficientes $(z_j^B - c_j)$, resulta que los problemas enteros con funciones objetivo:

$$-\sum_{j=1}^n (z_j^B - c_j)x_j \quad \text{y} \quad -\sum_{j=1}^n T_j x_j,$$

siendo la región factible la de (PE), son tales que cualquier solución óptima para la segunda función objetivo es óptima para la primera función objetivo. Existe pues $\varepsilon > 0$ tal que siempre que:

$$\max_{1 \leq j \leq n} |T_j - (z_j^B - c_j)| < \varepsilon,$$

entonces se verifica la propiedad enunciada *ut supra*. El número ε se toma cumpliendo también que: $\varepsilon < \delta/2$, véase (1.30), cf. p.41. Por comodidad en la notación, designemos por \mathbf{x}^* el conjunto de soluciones óptimas de (PE).

Probemos, en primer lugar, que se verifica el siguiente contenido:

$$\{B_k = B\} \cap \{\max_{1 \leq j \leq n} |W_{k,j}| < \varepsilon\} \subseteq \{\hat{x}^{(k)} \in \mathbf{x}^*\}. \quad (2.9)$$

En efecto, si la base k -ésima es la base B , entonces para los índices no básicos de B_k se cumple:

$$z_j^{B_k} - c_j < -\delta < -2\varepsilon,$$

y puesto que:

$$|W_{k,j}| < \varepsilon,$$

resulta que $T_{k,j} < -\varepsilon$ para los índices no básicos, y $T_{k,j} = 0$ para los índices básicos, siendo la tabla estimada del símplex una tabla dual-factible. Por la definición del algoritmo ED, el problema entero con función objetivo:

$$-\sum_{j=1}^n T_{k,j} x_j,$$

se resuelve, siendo su solución $\hat{x}^{(k)}$. Dado que, para cada $j \in \{1, \dots, n\}$, se tiene que:

$$|T_{k,j} - (z_j^{B_k} - c_j)| = |W_{k,j}| < \varepsilon,$$

resulta que la aproximación de los costes reducidos es lo bastante cercana para que se cumpla:

$$\hat{x}^{(k)} \in \mathbf{x}^*,$$

quedando probado el contenido (2.9).

Se verifica pues que:

$$\{\hat{x}^{(k)} \notin \mathbf{x}^*\} \subseteq \{B_k \neq B\} \cup \{\max_{1 \leq j \leq n} |W_{k,j}| \geq \varepsilon\},$$

y por tanto:

$$\Pr\{\hat{x}^{(k)} \notin \mathbf{x}^*\} \leq \Pr\{B_k \neq B\} + \Pr\{\max_{1 \leq j \leq n} |W_{k,j}| \geq \varepsilon\}.$$

El primer sumando del segundo miembro de esta desigualdad es, como se probó en el Teorema 1.16, cf. p.41, de orden $O(1/k)$; el segundo sumando

también es de orden $O(1/k)$, como se deduce del Lema 1.15, cf. p.40, tomando, en la notación allí empleada: $s = 0$. En conclusión:

$$\Pr\{\hat{x}^{(k)} \notin \mathbf{x}^*\} \leq O(1/k) + O(1/k),$$

obteniéndose el resultado buscado.

Q.E.D.

Con este resultado se concluye el estudio del algoritmo ED, disponiendo de resultados sobre convergencia casi segura y velocidad de esa convergencia.

2.3 Algoritmo EE

Así como el algoritmo ED resolvía el problema entero en dos fases, siendo la primera estocástica y la segunda determinista, el algoritmo que presentamos a continuación es estocástico tanto en la fase de resolución del problema continuo como en la fase de resolución del problema entero.

2.3.1 Definición del algoritmo

Un nuevo muestreo

El muestreo que se va a utilizar para este algoritmo no es el mismo que el que se empleó en la resolución del problema continuo o en el algoritmo ED. Aquel muestreo observaba los incrementos del objetivo en las direcciones no básicas de los vértices. El nuevo muestreo observa directamente los coeficientes de la función de coste. Recordemos que el objetivo planteado era el de resolver los problemas de programación lineal sin estimar los coeficientes de la función de coste. Pese a que ahora observemos variables aleatorias que estiman dichos coeficientes, no vamos a convertir el problema en uno de “resolver y estimar”, y vamos a mantener la filosofía de “estimar sin resolver”. La razón del cambio de muestreo se expondrá más adelante.

Se considera una sucesión de variables aleatorias n -dimensionales:

$$\mathbf{Y}_1, \mathbf{Y}_2, \dots$$

No debe confundirse la sucesión $(\mathbf{Y}_k)_{k \geq 1}$ que acaba de definirse con la sucesión que se definió para la resolución del problema de programación lineal

continua, en el capítulo primero. Las componentes de la variable \mathbf{Y}_k : $\mathbf{Y}_{k,j}$, para $1 \leq j \leq n$, son estimadores de los costes c_j , para $1 \leq j \leq n$:

$$\mathbf{Y}_{k,j} = c_j + e_j^{(k)},$$

siendo la variable $e_j^{(k)}$ el error de estimación. Se hacen las hipótesis:

$$\mathbf{E}e_j^{(k)} = 0 \quad \text{y} \quad \sup_{k \geq 1} \mathbf{E}(e_j^{(k)})^2 < +\infty,$$

y supondremos que, para cada $j \in \{1, \dots, n\}$, la sucesión $(e_j^{(k)})_{k \geq 1}$ es una diferencia de martingalas para la filtración:

$$(\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k))_{k \geq 1}.$$

A lo largo de este estudio, se impondrán, para poder probar ciertos resultados, otras condiciones sobre los errores $(e_j^{(k)})$.

Primera fase del algoritmo

Los detalles de esta primera fase se explican de forma somera porque el desarrollo es el mismo que para el algoritmo del problema continuo.

Se parte de una base factible B_1^1 del poliedro P :

$$\begin{aligned} Ax &= b \\ x &\geq \mathbf{0}. \end{aligned}$$

Sea $k \geq 1$. En general, la base B_k^1 tiene como índices básicos: I_k^1 , y como índices no básicos: J_k^1 . La base B_k^1 es $\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1})$ -medible para $k > 1$.

Tomando $1 \leq r \leq k$, la estimación del coste reducido:

$$z_j^{B_k^1} - c_j,$$

para $1 \leq j \leq n$, basada en la r -ésima observación muestral: \mathbf{Y}_r , es:

$$x_{k,j}^{r,1} = \sum_{i \in I_k^1} \mathbf{Y}_{r,i} \cdot y_{ij}^{B_k^1} - \mathbf{Y}_{r,j},$$

siendo el error de estimación cometido: $w_{k,j}^{r,1}$, definido por:

$$x_{k,j}^{r,1} = z_j^{B_k^1} - c_j + w_{k,j}^{r,1},$$

igual a:

$$w_{k,j}^{r,1} = \sum_{i \in I_k^1} e_i^{(r)} y_{ij}^{B_k^1} - e_j^{(r)}.$$

Obsérvese que cuando $j \in I_k^1$, entonces el coste reducido j -ésimo es igual a cero y la estimación se hace sin error.

La estimación de $z_j^{B_k^1} - c_j$ basada en las k primeras observaciones muestrales es:

$$T_{k,j}^1 = \frac{1}{k} \sum_{r=1}^k x_{k,j}^{r,1},$$

siendo el error cometido:

$$W_{k,j}^1 = \frac{1}{k} \sum_{r=1}^k w_{k,j}^{r,1}.$$

Analizando la cantidad:

$$\max_{1 \leq j \leq n} T_{k,j}^1,$$

según sea ésta positiva o nula, se determina la base B_{k+1}^1 , exactamente igual que en el algoritmo para el problema continuo.

El procedimiento descrito anteriormente se lleva a cabo para $1 \leq k < K_1$. Es decir, se hacen $K_1 - 1$ iteraciones del algoritmo para el problema continuo:

$$\begin{aligned} &\text{mín } c'x \\ &Ax = b \\ &x \geq \mathbf{0}. \end{aligned}$$

La cantidad K_1 será en general una cantidad aleatoria. Al llegar a la base $B_{K_1}^1$ finaliza la primera fase del algoritmo EE.

Siguientes fases del algoritmo

Si la solución básica identificada por la base $B_{K_1}^1$ no es entera, entonces se introduce un plano de corte, es decir, se añade una restricción más, que suprime la solución básica de $B_{K_1}^1$, pero que no elimina ninguna solución factible del problema entero. Tenemos ahora un nuevo problema de programación lineal planteado en un espacio distinto, puesto que hay $n+1$ variables y $m+1$ restricciones. Las bases tendrán ahora tamaño $m+1$, y de los $n+1$ costes seguimos muestreando los n costes de las variables originales ya que el coste de la nueva variable es conocido e igual a cero.

Se determina, en el nuevo problema de mayor dimensión, una base factible $B_{K_1+1}^2$. El sub-índice $K_1 + 1$ sigue la numeración de la primera fase: la última base de la primera fase era la K_1 -ésima. El super-índice 2 indica que nos encontramos en la segunda fase. Los índices básicos y no básicos de esta base son:

$$I_{K_1+1}^2 \quad \text{y} \quad J_{K_1+1}^2,$$

respectivamente. La elección de la base factible $B_{K_1+1}^2$ es medible respecto de las observaciones muestrales:

$$\mathbf{Y}_1, \dots, \mathbf{Y}_{K_1}. \quad (2.10)$$

Cada una de las observaciones muestrales de (2.10) permite estimar los costes reducidos de la base $B_{K_1+1}^2$. En concreto, para $1 \leq j \leq n + 1$, la estimación de:

$$z_j^{B_{K_1+1}^2} - c_j,$$

basada en la muestra \mathbf{Y}_r , para $1 \leq r \leq K_1$, es:

$$x_{K_1+1,j}^{r,2} = \sum_{i \in I_{K_1+1}^2} \mathbf{Y}_{r,i} \cdot y_{ij}^{B_{K_1+1}^2} - \mathbf{Y}_{r,j}, \quad (2.11)$$

siendo el error de estimación cometido:

$$w_{K_1+1,j}^{r,2} = \sum_{i \in I_{K_1+1}^2} e_i^{(r)} y_{ij}^{B_{K_1+1}^2} - e_j^{(r)}.$$

A partir de una nueva observación muestral: \mathbf{Y}_{K_1+1} , se hace una nueva estimación del coste reducido. Esta estimación, que se define de la manera natural, como en (2.11), se denota por:

$$x_{K_1+1,j}^{K_1+1,2},$$

y el error que comete es:

$$w_{K_1+1,j}^{K_1+1,2}.$$

La estimación del coste reducido j -ésimo es, utilizando toda la información disponible:

$$T_{K_1+1,j}^2 = \frac{1}{K_1 + 1} \sum_{r=1}^{K_1+1} x_{K_1+1,j}^{r,2}.$$

Hay que hacer dos observaciones importantes. La variable j , que en la primera fase tomaba valores en $\{1, \dots, n\}$, ahora los toma en $\{1, \dots, n+1\}$ ya que el problema de programación lineal tiene una variable más. En la expresión (2.11), se entiende que cuando alguno de los índices i o j corresponda a la variable de holgura, el valor de la estimación de su coste es cero con error igual a cero.

A partir de este instante, el procedimiento que se sigue es el mismo. Analizando la cantidad:

$$\max_{1 \leq j \leq n+1} T_{K_1+1,j}^2,$$

se determina una nueva base: $B_{K_1+2}^2$, se actualizan las $K_1 + 1$ observaciones muestrales hechas hasta el momento, y se incorpora la $(K_1 + 2)$ -ésima observación muestral. Así hasta el tiempo $K_2 > K_1$, que marca el final de la segunda fase, en el que se añade otro plano de corte, modificando nuevamente el problema de programación lineal, teniendo éste $n + 2$ variables y $m + 2$ restricciones a lo largo de la tercera fase. La notación que se emplea en las fases subsiguientes se rige por las normas ya habituales y no se detalla por no hacer tediosa su exposición. Lamentamos lo recargado de la notación, aunque esto es inevitable al considerar el problema de estimación dividido, por un lado, según las observaciones muestrales y, por otro lado, según las fases. En general, las variables:

$$x_{k,j}^{r,s} \quad \text{y} \quad w_{k,j}^{r,s}$$

representan la estimación y el error de estimación del coste reducido de la j -ésima variable de la k -ésima base, perteneciendo ésta a la s -ésima fase, basándose la estimación en la r -ésima observación muestral. La estimación del j -ésimo coste reducido de la k -ésima base con toda la información disponible en las k primeras muestras es:

$$T_{k,j}^s = \frac{1}{k} \sum_{r=1}^k x_{k,j}^{r,s},$$

siendo el error de estimación:

$$W_{k,j}^s = \frac{1}{k} \sum_{r=1}^k w_{k,j}^{r,s}.$$

Se verifica:

$$1 \leq r \leq k, \quad K_{s-1} < k \leq K_s, \quad 1 \leq j \leq n + s - 1 \quad \text{y} \quad s \geq 1, \quad (2.12)$$

donde se define $K_0 = 0$. Nótese que el índice s es, en cierto modo, superfluo porque una vez conocido k se puede determinar s . Para mayor claridad, se procurará mantener el índice que indica el número de fase: s .

Aún no se ha justificado por qué se ha cambiado el tipo de muestreo. Las siguientes observaciones explican el motivo.

Supongamos que se verifican las condiciones dadas en (2.12). El error de estimación: $w_{k,j}^{r,s}$, se puede escribir de la forma:

$$w_{k,j}^{r,s} = \sum_{i=1}^n e_i^{(r)} X_{j,i}^{(k)}, \quad (2.13)$$

siendo la variable $X_{j,i}^{(k)}$ medible en la σ -álgebra:

$$\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_{k-1}).$$

La expresión (2.13) no supone ninguna novedad respecto del muestreo considerado para el problema continuo. Lo interesante es observar que la dependencia de $X_{j,i}^{(k)}$ de la citada σ -álgebra es, en realidad, únicamente a través de la base B_k — B_k^s si se hace aparecer la fase— y de los planos de corte introducidos, lo que escribiremos:

$$w_{k,j}^{r,s} = \sum_{i=1}^n e_i^{(r)} X_{j,i}^{(B_k)}. \quad (2.14)$$

De haber conservado el mismo muestreo que para el algoritmo continuo, la expresión (2.14) hubiese sido:

$$w_{k,j}^{r,s} = \sum_{i=1}^n e_i^{(r)} X_{j,i}^{(B_k, B_r)}, \quad (2.15)$$

ya que el error basado en la r -ésima observación muestral que se comete en la base k -ésima depende también de la r -ésima base. Hay que hacer notar que las expresiones (2.14) y (2.15) son, en su propia esencia, totalmente diferentes, y una expresión del tipo (2.15), como será razonado, no hace viable una demostración de convergencia del algoritmo EE.

2.3.2 Convergencia del algoritmo

En esta sección estudiamos las propiedades asintóticas del algoritmo EE. Se analiza la consistencia de la estimación de la solución del problema (PE), así como el comportamiento límite de los errores cometidos.

Comentarios generales

Observemos en primer lugar las nuevas dificultades a las que nos enfrentamos en el problema de programación entera y que no habían surgido en el problema continuo.

La convergencia del algoritmo se basa en probar que, para cada variable del problema, el error que se comete, en la k -ésima etapa, al estimar su coste reducido en la k -ésima base converge a cero. Los errores de estimación en la k -ésima base, basados en los de la r -ésima observación muestral, se obtienen aplicando a estos últimos las operaciones de pivoteo que hacen pasar de la r -ésima base— la base de partida— a la k -ésima base— la base de llegada—. El procedimiento de demostración que se utilizó en el algoritmo para el problema continuo fue fijar la base de llegada. Una vez fijada ésta, los errores cometidos tenían estructura de diferencia de martingalas y se probaba su convergencia a cero. Al ser las posibles bases de llegada un número finito— había tantas como bases factibles tuviese el poliedro— se deducía que el error de estimación convergía a cero, por una propiedad del tipo de que el máximo de una cantidad finita de variables que convergen a cero también tiende a cero.

En el problema entero surge una complicación que no existía en el problema continuo. Cada vez que se añade un plano de corte, la dimensión del problema de programación lineal aumenta, como ya ha sido explicado. Las posibles bases de llegada son entonces una cantidad infinita, y al tener una cantidad infinita de variables que convergen a cero, no podemos asegurar que su supremo tiende a cero. Desde el punto de vista teórico, y tendremos ocasión de seguirlo comprobando, éste es el mayor inconveniente que presenta el problema entero. El problema continuo era de naturaleza finita, pero el problema entero, al estar basado en problemas de diferentes dimensiones, es de naturaleza infinita.

Para hacer frente a estas dificultades, surgieron varias posibilidades. La primera fue suponer que existía un número M tal que cualquier restricción del problema continuo original con M planos de corte fuese tal que su solución óptima fuera entera, o dicho en otros términos, cualesquiera M planos de corte que introdujésemos, la solución del problema continuo era entera. Esto hubiese resuelto el problema de que el número de bases fuera infinito. Esta solución fue descartada porque, como es evidente, supone una enorme pérdida de generalidad y obligaría a limitarse a una clase muy restringida de poliedros.

Otra posibilidad que se consideró fue que al iniciar la segunda fase, el

nuevo problema de programación lineal que se planteaba se comenzase a estimar desde cero por el algoritmo ya desarrollado para el problema continuo. Esta técnica también hubiese suprimido la dificultad del número infinito de bases ya que, pese a poder ser el número de fases infinito, al tratarse cada fase de forma independiente de las demás, el problema de naturaleza infinita se hubiera descompuesto en una cantidad infinita de problemas finitos, que sí hubiesen sido resolubles por el método habitual. Esta técnica hubiera, asimismo, permitido conservar el mismo tipo de muestreo que para el problema de programación lineal continua, sin necesidad de introducir el nuevo muestreo basado directamente en la función objetivo. Sin embargo, este método desaprovecha una gran cantidad de información. En concreto, cuando se sitúa en una cierta fase, no tiene en cuenta ninguna de las observaciones muestrales hechas en las fases anteriores, resolviendo pues el problema en cada fase independientemente de la información que las fases previas de hecho contienen sobre la fase actual. El número de observaciones muestrales necesarias se multiplicaría pues por el número de fases, lo que requiere un número excesivo de muestras, además de estar desperdiciándose mucha información.

La tercera posibilidad, que es por la que nos hemos decantado, es la que ya ha sido explicada.

Al avanzar una fase, se cambia de problema de programación lineal, aunque toda la información que se ha obtenido durante las fases previas se adapta al nuevo problema de programación lineal, es decir, a la fase actual, de tal forma que, cuando abordamos la nueva fase, no estamos partiendo de cero en un nuevo problema, sino que se trata de un nuevo problema del que tenemos ya toda la información que proporcionan los muestreos llevados a cabo en las anteriores fases, e iniciamos pues cada fase sirviéndonos de toda la información pasada.

Éste es el interés primordial del algoritmo EE: la información se conserva a lo largo de cada iteración del algoritmo. Se conserva la información al cambiar de vértice, dentro de una misma fase, y también se conserva la información al cambiar de fase.

Esta última vía supera a las dos primeras en generalidad y en aprovechamiento de la información, aunque paga el precio de tener que modificar el tipo de muestreo para obtener expresiones del tipo (2.14) y no del tipo (2.15), aunque, como ya ha quedado claro, no por observar los costes de la función objetivo se ha convertido el problema en uno de “resolver y estimar”, puesto que los métodos utilizados son similares a los empleados en el algoritmo para el problema continuo.

Condiciones para la convergencia de los errores

Vamos a dar ahora una definición formal de los instantes: K_1, K_2, \dots , que determinan los momentos en los que se cambia de fase. Se toma:

$$K_0 = 0,$$

y cada variable K_r , para $r \geq 1$, toma valores en el conjunto:

$$1, 2, 3, \dots, +\infty,$$

siendo K_r un tiempo de parada para la filtración:

$$(\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_k))_{k \geq 1}.$$

Se supone además que dado $\omega \in \Omega$, si todos los valores $K_r(\omega)$, para $r \geq 0$, son finitos entonces la sucesión $(K_r(\omega))_{r \geq 1}$ es estrictamente creciente. En el caso contrario, la sucesión $(K_r(\omega))_{r \geq 1}$ es estrictamente creciente mientras es finita y es constante e igual a $+\infty$ desde el primer índice en el que es infinita. Esto equivale a considerar que se permanece al menos una etapa en cada fase, y también implica que si:

$$K_{r-1}(\omega) < K_r(\omega) = +\infty,$$

entonces no vamos más allá de la r -ésima fase.

Definimos la sucesión de variables aleatorias: $(r_k)_{k \geq 1}$, siendo r_k el número de fase en el que nos encontramos en la k -ésima etapa del muestreo, es decir:

$$r_k = s \quad \text{si } K_{s-1} < k \leq K_s, \text{ para } s \geq 1.$$

Recordemos la expresión (2.14):

$$w_{k,j}^{r,s} = \sum_{i=1}^n e_i^{(r)} X_{j,i}^{(B_k^s)}.$$

Las cantidades $X_{j,i}^{(B_k^s)}$ representan los coeficientes de las combinaciones lineales que permiten expresar los errores de estimación de los costes reducidos de la base B_k^s de la s -ésima fase. Como el número de posibles bases en cada fase es finito, resulta que las cantidades $X_{j,i}^{(B_k^s)}$ están acotadas, para cada s fijo, existiendo pues una constante C_s tal que:

$$|X_{j,i}^{(B_k^s)}| \leq C_s,$$

para toda variable j , todo índice i , y toda base B_k^s de la s -ésima fase. Esto equivale a la propiedad, que ya ha sido comentada anteriormente, de que en cada fase, el problema es de naturaleza finita. Las condiciones suficientes de convergencia a cero de los errores de estimación están basadas en el comportamiento asintótico de la sucesión $(C_s)_{s \geq 1}$, que suponemos creciente.

Analicemos en primer lugar la convergencia a cero de los errores de estimación de los costes reducidos de las n variables originales, siendo la convergencia para las demás variables que se van añadiendo una consecuencia inmediata.

Fijado $j \in \{1, \dots, n\}$, el error cometido en la estimación del coste reducido j -ésimo de la base $B_k^{r_k}$ es:

$$W_{k,j}^{r_k} = \frac{1}{k} \sum_{r=1}^k w_{k,j}^{r,r_k}.$$

El siguiente resultado muestra que, prescindiendo de la propiedad de diferencia de martingalas de los errores muestrales, pero manteniendo la de ortogonalidad, se puede obtener la convergencia en probabilidad a cero de los errores.

PROPOSICIÓN 2.6: *Si eliminamos la hipótesis de diferencia de martingalas y suponemos que las variables $(e_i^{(k)})_{k \geq 1}$ son ortogonales, para cada índice $i \in \{1, \dots, n\}$, y además:*

$$C_k = o(k^{1/2}),$$

entonces:

$$W_{k,j}^{r_k} \xrightarrow{P} 0,$$

cuando $k \rightarrow +\infty$, para cada $j \in \{1, \dots, n\}$.

DEMOSTRACIÓN: Se cumple que:

$$\begin{aligned} W_{k,j}^{r_k} &= \frac{1}{k} \sum_{r=1}^k w_{k,j}^{r,r_k} \\ &= \frac{1}{k} \sum_{r=1}^k \sum_{i=1}^n e_i^{(r)} X_{j,i}^{(B_k^{r_k})} \\ &= \sum_{i=1}^n X_{j,i}^{(B_k^{r_k})} \frac{1}{k} \sum_{r=1}^k e_i^{(r)}, \end{aligned}$$

de donde:

$$\begin{aligned} |W_{k,j}^{r_k}| &\leq \sum_{i=1}^n |X_{j,i}^{(B_k^{r_k})}| \cdot \left| \frac{1}{k} \sum_{r=1}^k e_i^{(r)} \right| \\ &\leq \sum_{i=1}^n C_{r_k} \left| \frac{1}{k} \sum_{r=1}^k e_i^{(r)} \right|. \end{aligned}$$

La propiedad esencial es la siguiente:

$$C_{r_k} \leq C_k,$$

ya que, en la k -ésima etapa del muestreo, nunca estaremos en una fase posterior a la k -ésima. Para probar el resultado enunciado es suficiente con demostrar que, para cada $i \in \{1, \dots, n\}$ fijo, se cumple que:

$$\frac{C_k}{k} \sum_{r=1}^k e_i^{(r)} \xrightarrow{p} 0,$$

cuando $k \rightarrow +\infty$, lo cual se probará utilizando la desigualdad de Tchebychev.

Dado $\varepsilon > 0$, se verifica:

$$\Pr \left\{ \left| \frac{C_k}{k} \sum_{r=1}^k e_i^{(r)} \right| \geq \varepsilon \right\} \leq \frac{C_k^2}{k^2 \varepsilon^2} \mathbf{E} \left(\sum_{r=1}^k e_i^{(r)} \right)^2.$$

Por la ortogonalidad de las variables $e_i^{(r)}$ y por la acotación uniforme de sus segundos momentos, resulta que existe una constante D tal que:

$$\mathbf{E} \left(\sum_{r=1}^k e_i^{(r)} \right)^2 \leq kD,$$

y por tanto:

$$\Pr \left\{ \left| \frac{C_k}{k} \sum_{r=1}^k e_i^{(r)} \right| \geq \varepsilon \right\} \leq \frac{C_k^2}{k} \cdot \frac{D}{\varepsilon^2},$$

de donde se deduce el resultado buscado.

Q.E.D.

Volviendo a hacer la hipótesis de diferencia de martingalas para la sucesión:

$$(e_i^{(k)})_{k \geq 1},$$

y abarcando pues el caso de ortogonalidad, se cumple otro resultado. En el enunciado no se menciona la hipótesis de diferencia de martingalas por ser la inicialmente supuesta.

PROPOSICIÓN 2.7: Si la sucesión $(C_k)_{k \geq 1}$ verifica que:

$$\sum_{k=1}^{+\infty} \frac{C_k^2}{k^2} < +\infty,$$

entonces se cumple:

$$W_{k,j}^{r_k} \xrightarrow{\text{cs}} 0,$$

cuando $k \rightarrow +\infty$, para cada $j \in \{1, \dots, n\}$.

DEMOSTRACIÓN: Siguiendo los mismos pasos que en la demostración de la Proposición 2.6, se concluye que basta con probar que:

$$\frac{C_k}{k} \sum_{r=1}^k e_i^{(r)} \xrightarrow{\text{cs}} 0,$$

cuando $k \rightarrow +\infty$, lo cual, teniendo en cuenta las propiedades de la sucesión $(e_i^{(k)})_{k \geq 1}$, es una consecuencia del Teorema 13.11 de IBARROLA et al., cf. p.407, tomando, en la notación allí empleada: $B_k = k/C_k$.

Q.E.D.

La hipótesis sobre la sucesión $(C_k)_{k \geq 1}$ se cumple, por ejemplo, si existe un número $\alpha > 0$ tal que:

$$C_k = O(k^{(1/2)-\alpha}),$$

expresión que permite comparar las condiciones de las proposiciones 2.6 y 2.7.

Hay que observar que lo que verdaderamente garantiza las convergencias probadas en las dos proposiciones anteriores es la convergencia a cero, en probabilidad o casi seguramente, de:

$$\frac{C_{r_k}}{k} \sum_{r=1}^k e_i^{(r)},$$

habiéndose probado la de:

$$\frac{C_k}{k} \sum_{r=1}^k e_i^{(r)},$$

que se corresponde, de algún modo, con el peor caso posible, que es cuando se cambia de fase en cada etapa. Si, por ejemplo, se pudiese suponer que el número de fases está acotado, entonces las proposiciones 2.6 y 2.7 se verificarían con total generalidad. Así pues, las condiciones:

$$C_k = o(k^{1/2}) \quad \text{y} \quad \sum_{k=1}^{+\infty} \frac{C_k^2}{k^2} < +\infty$$

son las más exigentes bajo las que se cumplen los resultados de convergencia probados.

Existen remedios sencillos para rebajar estas condiciones sobre la sucesión $(C_k)_{k \geq 1}$, como es aplicar un método de retardo en el cambio de fase. Por ejemplo, se puede imponer que la sucesión de tiempos de parada $(K_r)_{r \geq 1}$ cumpla:

$$K_r \geq 2^r,$$

es decir, no se pasa a la fase $r + 1$ con menos de 2^r observaciones. Sin entrar en el detalle de los cálculos, aplicando este retardo exponencial, la condición:

$$C_k = o(k^{1/2})$$

de la Proposición 2.6 se convierte en:

$$C_k = o(2^{k/2}),$$

que es mucho menos exigente, cubriendo, entre otros, todos los casos de crecimientos potenciales:

$$C_k \sim k^a.$$

Por tanto, las condiciones impuestas a $(C_k)_{k \geq 1}$ no son realmente restrictivas, y se pueden llegar a cumplir utilizando un retardo de cambio de fase adecuado. Sin embargo, por no complicar la exposición, no impondremos ningún retardo.

Hagamos un comentario sobre la técnica de demostración de la Proposición 2.7. La base está en el análisis de:

$$\sum_{r=1}^k e_i^{(r)} X_{j,i}^{(B_k^{rk})}, \quad (2.16)$$

que no es el último término de un arreglo triangular de martingalas, como hubiese resultado cómodo para la demostración. Sin embargo, sí que se ha

podido acotar (2.16) por el último término de un arreglo triangular de martingalas, para el que se ha utilizado una ley fuerte de los grandes números. Por tanto, la Teoría de Martingalas no se aplica directamente a los términos de error sino a cotas superiores de los citados términos de error.

La Proposición 2.7 analiza la convergencia de los errores cometidos en la estimación de los costes reducidos de una de las n variables originales. No se ha estudiado aún el error para las variables que son añadidas— una en cada nueva fase— al problema de programación lineal. Fijemos un número de fase: $s > 1$, tal que:

$$\{\omega \in \Omega : K_{s-1}(\omega) < +\infty\} \neq \emptyset,$$

de forma que existen trayectorias en las que se alcanza la fase s . A lo largo de las fases s y siguientes, si las hubiere, se ha añadido una nueva variable al problema de programación lineal: la variable $(n + s - 1)$ -ésima, cuyos costes reducidos son estimados únicamente a partir de la etapa $(K_{s-1} + 1)$ -ésima del muestreo, siendo esta etapa finita para las trayectorias consideradas. El error que se comete en la estimación del $(n + s - 1)$ -ésimo coste reducido es:

$$W_{k,n+s-1}^{r_k} = \frac{1}{k} \sum_{r=1}^k w_{k,n+s-1}^{r,r_k},$$

para $k > K_{s-1}$. Se cumple el siguiente resultado de convergencia.

PROPOSICIÓN 2.8: *Si se verifica la condición:*

$$\sum_{k=1}^{+\infty} \frac{C_k^2}{k^2} < +\infty,$$

entonces:

$$W_{k,n+s-1}^{r_k} \xrightarrow{\text{cs}} 0,$$

cuando $k \rightarrow +\infty$, en el conjunto de trayectorias que llegan hasta la s -ésima fase.

DEMOSTRACIÓN: Es la misma que la de la Proposición 2.7, con la salvedad de que el índice k es mayor que K_{s-1} , y al tratarse de un límite casi seguro, se conserva la convergencia. Q.E.D.

Antes de estudiar la convergencia del algoritmo EE, necesitamos un último resultado. Las proposiciones 2.7 y 2.8 prueban que, para cada variable,

los errores que se cometen en las estimaciones de sus costes reducidos tienden casi seguramente a cero. Lo que ocurre es que a medida que avanza el muestreo, el número de variables va incrementándose, con lo cual estamos tratando con una cantidad que tiende a $+\infty$ de variables que tienden a cero. La pregunta es si el máximo error que se comete, es decir, el máximo de las citadas variables, también converge a cero.

TEOREMA 2.9: *Si se cumple que:*

$$\sum_{k=1}^{+\infty} \frac{C_k^2}{k^2} < +\infty,$$

entonces:

$$\max \left\{ |W_{k,j}^{r_k}| : 1 \leq j \leq n + r_k - 1 \right\} \xrightarrow{cs} 0,$$

cuando $k \rightarrow +\infty$.

DEMOSTRACIÓN: Es inmediata sin más que observar que las cotas que se obtuvieron en la demostración de la Proposición 2.6 son uniformes en j .

Q.E.D.

Una observación importante es que, cuando estamos en la fase s , de los $n + s - 1$ costes reducidos, $m + s - 1$ de ellos— los correspondientes a las variables básicas— son nulos y son estimados con error cero. Por tanto, de las cantidades $|W_{k,j}^{r_k}|$, para $1 \leq j \leq n + r_k - 1$, hay a lo sumo $n - m$ positivas, y pese a que el número de variables crezca, el número de costes reducidos que realmente estimamos es fijo, independientemente de la fase en la que nos encontremos.

Convergencia de la estimación

Fijado un entero: $s \geq 1$, que designa el número de fase, se define:

$$\delta_s = \min_{B \in \mathcal{B}^s} \min_{1 \leq j \leq n+s-1} \{|z_j^B - c_j| : z_j^B - c_j \neq 0\}, \quad (2.17)$$

donde \mathcal{B}^s , que es el conjunto de bases factibles de las s -ésimas fases, es un conjunto finito. Se cumple que $\delta_s > 0$. Hagamos la siguiente hipótesis:

$$\inf_{s \geq 1} \delta_s > 0, \quad (2.18)$$

y tomemos δ tal que:

$$\inf_{s \geq 1} \delta_s > \delta > 0.$$

A continuación, proponemos una elección de los tiempos de cambio de fase que hace que el algoritmo EE sea convergente. En la etapa k -ésima del muestreo, sea $\hat{x}^{(k)}$ la solución básica identificada por el algoritmo.

TEOREMA 2.10: Para $r \geq 0$, se define:

$$K_{r+1} = \min \left\{ k > K_r : \max_{1 \leq j \leq n+r_k-1} T_{k,j}^{r_k} \leq \delta/2 \text{ y } \hat{x}^{(k)} \text{ no es entera} \right\}.$$

Si se verifican las hipótesis (2.18) y:

$$\sum_{k=1}^{+\infty} \frac{C_k^2}{k^2} < +\infty,$$

entonces se cumple que, con probabilidad uno, existe un tamaño muestral a partir del cual $\hat{x}^{(k)}$ es una solución óptima de (PE).

DEMOSTRACIÓN: Tomemos una trayectoria $\omega \in \Omega$ para la que se verifique el Teorema 2.9.

Existe un tamaño muestral k_0 a partir del cual:

$$\max \left\{ |W_{k,j}^{r_k}| : 1 \leq j \leq n + r_k - 1 \right\} < \delta/2.$$

Probemos que para la trayectoria fijada, únicamente existe una cantidad finita de tiempos: $K_r(\omega)$, para $r \geq 1$, que son finitos. Lo haremos por reducción al absurdo, admitiendo que:

$$\forall r \geq 1, K_r(\omega) < +\infty.$$

Bajo esta suposición, hay pues una cantidad infinita de cambios de fase, produciéndose esos cambios en instantes en los que:

$$\max_{1 \leq j \leq n+r_k-1} T_{k,j}^{r_k} \leq \delta/2.$$

Para un instante de cambio de fase posterior a la etapa k_0 -ésima del muestreo, llamémoslo k , se verifica, para toda variable:

$$j \in \{1, \dots, n + r_k - 1\}$$

que:

$$T_{k,j}^{r_k} = z_j^{B_k^{r_k}} - c_j + W_{k,j}^{r_k},$$

y puesto que:

$$T_{k,j}^{r_k} \leq \delta/2 \quad \text{y} \quad |W_{k,j}^{r_k}| < \delta/2,$$

resulta que:

$$z_j^{B_k^{r_k}} - c_j < \delta < \delta_{r_k},$$

y se deduce de (2.17) que necesariamente, para toda variable j :

$$z_j^{B_k^{r_k}} - c_j \leq 0.$$

En conclusión, en los instantes de cambio de fase posteriores a la etapa de muestreo k_0 -ésima, las soluciones básicas son óptimas para el problema de programación lineal continua considerado en la fase que termina. Ninguno de los planos de corte introducidos previamente ha eliminado soluciones factibles enteras, aunque no hayan sido introducidos en bases óptimas. En las fases posteriores a la observación k_0 -ésima se está resolviendo el problema continuo, y no sólo estimando su solución. Es decir, de una fase en adelante se está aplicando el algoritmo de planos de corte, y no ya una aproximación de éste, y por la hipótesis de convergencia de dicho algoritmo, resulta contradictorio que todos los instantes:

$$(K_r(\omega))_{r \geq 1}$$

sean finitos, pues querría también decir que aplicando el algoritmo de planos de corte encontramos, para todos los sucesivos problemas continuos, soluciones óptimas que no son enteras, lo cual no es compatible con la convergencia finita del algoritmo de programación entera de planos de corte.

Hemos probado pues que existe una última fase: s_0 , es decir, un primer natural tal que:

$$K_{s_0}(\omega) = +\infty.$$

Se verifica entonces que, para toda etapa del muestreo: $k > K_{s_0-1}(\omega)$, se cumple que: $\hat{x}^{(k)}$ es entera, o bien:

$$\max \left\{ T_{k,j}^{r_k} : 1 \leq j \leq n + r_k - 1 \right\} > \delta/2. \quad (2.19)$$

Repitiendo una demostración análoga a la del Teorema 1.12, cf. p.35, se concluye que (2.19) no puede darse más que una cantidad finita de veces.

Resulta también que tomando k lo bastante grande, la base $B_k^{s_0}$ es óptima para el problema continuo de la fase s_0 y $\hat{x}^{(k)}$, su solución óptima, es entera. Puesto que el problema de la fase s_0 se ha conseguido a partir del problema original añadiendo planos de corte que no han eliminado ninguna solución factible entera, se concluye que $\hat{x}^{(k)}$ es una solución óptima del problema entero (PE). Q.E.D.

Todos los resultados anteriores al Teorema 2.10 son ciertos para cualquier elección de $(K_r)_{r \geq 1}$, pero definiendo los cambios de fase como en el enunciado, conseguimos un algoritmo EE convergente.

Convergencia en distribución

Recordemos bajo qué condiciones se probó la convergencia en distribución de los errores para el algoritmo del problema de programación lineal continua. Se distinguían dos casos, según fuese la solución óptima del problema única o no. Si la solución óptima era única, entonces el algoritmo se acababa deteniendo en la única base óptima. Si la solución óptima no era única, entonces se aplicaba el algoritmo γ -truncado, de tal forma que el algoritmo se detenía, dependiendo de cada trayectoria, en una de las bases óptimas.

Para utilizar razonamientos similares en el problema entero, observemos lo siguiente. Aunque la solución del problema entero sea única, esta única solución puede admitir varias representaciones en bases de diferente tamaño, pertenecientes pues a fases distintas. Por tanto, técnicamente, este problema es equivalente a uno con varias soluciones óptimas, en tanto que hay varias bases óptimas, que pueden ser una cantidad infinita. Debemos pues utilizar el método del algoritmo γ -truncado.

Admitamos que se verifica (2.18), véase p.101, y tomemos γ_0 tal que:

$$0 < \gamma_0 < \delta/2.$$

Aplicando un algoritmo truncado en los distintos cambios de base del algoritmo EE, se obtiene un resultado similar al Teorema 2.10.

TEOREMA 2.11: *Supongamos que se cumple (2.18) y sea γ tal que:*

$$0 < \gamma \leq \gamma_0.$$

Para $r \geq 0$, se define:

$$K_{r+1} = \min \left\{ k > K_r : \max_{1 \leq j \leq n+r_k-1} T_{k,j}^{r_k} \leq \gamma \text{ y } \hat{x}^{(k)} \text{ no es entera} \right\}.$$

Si se verifica que:

$$\sum_{k=1}^{+\infty} \frac{C_k^2}{k^2} < +\infty,$$

y se aplica el algoritmo γ -truncado, entonces se tiene que, con probabilidad uno, existe un tamaño muestral a partir del cual $\hat{x}^{(k)}$ es una solución óptima de (PE).

DEMOSTRACIÓN: La demostración es igual que la del Teorema 2.10 y se inspira también en la del Teorema 1.17, cf. p.46. Q.E.D.

Los cambios de fase se producen pues cuando el algoritmo truncado para el problema continuo repite un vértice que no es entero.

La peculiaridad del algoritmo γ -truncado es que, dada una trayectoria para la que se cumple el Teorema 2.9, existe una base óptima, que tiene asociada una solución entera óptima, en la que se detiene el algoritmo. Como el conjunto de posibles bases, teniendo en cuenta las diferentes fases, es numerable, el conjunto de bases cuya solución básica es una solución óptima de (PE) es también numerable. Sea \mathcal{B}^* el conjunto de esas bases óptimas. El conjunto medible:

$$\Gamma_B,$$

para cada $B \in \mathcal{B}^*$, es el conjunto de $\omega \in \Omega$ tales que el algoritmo EE se detiene en la base B . Se tiene que los conjuntos Γ_B son disjuntos y que:

$$\Pr\left\{\bigcup_{B \in \mathcal{B}^*} \Gamma_B\right\} = 1.$$

Dado $j \geq 1$, queremos determinar la distribución límite de:

$$\left(\sqrt{k}W_{k,j}^{r_k}\right)_{k \geq 1}. \quad (2.20)$$

Si la variable j no pertenece a la r_k -ésima fase, es decir, si:

$$j \geq n + r_k,$$

entonces se define el error $W_{k,j}^{r_k}$ como igual a cero, lo que se consigue haciendo:

$$X_{j,i}^B = 0$$

cuando la variable j no pertenece a la fase de la base B . Construyamos, en primer lugar, una sucesión asintóticamente equivalente en distribución a (2.20).

LEMA 2.12: *Bajo las condiciones del Teorema 2.11, se verifica, para cada variable $j \geq 1$, que:*

$$\sqrt{k}W_{k,j}^{rk} - \sum_{B \in \mathcal{B}^*} I_{\Gamma_B} \sum_{i=1}^n X_{j,i}^B \frac{1}{\sqrt{k}} \sum_{r=1}^k e_i^{(r)} \xrightarrow{\text{cs}} 0,$$

cuando $k \rightarrow +\infty$.

DEMOSTRACIÓN: Es una consecuencia inmediata del Teorema 2.11 y de las expresiones para $W_{k,j}^{rk}$ que se obtuvieron en la Proposición 2.6.

Q.E.D.

Apoyándonos en este lema, será suficiente con determinar la distribución límite de:

$$\sum_{B \in \mathcal{B}^*} I_{\Gamma_B} \sum_{i=1}^n X_{j,i}^B \frac{1}{\sqrt{k}} \sum_{r=1}^k e_i^{(r)},$$

lo cual se hará estudiando la convergencia en ley de cada término:

$$\sum_{i=1}^n X_{j,i}^B \frac{1}{\sqrt{k}} \sum_{r=1}^k e_i^{(r)},$$

para $B \in \mathcal{B}^*$. Obsérvese, y es un resultado inmediato, que la sucesión de índice doble:

$$Z_{kr}^B = \sum_{i=1}^n X_{j,i}^B \frac{1}{\sqrt{k}} e_i^{(r)},$$

para $1 \leq r \leq k$, $k \geq 1$, y fijado $j \geq 1$, es un arreglo triangular de diferencias de martingalas de media cero y de cuadrado integrable para las filtraciones:

$$(\sigma(\mathbf{Y}_1, \dots, \mathbf{Y}_r))_{1 \leq r \leq k}.$$

El teorema central del límite para el algoritmo EE es el siguiente, y su demostración resulta bastante técnica.

TEOREMA 2.13: *Fijado $j \geq 1$, bajo las hipótesis del Teorema 2.11 y si, para cada $B \in \mathcal{B}^*$, el arreglo triangular (Z_{kr}^B) verifica las condiciones del Teorema 1.19, cf. p.50, entonces la sucesión:*

$$\left(\sqrt{k}W_{k,j}^{rk} \right)_{k \geq 1}$$

converge en distribución a una variable aleatoria con función característica:

$$\mathbf{E} \exp\left\{-\frac{1}{2}\eta_{\mathbf{B}}^2 t^2\right\},$$

para $t \in \mathbb{R}$, donde \mathbf{B} es la variable que toma el valor $B \in \mathcal{B}^*$ en Γ_B , y la variable aleatoria $\eta_{\mathbf{B}}^2$ es el límite en probabilidad de $\sum_{r=1}^k (Z_{kr}^B)^2$.

DEMOSTRACIÓN: Sea $B \in \mathcal{B}^*$. Supongamos, como indica el enunciado, que:

$$\sum_{r=1}^k (Z_{kr}^B)^2 \xrightarrow{P} \eta_B^2.$$

Se cumple entonces que:

$$\sum_{i=1}^n X_{j,i}^B \frac{1}{\sqrt{k}} \sum_{r=1}^k e_i^{(r)}$$

converge en distribución a una variable aleatoria con función característica:

$$\mathbf{E} \exp\left\{-\frac{1}{2}\eta_B^2 t^2\right\},$$

en la variable $t \in \mathbb{R}$.

Al igual que se aplicó en el Teorema 1.26, cf. p.66, se verifica que:

$$\lim_{k \rightarrow +\infty} \mathbf{E}(I_{\Gamma_B} \exp\left\{it \sum_{i=1}^n X_{j,i}^B \frac{1}{\sqrt{k}} \sum_{r=1}^k e_i^{(r)}\right\}) = \mathbf{E}(I_{\Gamma_B} \exp\left\{-\frac{1}{2}\eta_B^2 t^2\right\}),$$

para $t \in \mathbb{R}$, es decir:

$$\lim_{k \rightarrow +\infty} \mathbf{E}(I_{\Gamma_B} \exp\left\{it \sum_{r=1}^k Z_{kr}^B\right\}) = \mathbf{E}(I_{\Gamma_B} \exp\left\{-\frac{1}{2}\eta_B^2 t^2\right\}). \quad (2.21)$$

Probemos el siguiente resultado de intercambio de límite y sumatorio:

$$\sum_{B \in \mathcal{B}^*} \lim_{k \rightarrow +\infty} \mathbf{E}(I_{\Gamma_B} \exp\left\{it \sum_{r=1}^k Z_{kr}^B\right\}) = \lim_{k \rightarrow +\infty} \sum_{B \in \mathcal{B}^*} \mathbf{E}(I_{\Gamma_B} \exp\left\{it \sum_{r=1}^k Z_{kr}^B\right\}), \quad (2.22)$$

donde se supone que las bases de \mathcal{B}^* están numeradas:

$$\mathcal{B}^* = \{B_{(1)}, B_{(2)}, \dots\}, \quad (2.23)$$

entendiéndose que el sumatorio sobre \mathcal{B}^* sigue el orden de (2.23). Admitiremos que \mathcal{B}^* es infinito, ya que de ser finito el resultado es inmediato.

El primer miembro de (2.22) verifica:

$$\begin{aligned} \sum_{a=1}^{+\infty} \lim_{k \rightarrow +\infty} \mathbf{E}(I_{\Gamma_{B(a)}} \exp\{it \sum_{r=1}^k Z_{kr}^{B(a)}\}) &= \sum_{a=1}^{+\infty} \mathbf{E}(I_{\Gamma_{B(a)}} \exp\{-\frac{1}{2}\eta_{B(a)}^2 t^2\}) \\ &= \mathbf{E} \sum_{a=1}^{+\infty} (I_{\Gamma_{B(a)}} \exp\{-\frac{1}{2}\eta_{B(a)}^2 t^2\}), \end{aligned}$$

como se deduce de (2.21) y del Teorema de la Convergencia Dominada, estando pues bien definido el primer miembro de (2.22).

Fijemos $t \in \mathbb{R}$. Dado $\varepsilon > 0$, existe un entero A_0 tal que:

$$\sum_{a > A_0} \Pr\{\Gamma_{B(a)}\} < \varepsilon/3.$$

Existe también un entero k_0 tal que para $k > k_0$:

$$\left| \sum_{a=1}^{A_0} \mathbf{E}(I_{\Gamma_{B(a)}} \exp\{it \sum_{r=1}^k Z_{kr}^{B(a)}\}) - \sum_{a=1}^{A_0} \mathbf{E}(I_{\Gamma_{B(a)}} \exp\{-\frac{1}{2}\eta_{B(a)}^2 t^2\}) \right| < \varepsilon/3, \quad (2.24)$$

en virtud de (2.21). Se verifica que:

$$\left| \sum_{a=1}^{+\infty} \mathbf{E}(I_{\Gamma_{B(a)}} \exp\{-\frac{1}{2}\eta_{B(a)}^2 t^2\}) - \sum_{a=1}^{A_0} \mathbf{E}(I_{\Gamma_{B(a)}} \exp\{-\frac{1}{2}\eta_{B(a)}^2 t^2\}) \right|$$

es igual a:

$$\sum_{a > A_0} \mathbf{E}(I_{\Gamma_{B(a)}} \exp\{-\frac{1}{2}\eta_{B(a)}^2 t^2\}),$$

que, puesto que las funciones exponenciales están acotadas por uno, es menor o igual que:

$$\sum_{a > A_0} \Pr\{\Gamma_{B(a)}\} < \varepsilon/3,$$

luego:

$$\left| \sum_{a=1}^{+\infty} \mathbf{E}(I_{\Gamma_{B(a)}} \exp\{-\frac{1}{2}\eta_{B(a)}^2 t^2\}) - \sum_{a=1}^{A_0} \mathbf{E}(I_{\Gamma_{B(a)}} \exp\{-\frac{1}{2}\eta_{B(a)}^2 t^2\}) \right| < \varepsilon/3. \quad (2.25)$$

Para cada $k > k_0$ se tiene, por el mismo motivo, que:

$$\left| \sum_{a=1}^{+\infty} \mathbf{E}(I_{\Gamma_{B^{(a)}}} \exp\{\mathbf{it} \sum_{r=1}^k Z_{kr}^{B^{(a)}}\}) - \sum_{a=1}^{A_0} \mathbf{E}(I_{\Gamma_{B^{(a)}}} \exp\{\mathbf{it} \sum_{r=1}^k Z_{kr}^{B^{(a)}}\}) \right| < \varepsilon/3. \quad (2.26)$$

En conclusión, para $k > k_0$, sumando (2.24), (2.25) y (2.26), se verifica que:

$$\left| \sum_{a=1}^{+\infty} \mathbf{E}(I_{\Gamma_{B^{(a)}}} \exp\{\mathbf{it} \sum_{r=1}^k Z_{kr}^{B^{(a)}}\}) - \sum_{a=1}^{+\infty} \mathbf{E}(I_{\Gamma_{B^{(a)}}} \exp\{-\frac{1}{2}\eta_{B^{(a)}}^2 t^2\}) \right| < \varepsilon,$$

quedando pues probado (2.22).

Se cumple, por el Teorema de la Convergencia Dominada, que:

$$\mathbf{E} \exp\{\mathbf{it} \sum_{B \in \mathcal{B}^*} I_{\Gamma_B} \sum_{r=1}^k Z_{kr}^B\} = \sum_{B \in \mathcal{B}^*} \mathbf{E}(I_{\Gamma_B} \exp\{\mathbf{it} \sum_{r=1}^k Z_{kr}^B\}),$$

cuyo límite, cuando $k \rightarrow +\infty$, es igual a:

$$\mathbf{E} \sum_{B \in \mathcal{B}^*} (I_{\Gamma_B} \exp\{-\frac{1}{2}\eta_B^2 t^2\}),$$

que se puede escribir como:

$$\mathbf{E} \exp\{-\frac{1}{2}\eta_{\mathcal{B}}^2 t^2\}, \quad (2.27)$$

de donde:

$$\sum_{B \in \mathcal{B}^*} I_{\Gamma_B} \sum_{r=1}^k Z_{kr}^B$$

converge en ley a una variable aleatoria con función característica (2.27), lo que, teniendo en cuenta el Lema 2.12, permite concluir el resultado.

Q.E.D.

Imponiendo condiciones específicas sobre las variables $e_i^{(r)}$ se verifican las hipótesis del Teorema 2.13.

COROLARIO 2.14: *Bajo las condiciones del Teorema 2.11 y si las variables $e_i^{(r)}$, para $1 \leq i \leq n$ y $r \geq 1$, son independientes e idénticamente distribuidas con varianza igual a σ^2 , entonces:*

$$\left(\sqrt{k} W_{k,j}^{rk} \right)_{k \geq 1}$$

converge en distribución a una variable aleatoria que tiene función característica:

$$\sum_{B \in \mathcal{B}^*} \Pr\{\Gamma_B\} \exp\left\{-\frac{1}{2}\sigma^2 \sum_{i=1}^n (X_{j,i}^B)^2 t^2\right\},$$

para $t \in \mathbb{R}$, que es una mixtura de variables aleatorias normales de media cero y de varianza:

$$\sigma^2 \sum_{i=1}^n (X_{j,i}^B)^2,$$

siendo las probabilidades de mixtura:

$$\Pr\{\Gamma_B\},$$

para $B \in \mathcal{B}^*$.

DEMOSTRACIÓN: Es en parte la misma que la de la Proposición 1.21, cf. p.53, y es inmediata una vez comprobadas las hipótesis del Teorema 1.19, cf. p.50. Q.E.D.

Con estos dos resultados concluye el estudio del algoritmo EE, habiéndose probado pues un teorema central del límite para todos los errores de estimación de los costes reducidos. Los pasos dados para demostrar el Teorema 2.13 son similares a los del Teorema 1.26, cf. p.66. Hay que hacer notar que, sin embargo, han surgido dificultades técnicas. En el Teorema 1.26, el conjunto de bases considerado era finito, mientras que en el Teorema 2.13 puede ser infinito, y esto ha añadido la complicación de tener que probar un resultado de intercambio de límite y sumatorio infinito que en el caso del Teorema 1.26 no era necesario.

2.4 Conclusiones

Tomando como punto de partida el algoritmo que se desarrolló para la resolución de un problema de programación lineal continua en el primer capítulo, se ha estudiado el problema de la estimación de la solución de un problema de programación lineal entera por el algoritmo de planos de corte de Gomory, a fin de utilizar los resultados ya conocidos en el problema continuo. Se han definido dos algoritmos de resolución del problema entero que no se basan

en la estimación de la función objetivo sino en aproximarse iterativamente, y con un aprovechamiento total de la información pasada, a las soluciones de problemas continuos.

La principal diferencia entre el problema continuo y el problema entero es que, mientras que el primero es de naturaleza finita, el segundo es de tipo infinito. Para el problema continuo solamente existía una cantidad finita de bases; para el problema entero son una cantidad infinita. El algoritmo ED evita esta problemática puesto que, tras una perturbación por potencias de ε , aproxima una solución del problema continuo y luego procede de forma determinista. Es el algoritmo EE el que se enfrenta al problema infinito. Las dificultades encontradas son, entre otras, que los términos de error ya no se pueden expresar a partir de diferencias de martingalas, aunque sí se pueden acotar por martingalas que convergen a cero. Otro inconveniente que ha surgido es que las bases óptimas son, en general, infinitas, y esto incluso en el caso de ser la solución del problema entero única, lo que ha complicado notablemente la demostración del teorema central del límite, que se basa en el concepto de convergencia débil en L_1 de funciones del tipo:

$$\exp\{itZ_k\}.$$

Todo ello se ha conseguido exigiendo condiciones relativamente débiles a los errores de estimación, que son diferencias de martingalas.

En conclusión, en este estudio del problema entero se ha puesto de manifiesto que problemas entero y continuo son, en realidad, totalmente diferentes, habiéndose resuelto las dificultades propias del problema entero utilizando técnicas específicas, no siendo por tanto este estudio una mera aplicación del llevado a cabo para el problema continuo, y teniendo pues interés por sí mismo.

CAPÍTULO 3

Problema de Parada Óptima

3.1 Introducción

3.1.1 Objetivos

Se considera un problema de parada óptima en una cadena de Markov finita con pagos no negativos. Se sabe que existe una regla de decisión óptima que se calcula a partir del valor del problema de decisión: la mínima función excesiva de la cadena de Markov.

Supondremos que las probabilidades de transición de la cadena de Markov son desconocidas. Estas probabilidades serán estimadas, a partir de la observación de las transiciones de la cadena de Markov, según la teoría de la máxima verosimilitud desarrollada en BILLINGSLEY. Se obtendrán resultados de consistencia casi segura y de distribuciones asintóticas para estas estimaciones.

Fijada una regla de decisión particular, se estudia la estimación de su valor a partir de las observaciones muestrales de las transiciones de la cadena. Dicha estimación resultará ser consistente casi seguramente y se obtendrá su distribución límite.

Se abordará el problema de la estimación de una regla de decisión óptima analizando la consistencia de esta estimación.

Se tratará también la estimación del valor del problema. Se probará la consistencia del estimador del valor del problema que se obtiene de forma natural a partir de la observación de la cadena, aunque la distribución límite no se establecerá con toda generalidad sino imponiendo ciertas condiciones

adicionales sobre los parámetros del problema. Por último, se introduce el método de la perturbación. Este método, mediante una alteración infinitesimal de los parámetros del problema, proporcionará otros estimadores del valor del problema de los que, además de probar que son consistentes casi seguramente, se verá que verifican un teorema central del límite, esta vez sin tener que imponer ninguna condición adicional sobre los parámetros.

Los resultados teóricos sobre problemas de decisión markovianos y problemas de parada óptima, así como gran parte de la notación, están tomados de PUTERMAN y en particular del Capítulo 7.

3.1.2 Notación

El sistema considerado evoluciona según una cadena de Markov y en cada estado se pueden tomar dos acciones: parar y recibir un pago según el estado en el que se encuentre el sistema, o bien continuar y efectuar una transición más. En este tipo de problemas es usual considerar que el pago en los estados recurrentes es menor o igual que cero. Nosotros supondremos además que el pago en los estados de transición de la cadena de Markov es mayor o igual que cero. En estas circunstancias, puesto que una vez que la cadena alcance los estados recurrentes lo óptimo será no detenerse nunca, podemos admitir, sin pérdida de generalidad, que el pago en los estados recurrentes es igual a cero.

Supongamos que la matriz de transición de la cadena de Markov es (p_{ij}) y que la cadena tiene estados de transición \mathcal{T} y estados recurrentes \mathcal{R} . Formamos una nueva cadena de Markov cuyo conjunto de estados es: $\mathcal{T} \cup \{s'\}$, donde el estado s' es absorbente y para $i \in \mathcal{T}$ se cumple:

$$p_{is'} = \sum_{j \in \mathcal{R}} p_{ij},$$

siendo las transiciones entre los estados de \mathcal{T} como en la cadena original. Se puede decir que hemos hecho de todos los estados recurrentes un único estado. Esta modificación de la cadena no va a afectar al problema de parada óptima puesto que una vez que el sistema entraba en los estados recurrentes, al tener éstos pago cero, era indiferente la acción que se tomara: parar o continuar.

Formulamos el problema de la siguiente manera. Sea:

$$S = \{s_1, \dots, s_m\},$$

correspondiéndose S con \mathcal{T} en lo expuesto anteriormente, y sea $\bar{S} = S \cup \{s'\}$. Usualmente se denotará s' por s_{m+1} , de forma que:

$$\bar{S} = \{s_1, \dots, s_{m+1}\}.$$

Sobre el conjunto \bar{S} se define una cadena de Markov con matriz de transición:

$$\bar{P} = \begin{pmatrix} P & T \\ 0 \cdots 0 & 1 \end{pmatrix}.$$

La matriz \bar{P} es de dimensiones $(m+1) \times (m+1)$, la matriz P es $m \times m$ y T es $m \times 1$, cumpliendo $T = \mathbf{1} - P\mathbf{1}$, donde $\mathbf{1}$ es un vector columna con cada componente igual a uno. La matriz P representa las transiciones entre los estados de transición de la cadena de Markov. Se tiene pues que:

$$P = (p_{ij})_{i,j=1,\dots,m}.$$

La matriz $T = (p_{i,m+1})_{i=1,\dots,m}$ representa las probabilidades de absorción en una etapa. Para los elementos de las matrices P y T se utilizarán indistintamente las notaciones p_{ij} y $p(j|i)$, con $1 \leq i \leq m$ y $1 \leq j \leq m+1$.

Sea $g \in \mathbb{R}^m$ el vector de pagos: $g = (g(s))_{s \in S}$, siendo, para cada $s \in S$:

$$g(s) \geq 0.$$

Los conjuntos de acciones son, para $s \in S$:

$$A_s = \{0, 1\} \quad \text{y} \quad A_{s_{m+1}} = \{0\},$$

donde la acción 0 significa “parar” y la acción 1 significa “continuar”. Las probabilidades de transición entre los estados de \bar{S} , según la acción tomada, son, cuando $1 \leq i \leq m$ y $1 \leq j \leq m+1$:

$$p(j|i, 1) = p(j|i) = p_{ij}, \quad p(m+1|i, 0) = 1, \quad p(m+1|m+1, 0) = 1,$$

y las funciones de pago son, cuando $s \in S$:

$$r(s, 1) = 0, \quad r(s, 0) = g(s), \quad \text{y} \quad r(s_{m+1}, 0) = 0.$$

Una regla de decisión d es una función cuyo dominio es \bar{S} y tal que para cada $s \in \bar{S}$ se tiene que $d(s) \in A_s$, o en otros términos: $d(s) = 0$ si la decisión que se toma en s es la de parar, o bien: $d(s) = 1$ si la decisión escogida es continuar. Nótese que siempre se tiene que:

$$d(s_{m+1}) = 0.$$

Se denota por D el conjunto de reglas de decisión. Asociada a una regla de decisión $d \in D$ consideramos la política estacionaria d^∞ que aplica en todas las épocas de decisión la regla d .

Las políticas son evaluadas según su pago esperado. El valor de la política estacionaria d^∞ , siendo el estado inicial s , se denota por:

$$v_d(s).$$

Informalmente diremos que $v_d(s)$ es el valor de d cuando el estado inicial es s . Los valores $v_d(s)$ son las componentes del vector v_d .

Tomando el supremo de los valores de todas las políticas π , pero sin necesariamente restringirnos a las políticas estacionarias, obtenemos el valor del problema de decisión:

$$v^*(s) = \sup_{\pi} v_{\pi}(s),$$

siendo $v^*(s)$ la máxima ganancia esperada cuando s es el estado inicial del sistema. Los valores $v^*(s)$ son las componentes del vector v^* .

Se puede probar que existe una regla $d \in D$ tal que su valor coincide con el valor del problema, es decir, una regla con la que se alcanza la máxima ganancia:

$$\forall s \in \bar{S}, v^*(s) = v_d(s).$$

De una tal regla d diremos que es óptima.

En ocasiones se considerará, por comodidad, que P es un vector de m^2 componentes. Las m primeras componentes de dicho vector serán las componentes de la primera fila de P , las m siguientes componentes del vector serán las de la segunda fila de P , etc.. Se supone que la matriz P es desconocida y que tanto P como T tienen todos sus elementos positivos, es decir que:

$$P \in T_m \times \cdots \times T_m,$$

siendo:

$$T_m = \{(x_1, \dots, x_m) \in \mathbb{R}^m : \sum_{i=1}^m x_i < 1 \text{ y } \forall i \in \{1, \dots, m\}, x_i > 0\}.$$

Así pues, la cadena de Markov dada por la matriz \bar{P} tiene un estado absorbente siendo todos los demás estados transitorios y comunicantes en una etapa entre sí.

La norma del supremo de un vector de \mathbb{R}^m es el máximo de los valores absolutos de sus coordenadas, es decir, la norma del supremo de:

$$\mathbf{y} = (y_i)_{1 \leq i \leq m},$$

es:

$$\|\mathbf{y}\| = \max_{1 \leq i \leq m} |y_i|.$$

Para una matriz $A = (a_{ij})$ de dimensiones $m \times m$ se define $\|A\|_B$ como:

$$\|A\|_B = \max_{1 \leq i \leq m} \sum_{j=1}^m |a_{ij}|.$$

Si C es otra matriz $m \times m$ se verifica, véase PUTERMAN, p.606, la siguiente desigualdad:

$$\|AC\|_B \leq \|A\|_B \|C\|_B.$$

Se considera un espacio de probabilidad $(\Omega, \mathcal{F}, \Pr)$.

3.2 Estimación del problema

3.2.1 Estimación de las probabilidades de transición

La estimación de la matriz P se hará por el método de la máxima verosimilitud. Los resultados sobre comportamiento límite y distribución asintótica del estimador de máxima verosimilitud están enunciados para cadenas ergódicas. La cadena con matriz de transición \bar{P} tiene un único estado recurrente y por tanto no se le pueden aplicar estos resultados. Para convertirla en ergódica es necesario que cada vez que el sistema esté en el estado s_{m+1} vuelva a un estado de transición, es decir, a un estado de S . Consideramos pues una cadena de Markov cuya matriz de transición sea:

$$\tilde{P} = \begin{pmatrix} P & T \\ q_1 \cdots q_m & 0 \end{pmatrix}.$$

Esto supone que el muestreo que se va a llevar a cabo es tal que cuando se produzca la absorción por el estado s_{m+1} se reinicializa la cadena en un estado de transición obtenido mediante el sorteo dado por las probabilidades q_1, \dots, q_m . Para cualquier elección de las probabilidades q_1, \dots, q_m esta

cadena es irreducible y aperiódica. Se supone que q_1, \dots, q_m están fijados y son conocidos, y por tanto sólo hay que estimar la matriz P , que pertenece al espacio paramétrico $T_m \times \dots \times T_m$, que es un abierto de \mathbb{R}^{m^2} . Suponemos que se parte de un estado cualquiera x_1 y que la sucesión de estados observada, después de n transiciones, es:

$$x_1, \dots, x_n, x_{n+1}.$$

Se definen las variables aleatorias n_{ij} , para $1 \leq i \leq m$ y $1 \leq j \leq m$, como el número de transiciones observadas del estado s_i al estado s_j y se define también n_i , para $1 \leq i \leq m$, como el número de veces que, en las etapas de muestreo $1, \dots, n$, el sistema ha estado en s_i . Los estimadores de máxima verosimilitud son:

$$\hat{p}_{ij}^{(n)} = \frac{n_{ij}}{n_i},$$

para $1 \leq i \leq m$ y $1 \leq j \leq m$. En el caso de ser $n_i = 0$, el estimador de máxima verosimilitud es cualquiera. Se denota la estimación de máxima verosimilitud de P por $\hat{P}^{(n)}$, que es la matriz $m \times m$ formada con las estimaciones de máxima verosimilitud de las componentes de P . Sea $(p_1, \dots, p_m, p_{m+1})$ la distribución estacionaria de la cadena con matriz de transición \tilde{P} .

TEOREMA 3.1: *El estimador de P : $\hat{P}^{(n)}$, es consistente casi seguramente cuando $n \rightarrow +\infty$.*

DEMOSTRACIÓN: Está basada en el Teorema 1.1, p.6, de BILLINGSLEY. Fijados i y j de $\{1, \dots, m\}$ se definen las funciones η y ξ , para cualesquiera dos estados x_1 y x_2 de \bar{S} , como:

$$\eta(x_1, x_2) = \begin{cases} 1, & \text{si } x_1 = s_i \text{ y } x_2 = s_j, \\ 0, & \text{en los demás casos,} \end{cases}$$

y:

$$\xi(x_1, x_2) = \begin{cases} 1, & \text{si } x_1 = s_i, \\ 0, & \text{en los demás casos.} \end{cases}$$

Se cumple que:

$$\frac{1}{n} \sum_{k=1}^n \eta(x_k, x_{k+1}) \xrightarrow{\text{cs}} \mathbf{E}\eta(x_1, x_2),$$

cuando $n \rightarrow +\infty$, donde la esperanza está tomada respecto de la distribución estacionaria de la cadena, es decir, suponiendo que la distribución de x_1 es

la distribución estacionaria de \tilde{P} ; un resultado análogo es cierto para ξ . Se deduce de estos resultados que:

$$\frac{n_{ij}}{n} \xrightarrow{\text{cs}} p_i p_{ij} \quad \text{y} \quad \frac{n_i}{n} \xrightarrow{\text{cs}} p_i.$$

Puesto que la cadena de Markov con matriz de transición \tilde{P} es ergódica, se tiene que $p_i > 0$, y por tanto, con probabilidad uno, a partir de un cierto n se tiene que $n_i > 0$, y se cumple que:

$$\hat{p}_{ij}^{(n)} = \frac{n_{ij}}{n_i} \xrightarrow{\text{cs}} p_{ij},$$

cuando $n \rightarrow +\infty$, de donde $\hat{P}^{(n)}$ converge casi seguramente a P . Q.E.D.

Para estudiar la distribución límite debemos determinar la matriz de información de Fisher. Consideramos i y j verificando:

$$1 \leq i \leq m \quad \text{y} \quad 1 \leq j \leq m,$$

y sean x_1 y x_2 la primera y segunda observación del sistema, respectivamente, que son por tanto realizaciones de variables aleatorias. Calculamos la derivada del logaritmo de la verosimilitud respecto de cada una de las variables del espacio paramétrico. La probabilidad de que en la cadena con matriz de transición \tilde{P} se efectúe una transición del estado x_1 al x_2 se escribirá $\tilde{p}_{x_1 x_2}$. Se cumple que:

$$\frac{\partial \log \tilde{p}_{x_1 x_2}}{\partial p_{ij}} = \begin{cases} 1/p_{ij}, & \text{si } (s_i, s_j) = (x_1, x_2), \\ -1/p_{i \ m+1}, & \text{si } s_i = x_1 \text{ y } s_{m+1} = x_2, \\ 0, & \text{en los demás casos.} \end{cases}$$

Si fijamos otro par (i', j') de enteros que cumplan:

$$1 \leq i' \leq m \quad \text{y} \quad 1 \leq j' \leq m,$$

se tiene que:

$$\frac{\partial \log \tilde{p}_{x_1 x_2}}{\partial p_{ij}} \frac{\partial \log \tilde{p}_{x_1 x_2}}{\partial p_{i' j'}} = \begin{cases} 1/p_{ij}^2, & \text{si } (s_i, s_j) = (s_{i'}, s_{j'}) = (x_1, x_2), \\ 1/p_{i \ m+1}^2, & \text{si } s_i = s_{i'} = x_1 \text{ y } s_{m+1} = x_2, \\ 0, & \text{en los demás casos.} \end{cases}$$

Por tanto, condicionando por el valor de x_1 , se verifica que el valor de:

$$\mathbf{E}_{x_1} \left[\frac{\partial \log \tilde{p}_{x_1 x_2}}{\partial p_{ij}} \frac{\partial \log \tilde{p}_{x_1 x_2}}{\partial p_{i' j'}} \right]$$

es:

$$\begin{cases} 1/p_{ij} + 1/p_{i\ m+1}, & \text{si } s_i = s_{i'} = x_1 \text{ y } j = j', \\ 1/p_{i\ m+1}, & \text{si } s_i = s_{i'} = x_1 \text{ y } j \neq j', \\ 0, & \text{en los demás casos.} \end{cases}$$

Si tomamos esperanzas cuando la distribución de x_1 es la distribución estacionaria de la cadena con matriz de transición \tilde{P} , se tiene que el valor de:

$$\mathbf{E} \left[\frac{\partial \log \tilde{p}_{x_1 x_2}}{\partial p_{ij}} \frac{\partial \log \tilde{p}_{x_1 x_2}}{\partial p_{i'j'}} \right]$$

es:

$$\begin{cases} p_i(1/p_{ij} + 1/p_{i\ m+1}), & \text{si } i = i' \text{ y } j = j', \\ p_i/p_{i\ m+1}, & \text{si } i = i' \text{ y } j \neq j', \\ 0, & \text{si } i \neq i'. \end{cases}$$

La matriz de información de Fisher¹: $I_{(i,j),(i',j')}(P)$, que es una matriz de dimensiones $m^2 \times m^2$, donde se supone que los parámetros p_{ij} están ordenados primero según i y luego según j , es diagonal por cajas:

$$I(P) = \begin{pmatrix} p_1 C_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & p_2 C_2 & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & p_m C_m \end{pmatrix}.$$

Las matrices C_i , que son $m \times m$, vienen dadas por:

$$C_i = \left(\delta_{jj'} \frac{1}{p_{ij}} + \frac{1}{p_{i\ m+1}} \right)_{1 \leq j, j' \leq m},$$

siendo sus inversas:

$$C_i^{-1} = (p_{ij}(\delta_{jj'} - p_{ij'}))_{1 \leq j, j' \leq m},$$

y la inversa de la matriz de información es:

$$I(P)^{-1} = \begin{pmatrix} \frac{1}{p_1} C_1^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \frac{1}{p_2} C_2^{-1} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \mathbf{0} & \cdots & \frac{1}{p_m} C_m^{-1} \end{pmatrix}.$$

¹Esta matriz depende de P y de q_1, \dots, q_m aunque no se hace explícita la dependencia de los parámetros q_1, \dots, q_m por suponerse fijados de antemano.

Proponemos a continuación una interpretación de la matriz de Fisher. La matriz de información es diagonal por cajas y cada caja corresponde a la matriz de información de un muestreo multinomial de probabilidades (p_i) , estando esta matriz afectada del coeficiente p_i , probabilidad estacionaria del estado s_i . Esto quiere decir que lo que estamos haciendo son varios muestreos multinomiales independientes y cada uno de ellos lo hacemos con una probabilidad—límite— p_i , ya que del muestreo correspondiente a (p_i) se toma una muestra siempre que la cadena esté en el estado s_i , lo que ocurre, en el límite, con dicha probabilidad. Hay que observar que la suma de estas probabilidades es: $p_1 + \dots + p_m < 1$, ya que los muestreos multinomiales necesarios para la estimación del parámetro son los correspondientes a los m primeros estados de \bar{S} .

El siguiente resultado proporciona la distribución límite del estimador de máxima verosimilitud de P .

TEOREMA 3.2: *El estimador $\hat{P}^{(n)}$ cumple:*

$$\sqrt{n} \left(\hat{P}^{(n)} - P \right) \xrightarrow{d} \mathcal{N} \left(\mathbf{0}, I(P)^{-1} \right),$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: Comprobamos, para la cadena cuya matriz de transición es \tilde{P} , la condición 5.1, p.23, de BILLINGSLEY. El conjunto D de pares (i, j) , con $1 \leq i \leq m+1$ y $1 \leq j \leq m+1$, tales que $\tilde{p}_{ij} > 0$, siendo \tilde{p}_{ij} el término (i, j) de \tilde{P} , no depende del valor de P , y se cumple que:

$$D = \{1, \dots, m+1\}^2 - \{(m+1, m+1)\},$$

que tiene $d = (m+1)^2 - 1$ elementos. Las funciones \tilde{p}_{ij} son de clase tres en todo el espacio paramétrico $T_m \times \overset{m}{\dots} \times T_m$. La matriz de término general:

$$\frac{\partial \tilde{p}_{ij}}{\partial p_{i'j'}},$$

siendo $(i, j) \in D$, $1 \leq i' \leq m$ y $1 \leq j' \leq m$, tiene rango m^2 , que es la dimensión del espacio paramétrico. Por último, para cualquier valor del parámetro P , la cadena tiene un único conjunto ergódico y no hay estados de transición. Se cumple pues dicha condición 5.1.

Estamos en las hipótesis del Teorema 2.2, p.13, de BILLINGSLEY y es cierto el resultado enunciado. Q.E.D.

Nótese que la distribución límite del estimador depende de la distribución invariante de la cadena. Se puede obtener otro estimador cuya distribución límite no depende de la distribución estacionaria de la cadena y sólo depende de P , siendo por tanto la distribución límite de este estimador independiente de la elección de las probabilidades q_1, \dots, q_m .

PROPOSICIÓN 3.3: *Se verifica, si $1 \leq i \leq m$, el siguiente límite:*

$$\sqrt{n_i} \left(\hat{p}_{ij}^{(n)} - p_{ij} \right)_{1 \leq j \leq m} \xrightarrow{d} \mathcal{N} \left(\mathbf{0}, C_i^{-1} \right),$$

cuando $n \rightarrow +\infty$, siendo las distribuciones límite independientes.

DEMOSTRACIÓN: Por distribuciones límite independientes se entiende que el límite en ley conjunto, al variar i , es una normal cuya matriz de covarianzas es diagonal por cajas, con cajas C_i^{-1} . Se demostró en el Teorema 3.1 que:

$$\frac{n_i}{n} \xrightarrow{cs} p_i,$$

cuando $n \rightarrow +\infty$, de donde:

$$\sqrt{\frac{n_i}{np_i}} \xrightarrow{cs} 1.$$

Puesto que:

$$\sqrt{n} \left(\hat{p}_{ij}^{(n)} - p_{ij} \right)_{1 \leq j \leq m} \xrightarrow{d} \mathcal{N} \left(\mathbf{0}, \frac{1}{p_i} C_i^{-1} \right),$$

cuando $n \rightarrow +\infty$, se cumple, aplicando el resultado del problema 64 de la página 230 de LAHA et al., que:

$$\sqrt{n_i} \left(\hat{p}_{ij}^{(n)} - p_{ij} \right)_{1 \leq j \leq m} \xrightarrow{d} \mathcal{N} \left(\mathbf{0}, C_i^{-1} \right),$$

resultando, para los distintos valores de i , las distribuciones límite independientes. Q.E.D.

Esta proposición prueba que en cada uno de los muestreos multinomiales que hacemos no hay que corregir el error de la estimación por el factor \sqrt{n} , que es el número total de observaciones, sino que basta con corregirlo por $\sqrt{n_i}$, que es el número de veces que se ha hecho el muestreo de la multinomial (p_i). Nótese bien que el factor de corrección $\sqrt{n_i}$ es una variable aleatoria.

3.2.2 Estimación del valor de una regla

Recuérdese que D es el conjunto de reglas markovianas deterministas, es decir:

$$D = A_{s_1} \times \cdots \times A_{s_{m+1}}.$$

El conjunto $d^{-1}(1)$ es el conjunto de estados en los que, siguiendo la regla d , se decide continuar y los estados de $d^{-1}(0)$ son aquéllos en los que se decide parar. Para evitar situaciones triviales supondremos que:

$$d^{-1}(1) \neq \emptyset.$$

Sea l el número de elementos de $d^{-1}(1)$. Por comodidad en la notación y sin pérdida de generalidad admitiremos que estos l estados son los l primeros estados de S , es decir:

$$d^{-1}(1) = \{s_1, \dots, s_l\},$$

siendo $1 \leq l \leq m$,

Se define \bar{P}_d como la matriz de transición de la cadena de Markov asociada a esta regla de decisión. La matriz \bar{P}_d tiene pues elementos: $p(j|i, d(s_i))$, para $1 \leq i \leq m+1$ y $1 \leq j \leq m+1$. Se tiene que:

$$\bar{P}_d = \begin{pmatrix} P_d & T_d \\ 0 \cdots 0 & 1 \end{pmatrix},$$

donde P_d es la matriz $m \times m$:

$$P_d = \begin{pmatrix} Q_d \\ \mathbf{0} \end{pmatrix},$$

siendo Q_d la matriz $l \times m$ formada por las l primeras filas de la matriz P .

Sea $\bar{r}_d \in \mathbb{R}^{m+1}$ el vector de pagos de la regla d , es decir, el vector \bar{r}_d tiene componentes $r(j, d(s_j))$. Sus primeras l componentes, así como la última, son nulas, y las $m-l$ restantes son iguales a las correspondientes componentes del vector g , es decir:

$$\bar{r}_d = \begin{pmatrix} r_d \\ 0 \end{pmatrix},$$

donde $r_d \in \mathbb{R}^m$ verifica:

$$r_d = \begin{pmatrix} \mathbf{0} \\ g_{d^{-1}(0)} \end{pmatrix},$$

designando $g_{d^{-1}(0)}$ las componentes de g correspondientes a los estados de S que están en el conjunto $d^{-1}(0)$.

Se sabe, véase Teorema 7.2.3, p.286, de PUTERMAN, que el valor de la regla d es la mínima solución no negativa en \mathbb{R}^{m+1} , en la incógnita v , de la ecuación:

$$v = \bar{r}_d + \bar{P}_d v. \quad (3.1)$$

Obsérvese que, en problemas con descuento, la solución de esta ecuación es única puesto que la ecuación (3.1) se escribe: $v = \bar{r}_d + \lambda \bar{P}_d v$, donde $0 \leq \lambda < 1$, y la matriz $I - \lambda \bar{P}_d$ tiene inversa, siendo esta solución:

$$(I - \lambda \bar{P}_d)^{-1} \bar{r}_d.$$

En el caso sin descuento, no existe la inversa de $I - \bar{P}_d$ por lo que la solución no es única y es la mínima de entre todas las soluciones no negativas. Sin embargo, si Q es la sub-matriz que corresponde a las transiciones entre los estados de transición de una cadena de Markov, entonces Q es tal que sí existe $(I - Q)^{-1}$. Por tanto, una vez fijado el valor de v en los estados recurrentes, su valor en los estados de transición vendrá unívocamente determinado. Puesto que el valor de d es la mínima solución no negativa de (3.1), parece natural que haciendo v igual a cero en los estados recurrentes obtendremos, ya de forma unívoca, dicha mínima solución. Esto nos sugiere el siguiente resultado:

PROPOSICIÓN 3.4: *El valor $v_d \in \mathbb{R}^{m+1}$ de una regla $d \in D$ es igual a:*

$$\begin{pmatrix} (I - P_d)^{-1} r_d \\ 0 \end{pmatrix}. \quad (3.2)$$

DEMOSTRACIÓN: La cadena de Markov cuya matriz de transición es \bar{P}_d tiene por estados de transición los de S y por único estado recurrente s_{m+1} . Por tanto existe la inversa de $I - P_d$.

Supongamos que el vector $v \in \mathbb{R}^{m+1}$ viene dado por la expresión (3.2). Se tiene que v es no negativo y verifica la ecuación:

$$v = \bar{r}_d + \bar{P}_d v.$$

Sea v' una solución no negativa de (3.1) que cumpla $v' \leq v$. Dado que:

$$v(s_{m+1}) = 0,$$

se tiene necesariamente que $v'(s_{m+1}) = 0$. En la ecuación (3.1), puesto que el valor de v' ya es conocido en los estados recurrentes de \bar{P}_d , es decir, en s_{m+1} ,

el valor de v' está unívocamente determinado en los estados de transición y es igual a v . Se cumple pues que $v = v'$ y el vector v es por tanto la mínima solución de (3.1). Q.E.D.

Obsérvese de qué forma se ha mostrado que v es la mínima solución no negativa de (3.1). No se ha probado, en realidad, que v sea la solución mínima sino que se ha demostrado que cualquier solución no negativa v' menor o igual que v — coordenada a coordenada— es igual necesariamente a v . Esto, unido al hecho de que ya sabemos que la ecuación (3.1) tiene una solución no negativa mínima, nos proporciona el resultado buscado. Nótese que, puesto que la matriz P_d tiene sus últimas $m-l$ filas nulas y $v_d(s_{m+1}) = 0$, se tiene que v_d coincide con \bar{r}_d en los estados s_{l+1}, \dots, s_m .

El lema siguiente permite analizar cómo es la dependencia de v_d respecto de P o de Q_d .

LEMA 3.5: Sea $r \in \mathbb{R}^m$ y sea Q una matriz $m \times m$ de la forma:

$$Q = \begin{pmatrix} Q_1 \\ \mathbf{0} \end{pmatrix},$$

donde $Q_1 \in T_m \times \dots \times T_m$ es una matriz $l \times m$, para algún l que cumpla $1 \leq l \leq m$. La ecuación:

$$\mathbf{v} = r + Q\mathbf{v}, \quad (3.3)$$

define $\mathbf{v} \in \mathbb{R}^m$ como función implícita de Q_1 , siendo esta función implícita $C^\infty(T_m \times \dots \times T_m)$.

DEMOSTRACIÓN: Supongamos que $Q_1 = (Q_{11} \quad Q_{12})$, siendo Q_{11} una matriz $l \times l$ y Q_{12} una matriz $l \times (m-l)$, y sean:

$$\mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} \quad \text{y} \quad r = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix},$$

separando las l primeras coordenadas de \mathbf{v} y r de las $m-l$ últimas. La ecuación (3.3) se escribe entonces:

$$\mathbf{v}_2 = r_2 \quad (3.4)$$

y

$$\mathbf{v}_1 = r_1 + Q_{11}\mathbf{v}_1 + Q_{12}r_2. \quad (3.5)$$

Consideramos la función C definida de $\mathbb{R}^l \times (T_m \times \cdots \times T_m)$ en \mathbb{R}^l cuya expresión es:

$$C(\mathbf{v}_1, Q_1) = r_1 + (Q_{11} - I)\mathbf{v}_1 + Q_{12}r_2.$$

La ecuación $C(\mathbf{v}_1, Q_1) = \mathbf{0}$ define \mathbf{v}_1 como función implícita de Q_1 . Esta función viene dada por:

$$\mathbf{v}_1 = (I - Q_{11})^{-1}r_1 + (I - Q_{11})^{-1}Q_{12}r_2.$$

La función C es C^∞ en su dominio y se cumple que:

$$\frac{\partial C(\mathbf{v}_1, Q_1)}{\partial \mathbf{v}_1} = Q_{11} - I,$$

y la matriz $Q_{11} - I$ tiene determinante no nulo ya que Q_{11} es una matriz positiva con autovalor de Frobenius-Perron menor que uno, puesto que:

$$Q_{11} \in T_l \times \cdots \times T_l.$$

Se deduce pues que, en la ecuación (3.5), el vector \mathbf{v}_1 se puede poner como función C^∞ de Q_1 y como, en virtud de la ecuación (3.4) se tiene que \mathbf{v}_2 es constante, se concluye que \mathbf{v} es función C^∞ de Q_1 . Q.E.D.

Nótese que la función implícita del lema anterior es en realidad:

$$\Psi(Q) = \Psi(Q_1) = (I - Q)^{-1}r.$$

LEMA 3.6: *Suponemos las hipótesis del Lema 3.5 y consideramos la función implícita Ψ definida por:*

$$\Psi(Q) = (I - Q)^{-1}r = \sum_{t=0}^{\infty} Q^t r. \quad (3.6)$$

Existe un entorno V_{Q_1} de Q_1 contenido en $T_m \times \cdots \times T_m$ tal que si dQ_1 es una matriz $l \times m$ tal que:

$$Q_1 + dQ_1 \in V_{Q_1},$$

y definimos la matriz dQ de dimensiones $m \times m$ mediante:

$$dQ = \begin{pmatrix} dQ_1 \\ \mathbf{0} \end{pmatrix},$$

entonces:

$$\Psi(Q + dQ) = \Psi(Q) + D(Q)dQ + \rho(Q, dQ), \quad (3.7)$$

donde, en el segundo miembro de (3.7), dQ es un vector columna de \mathbb{R}^{m^2} , la matriz $D(Q)$ es $m \times m^2$ con sus últimas $m - l$ filas nulas, y $\rho(Q, dQ)$ tiene sus últimas $m - l$ filas nulas y verifica:

$$\|\rho(Q, dQ)\| \leq K \|dQ\|^2,$$

siendo la norma considerada la del supremo, y donde la constante K depende de r y de V_{Q_1} .

DEMOSTRACIÓN: Puesto que $Q_1 \in T_m \times \dots \times T_m$, resulta que la matriz Q verifica que $\|Q\|_B < 1$, y existe una bola abierta $V_{Q_1} \subseteq \mathbb{R}^{m \cdot l}$ que es un entorno de Q_1 y que está contenida en $T_m \times \dots \times T_m$ tal que: si $R_1 \in V_{Q_1}$, y definimos la matriz R , de dimensiones $m \times m$, como:

$$R = \begin{pmatrix} R_1 \\ \mathbf{0} \end{pmatrix},$$

se cumple que:

$$\sup_{R_1 \in V_{Q_1}} \|R\|_B = \lambda < 1. \quad (3.8)$$

Para un entero $t \geq 0$, se denotará por $R^t(i, j)$ el término (i, j) de la t -ésima potencia de R : R^t .

- Fijemos dos pares (i, j) y (i_0, j_0) de $\{1, \dots, l\} \times \{1, \dots, m\}$. Probemos que se cumple, para todo $t \geq 0$ y para todo $R_1 \in V_{Q_1}$, que:

$$\left| \frac{\partial R^t(i, j)}{\partial r_{i_0 j_0}} \right| \leq t \lambda^{t-1}.$$

Lo demostraremos por inducción. Para $t = 0$ el resultado es cierto, donde se adopta el convenio habitual de que $R^0 = I$. Supongamos el resultado verdadero para un cierto t . Se tiene que:

$$R^{t+1}(i, j) = \sum_{k=1}^m R(i, k) R^t(k, j),$$

y por tanto:

$$\frac{\partial R^{t+1}(i, j)}{\partial r_{i_0 j_0}} = \sum_{k=1}^l R(i, k) \frac{\partial R^t(k, j)}{\partial r_{i_0 j_0}} + \delta_{i i_0} R^t(j_0, j).$$

Cada término $\frac{\partial R^t(k, j)}{\partial r_{i_0 j_0}}$ está acotado por $t\lambda^{t-1}$, en virtud de la hipótesis de inducción, y dado que, según (3.8): $\|R\|_B \leq \lambda$, se tiene que:

$$\sum_{k=1}^l R(i, k) \leq \lambda,$$

deduciéndose que:

$$\left| \sum_{k=1}^l R(i, k) \frac{\partial R^t(k, j)}{\partial r_{i_0 j_0}} \right| \leq t\lambda^t.$$

Por otra parte, se cumple que:

$$\|R^t\|_B \leq \|R\|_B^t \leq \lambda^t,$$

al ser $\|R\|_B \leq \lambda$, y por tanto: $R^t(j_0, j) \leq \lambda^t$. En conclusión:

$$\left| \frac{\partial R^{t+1}(i, j)}{\partial r_{i_0 j_0}} \right| \leq t\lambda^t + \lambda^t = (t+1)\lambda^t,$$

con lo cual el resultado buscado es cierto para $t+1$, luego es cierto para todo t , como queríamos demostrar.

Si suponemos que $\|r\| \leq M$, donde se recuerda que la norma considerada es la del supremo, entonces se cumple, para todo $t \geq 0$, para todo $R_1 \in V_{Q_1}$, y para todo i con $1 \leq i \leq l$, que:

$$\left| \frac{\partial (R^t r)(i)}{\partial r_{i_0 j_0}} \right| \leq mMt\lambda^{t-1},$$

siendo $(R^t r)(i)$ la coordenada i -ésima de $R^t r$. Obsérvese que, para cada t , esta cota es uniforme en V_{Q_1} .

- Fijamos otro par $(i'_0, j'_0) \in \{1, \dots, l\} \times \{1, \dots, m\}$. Probemos que para todo $t \geq 0$ y para todo $R_1 \in V_{Q_1}$ se cumple:

$$\left| \frac{\partial^2 R^t(i, j)}{\partial r_{i'_0 j'_0} \partial r_{i_0 j_0}} \right| \leq t(t-1)\lambda^{t-2}.$$

Se hará la demostración por inducción. Para $t = 0$ el resultado es cierto. Lo supondremos cierto para un $t \geq 0$. Se verifica que:

$$\frac{\partial^2 R^{t+1}(i, j)}{\partial r_{i_0 j_0}' \partial r_{i_0 j_0}} = \sum_{k=1}^l R(i, k) \frac{\partial^2 R^t(k, j)}{\partial r_{i_0 j_0}' \partial r_{i_0 j_0}} + \delta_{ii_0} \frac{\partial R^t(j_0, j)}{\partial r_{i_0 j_0}} + \delta_{ii_0} \frac{\partial R^t(j_0, j)}{\partial r_{i_0 j_0}'}. \quad (3.9)$$

Aplicando la hipótesis de inducción y mediante un razonamiento análogo al utilizado previamente, se deduce que:

$$\left| \sum_{k=1}^l R(i, k) \frac{\partial^2 R^t(k, j)}{\partial r_{i_0 j_0}' \partial r_{i_0 j_0}} \right| \leq t(t-1)\lambda^{t-1}.$$

Del resultado anterior de acotación de primeras derivadas parciales y del hecho de que R^t tiene sus $m-l$ últimas filas nulas, se deduce también que cada uno de los dos últimos sumandos de (3.9) está acotado por $t\lambda^{t-1}$, de donde:

$$\left| \frac{\partial^2 R^t(i, j)}{\partial r_{i_0 j_0}' \partial r_{i_0 j_0}} \right| \leq t(t-1)\lambda^{t-1} + 2t\lambda^{t-1} = (t+1)t\lambda^{t-1},$$

siendo pues la hipótesis cierta para $t+1$, y quedando probado el resultado deseado para todo $t \geq 0$.

Se cumple además, para todo $t \geq 0$, para todo $R_1 \in V_{Q_1}$ y para todo i con $1 \leq i \leq l$, que:

$$\left| \frac{\partial^2 (R^t r)(i)}{\partial r_{i_0 j_0}' \partial r_{i_0 j_0}} \right| \leq mMt(t-1)\lambda^{t-2}, \quad (3.10)$$

donde se observa que, fijado t , esta cota es uniforme en V_{Q_1} .

- Nótese que Ψ es función de R , pero que en realidad lo es también de R_1 . Así, la función $R_1 \mapsto \Psi(R_1) = \Psi(R)$ está definida sobre el abierto V_{Q_1} y podemos hablar de derivada de Ψ respecto de R_1 .

Si $1 \leq i_0 \leq l$, $1 \leq j_0 \leq m$ y $1 \leq i \leq l$, se tiene que la función de la variable R_1 dada por $\frac{\partial (R^t r)(i)}{\partial r_{i_0 j_0}}$ es continua en V_{Q_1} , y se verifica también que la serie:

$$\sum_{t=0}^{\infty} \frac{\partial (R^t r)(i)}{\partial r_{i_0 j_0}}$$

es uniformemente convergente en V_{Q_1} , ya que:

$$\sum_{t=0}^{\infty} t\lambda^{t-1} < \infty,$$

y dicha serie converge uniformemente a una función continua. Se deduce de esto que:

$$\frac{\partial\Psi(R_1)(i)}{\partial r_{i_0j_0}} = \sum_{t=0}^{\infty} \frac{\partial(R^t r)(i)}{\partial r_{i_0j_0}},$$

para todo $R_1 \in V_{Q_1}$.

Se cumple asimismo que si $1 \leq i'_0 \leq l$ y $1 \leq j'_0 \leq m$, entonces la serie de funciones continuas en V_{Q_1} :

$$\sum_{t=0}^{\infty} \frac{\partial^2(R^t r)(i)}{\partial r_{i'_0j'_0} \partial r_{i_0j_0}},$$

converge uniformemente en V_{Q_1} a una función continua, ya que:

$$\sum_{t=0}^{\infty} t(t-1)\lambda^{t-2} < \infty.$$

Se verifica entonces que:

$$\frac{\partial^2\Psi(R_1)(i)}{\partial r_{i'_0j'_0} \partial r_{i_0j_0}} = \sum_{t=0}^{\infty} \frac{\partial^2(R^t r)(i)}{\partial r_{i'_0j'_0} \partial r_{i_0j_0}}. \quad (3.11)$$

Se deduce de (3.11) y de (3.10) que:

$$\sup_{R_1 \in V_{Q_1}} \left| \frac{\partial^2\Psi(R_1)(i)}{\partial r_{i'_0j'_0} \partial r_{i_0j_0}} \right| \leq \frac{2mM}{(1-\lambda)^3}.$$

El conjunto V_{Q_1} es una bola abierta de $T_m \times \dots \times T_m$ y es por tanto un convexo. Se le puede aplicar a la función $R_1 \mapsto \Psi(R)(i)$, con $1 \leq i \leq l$, el desarrollo de Taylor de segundo orden en el punto $Q_1 + dQ_1 \in V_{Q_1}$ centrado en el punto $Q_1 \in V_{Q_1}$. Se tiene que:

$$\begin{aligned} \Psi(Q + dQ)(i) &= \Psi(Q_1 + dQ_1)(i) = \Psi(Q_1)(i) + \\ &+ \left(\frac{\partial\Psi(Q_1)(i)}{\partial r_{i_0j_0}} \right) dQ_1 + \frac{1}{2} (dQ_1)' \left(\frac{\partial^2\Psi(Q_1 + \theta_i dQ_1)(i)}{\partial r_{i'_0j'_0} \partial r_{i_0j_0}} \right) dQ_1, \end{aligned} \quad (3.12)$$

para algún $\theta_i \in [0, 1]$. En la ecuación (3.12), dQ_1 es un vector columna con $m \cdot l$ componentes y $(dQ_1)'$ es su traspuesta. El último sumando de (3.12) está acotado por:

$$\frac{1}{2} \frac{2mM}{(1-\lambda)^3} m^2 l^2 \|dQ_1\|^2 \leq \frac{m^5 M}{(1-\lambda)^3} \|dQ_1\|^2 = K \|dQ_1\|^2.$$

Puesto que esta cota es válida para todo i , se deduce el resultado buscado para las l primeras filas de Ψ . Obsérvese que al ser la función Ψ constante en sus $m-l$ últimas coordenadas, se toma como matriz $D(Q)$ la matriz:

$$\left(\frac{\partial \Psi(Q_1)(i)}{\partial r_{i_0 j_0}} \right)_{1 \leq i \leq l}$$

completada con $m-l$ filas nulas, y se toman como l primeras filas de $\rho(Q, dQ)$ el último sumando de (3.12) completado con $m-l$ coordenadas nulas, para llegar a la fórmula deseada. Q.E.D.

Obsérvese que en la demostración de este lema se ha probado en realidad que se puede intercambiar suma con primeras y segundas derivadas parciales en la expresión (3.6).

Fijada una regla de decisión $d \in D$, se define la aplicación φ^d como aquélla que asocia a P_d el valor de la regla d en los estados de S . Nótese que φ^d es en realidad función de Q_d . Se tiene pues, según la Proposición 3.4, que:

$$\varphi^d(P_d) = \varphi^d(Q_d) = (v_d(s_j))_{1 \leq j \leq m} = (I - P_d)^{-1} r_d.$$

Se observa que la función φ^d es la función implícita definida en el Lema 3.5 y que es C^∞ en $T_m \times \cdots \times T_m$ como función de Q_d .

Sabemos que la matriz Q_d está formada con las filas de P correspondientes a los estados de $d^{-1}(1)$. Después de n transiciones, la estimación que se hace de Q_d es la matriz $l \times m$ formada con las filas de $\hat{P}^{(n)}$ de los estados de $d^{-1}(1)$. La estimación de Q_d es $\hat{Q}_d^{(n)}$ y se define $\hat{P}_d^{(n)}$, matriz de dimensiones $m \times m$, como:

$$\hat{P}_d^{(n)} = \begin{pmatrix} \hat{Q}_d^{(n)} \\ \mathbf{0} \end{pmatrix}.$$

La estimación que, tras n transiciones, se hace del valor de d : v_d , es:

$$\hat{v}_d^n = \begin{pmatrix} \varphi^d(\hat{P}_d^{(n)}) \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \varphi^d(\hat{Q}_d^{(n)}) \\ \mathbf{0} \end{pmatrix}.$$

Sabemos que $\varphi^d(Q_d)$ está definida cuando $Q_d \in T_m \times \cdots \times T_m$. Del Teorema 3.1 se deduce que para casi toda trayectoria de las observaciones existe un n a partir del cual $\hat{Q}_d^{(n)} \in T_m \times \cdots \times T_m$, y por tanto, para un tamaño muestral suficientemente grande, la estimación \hat{v}_d^n está bien definida. Nótese que \hat{v}_d^n coincide con v_d en los estados de S que están en $d^{-1}(0)$, siendo igual a g en esos estados. Por tanto, \hat{v}_d^n es un “verdadero” estimador de v_d en los estados de $d^{-1}(1) = \{s_1, \dots, s_l\}$.

El siguiente teorema estudia la estimación del valor de una regla d .

TEOREMA 3.7: *Se cumple para $s \in d^{-1}(0)$ que:*

$$\hat{v}_d^n(s) = v_d(s),$$

y para $s \in d^{-1}(1)$ que:

$$\hat{v}_d^n(s) \xrightarrow{\text{cs}} v_d(s),$$

cuando $n \rightarrow +\infty$. Restringiéndonos a los estados de $d^{-1}(1)$, se cumple:

$$\sqrt{n}(\hat{v}_d^n - v_d)|_{d^{-1}(1)} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma_d),$$

cuando $n \rightarrow +\infty$, siendo Σ_d una matriz $l \times l$.

DEMOSTRACIÓN: En el caso de que $\hat{Q}_d^{(n)} \notin T_m \times \cdots \times T_m$, se hace el convenio de que \hat{v}_d^n sea igual a g en los estados de S que están en $d^{-1}(0)$ e igual a cero en s_{m+1} . Cuando $\hat{Q}_d^{(n)} \in T_m \times \cdots \times T_m$ no es necesario este convenio, y en todo caso se cumple que si $s \in d^{-1}(0)$:

$$\hat{v}_d^n(s) = v_d(s).$$

La función φ^d es continua y puesto que $\hat{Q}_d^{(n)} \xrightarrow{\text{cs}} Q_d$, se cumple que:

$$\hat{v}_d^n(s) \xrightarrow{\text{cs}} v_d(s),$$

cuando $n \rightarrow +\infty$, si $s \in d^{-1}(1)$, como queríamos demostrar.

Para hallar la distribución límite del estimador, se aplica el Lema 3.6 donde, según la notación del lema, hacemos:

$$\begin{aligned} r &= r_d, & \Psi &= \varphi^d, & Q_1 &= Q_d, & Q &= P_d \\ dQ_1 &= \hat{Q}_d^{(n)} - Q_d & \text{y} & & dQ &= \hat{P}_d^{(n)} - P_d. \end{aligned}$$

Para casi todas las trayectorias de las estimaciones existe un tamaño muestral a partir del cual:

$$\hat{Q}_d^{(n)} \in V_{Q_d},$$

y para esas trayectorias, y con n suficientemente grande, se puede escribir:

$$\varphi^d(\hat{P}_d^{(n)}) = \varphi^d(P_d) + D(P_d)(\hat{P}_d^{(n)} - P_d) + \rho(P_d, \hat{P}_d^{(n)}).$$

Fijándonos en las l filas correspondientes a los estados de $d^{-1}(1)$, se cumple, para las restricciones de la matriz $D(P_d)$ y del vector ρ , que:

$$(\hat{v}_d^n - v_d)|_{d^{-1}(1)} = D_d(P_d)(\hat{Q}_d^{(n)} - Q_d) + \rho_d(P_d, \hat{P}_d^{(n)}),$$

donde:

$$\|\rho_d(P_d, \hat{P}_d^{(n)})\| \leq \|\rho(P_d, \hat{P}_d^{(n)})\| \leq K\|\hat{P}_d^{(n)} - P_d\|^2 = K\|\hat{Q}_d^{(n)} - Q_d\|^2, \quad (3.13)$$

siendo la norma considerada la del supremo. Del Teorema 3.2 se deduce que:

$$\sqrt{n}(\hat{Q}_d^{(n)} - Q_d) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma_{Q_d}),$$

cuando $n \rightarrow +\infty$, y por tanto:

$$n^{1/4}\|\hat{Q}_d^{(n)} - Q_d\| = \frac{1}{n^{1/4}}\|\sqrt{n}(\hat{Q}_d^{(n)} - Q_d)\| \xrightarrow{p} 0,$$

luego:

$$\sqrt{n}\|\hat{Q}_d^{(n)} - Q_d\|^2 \xrightarrow{p} 0. \quad (3.14)$$

Sabemos que:

$$\lim_{n \rightarrow \infty} \Pr\{\hat{Q}_d^{(n)} \notin V_{Q_d}\} = 0,$$

y de (3.13) se concluye que:

$$\sqrt{n}(\hat{v}_d^n - v_d)|_{d^{-1}(1)} - \sqrt{n}D_d(P_d)(\hat{Q}_d^{(n)} - Q_d)$$

está acotado, con probabilidad que tiende a uno, por la variable:

$$\sqrt{n}\|\hat{Q}_d^{(n)} - Q_d\|^2,$$

que, según (3.14), tiende en probabilidad a cero. Por tanto:

$$\sqrt{n}(\hat{v}_d^n - v_d)|_{d^{-1}(1)} - \sqrt{n}D_d(P_d)(\hat{Q}_d^{(n)} - Q_d) \xrightarrow{p} \mathbf{0},$$

cuando $n \rightarrow +\infty$. Puesto que $\sqrt{n}D_d(P_d)(\hat{Q}_d^{(n)} - Q_d)$ tiene límite en ley, que es normal de media $\mathbf{0}$ y covarianzas $\Sigma_d = D_d(P_d)\Sigma_{Q_d}(D_d(P_d))'$, resulta que:

$$\sqrt{n}(\hat{v}_d^n - v_d)|_{d^{-1}(1)}$$

tiene límite en ley, que es también $\mathcal{N}(\mathbf{0}, \Sigma_d)$.

Q.E.D.

3.2.3 Estimación de una regla óptima

Se sabe que existe una política determinista estacionaria óptima. Puesto que para cualquier $d \in D$ se tiene que $v_d(s_{m+1}) = 0$, resulta que una regla d será óptima cuando para toda otra regla $d' \in D$ se cumpla:

$$\varphi^d(P_d) \geq \varphi^{d'}(P_{d'}). \quad (3.15)$$

A veces expresaremos la aplicación φ^d como función de P y (3.15) se escribe entonces:

$$\varphi^d(P) \geq \varphi^{d'}(P),$$

siendo las funciones $P \mapsto \varphi^d(P)$ continuas en $T_m \times \cdots \times T_m$.

DEFINICIONES:

Para cada $d \in D$, definimos el conjunto A^d como el conjunto de:

$$P \in T_m \times \cdots \times T_m$$

tales que la regla d es óptima.

Dado $E \subseteq D$, se define A^E como el conjunto de $P \in T_m \times \cdots \times T_m$ tales que las reglas óptimas son las reglas de E . Obsérvese que no es lo mismo A^d que $A^{\{d\}}$.

Se define también el conjunto \tilde{A}^E como el conjunto de $P \in T_m \times \cdots \times T_m$ tales que las reglas óptimas para P están en E .

PROPOSICIÓN 3.8:

1. El conjunto $T_m \times \cdots \times T_m$ se puede poner como unión disjunta de los conjuntos A^E ;
2. para toda $d \in D$, el conjunto A^d es cerrado en $T_m \times \cdots \times T_m$;
3. para todo $E \subseteq D$, el conjunto \tilde{A}^E es abierto en $T_m \times \cdots \times T_m$.

DEMOSTRACIÓN: La primera afirmación se deduce de que para el problema de decisión markoviano siempre existe alguna regla determinista estacionaria que es óptima.

Para ver que A^d es cerrado observamos que:

$$A^d = \bigcap_{d' \in D} \left\{ P \in T_m \times \cdots \times T_m : \varphi^d(P) \geq \varphi^{d'}(P) \right\},$$

De la continuidad de las aplicaciones $\varphi^{d'}$ se deduce que el conjunto A^d es cerrado.

Puesto que:

$$\tilde{A}^E = \bigcap_{d \notin E} (A^d)^c,$$

resulta que \tilde{A}^E es abierto.

Q.E.D.

Sabemos que, con probabilidad uno, a partir de un cierto tamaño muestral: $\hat{P}^{(n)} \in T_m \times \cdots \times T_m$. Se define entonces E_n como el conjunto de reglas que son óptimas cuando las probabilidades de transición vienen dadas por $\hat{P}^{(n)}$, es decir:

$$\hat{P}^{(n)} \in A^{E_n}.$$

Tras observar n transiciones, la estimación de una regla óptima es una regla $d_n \in E_n$. El siguiente resultado prueba que, casi seguramente, este procedimiento de estimación acaba obteniendo una regla óptima para P .

TEOREMA 3.9: *Sea $d_n \in D$ una regla óptima cuando las probabilidades de transición vienen dadas por $\hat{P}^{(n)}$. Se cumple que la probabilidad de que, a partir de un cierto tamaño muestral, la regla d_n sea óptima es igual a uno.*

DEMOSTRACIÓN: Sea E tal que $P \in A^E$. Se tiene entonces que $P \in \tilde{A}^E$ y puesto que, según la Proposición 3.8, \tilde{A}^E es abierto en $T_m \times \cdots \times T_m$, existe, para casi toda trayectoria, un tamaño muestral a partir del cual $\hat{P}^{(n)} \in \tilde{A}^E$, luego $E_n \subseteq E$, y por tanto cualquier regla óptima para $\hat{P}^{(n)}$ está en E y es pues óptima para P . Q.E.D.

3.2.4 Estimación del valor del problema

Sea v^* el valor del problema de decisión. Sabemos que v^* es la mínima solución no negativa en \mathbb{R}^{m+1} de:

$$v \geq \max_{d \in D} \{ \bar{r}_d + \bar{P}_d v \},$$

y también que v^* es la mínima solución no negativa en \mathbb{R}^{m+1} de:

$$v = \max_{d \in D} \{ \bar{r}_d + \bar{P}_d v \}. \quad (3.16)$$

La ecuación (3.16) se escribe:

$$\begin{cases} v(s_i) = \text{máx} \left\{ g(s_i), \sum_{j=1}^{m+1} p(j|i)v(s_j) \right\}, & \text{si } s_i \in S, \\ v(s_{m+1}) = r(s_{m+1}, 0) = 0. \end{cases} \quad (3.17)$$

De (3.17) se deduce que $v^*(s_{m+1}) = 0$ y por tanto, para determinar v^* , deberemos hallar el mínimo vector $v \in \mathbb{R}^m$ que verifique:

$$v(s_i) = \text{máx} \left\{ g(s_i), \sum_{j=1}^m p(j|i)v(s_j) \right\}, \text{ si } s_i \in S, \quad (3.18)$$

y tal que $v \geq \mathbf{0}$. El siguiente resultado prueba que (3.18) tiene una única solución mayor o igual que $\mathbf{0}$ y que no será pues necesario determinar el mínimo de todas las soluciones mayores o iguales que $\mathbf{0}$ de (3.18). Obsérvese que al ser $g \geq \mathbf{0}$, es superfluo decir que la solución sea mayor o igual que $\mathbf{0}$ y (3.18) tiene por tanto una única solución en \mathbb{R}^m .

TEOREMA 3.10: *Si $(v(s))_{s \in S}$ verifica (3.18) entonces:*

$$v^*(s) = v(s), \text{ para } s \in S.$$

DEMOSTRACIÓN: Una política markoviana determinista: π , es una sucesión de reglas de decisión de D , es decir:

$$\pi = (d_1, d_2, \dots).$$

La política π aplica en la primera época de decisión la regla d_1 , en la segunda época la regla d_2 , y así sucesivamente. Se denota por P_π^N la matriz de transición de la política π tras N etapas, que es el producto de las matrices de transición de las reglas d_1, \dots, d_n .

Se cumple, véase el Teorema 7.2.4, p.287, de PUTERMAN, el siguiente resultado: *si $v' \in \mathbb{R}^{m+1}$ verifica (3.16) y, además, para cualquier política markoviana determinista π se cumple:*

$$\limsup_{N \rightarrow \infty} P_\pi^N v' = \mathbf{0}, \quad (3.19)$$

entonces $v' = v^*$.

Aplicamos esta propiedad para probar el teorema. Dada $d \in D$, la matriz P_d tiene algunas de sus filas iguales a P y sus otras filas son nulas, por lo que:

$$\beta = \max_{d \in D} \|P_d\|_B < 1.$$

Por tanto, para toda $d \in D$ y para todo $j \in \{1, \dots, m\}$, se cumple que:

$$p(m+1|j, d(s_j)) \geq 1 - \beta,$$

y para una política markoviana determinista: π , se verifica que:

$$\sum_{k=1}^m P_{\pi}^N(k|j) \leq \beta^N.$$

Se deduce que existe el límite de P_{π}^N cuando $N \rightarrow \infty$ y que:

$$\lim_{N \rightarrow \infty} P_{\pi}^N = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Sea $(v(s))_{s \in S} \in \mathbb{R}^m$ una solución de (3.18) y sea $v' \in \mathbb{R}^{m+1}$ definido por:

$$v' = \begin{pmatrix} v \\ 0 \end{pmatrix}.$$

El vector v' es entonces solución de (3.16) y, dado que su última coordenada es nula, verifica (3.19), cumpliéndose $v' = v^*$. Q.E.D.

En los problemas de decisión markovianos con descuento, la ecuación de optimalidad (3.16), que se escribía:

$$v = \max_{d \in D} \{ \bar{r}_d + \lambda \bar{P}_d v \},$$

caracterizaba unívocamente el valor del problema. Un inconveniente de los problemas sin descuento con el criterio de la recompensa total es que la ecuación de optimalidad no caracteriza el valor del problema. En los modelos positivos, como el que está siendo considerado, el valor es la mínima solución de la ecuación de optimalidad, es decir, que el hecho de que un vector verifique (3.16) aún le deja cierta libertad. La condición (3.19) significa que las soluciones de (3.16) pierden esa "libertad" cuando valen cero en,

dicho informalmente, los estados recurrentes de las políticas markovianas deterministas. La demostración del Teorema 3.10 consiste esencialmente en probar que el único estado recurrente para cualquier política markoviana determinista es s_{m+1} , y en virtud de (3.19), una solución de (3.16) que valga cero en s_{m+1} es en realidad igual a v^* . Esto es lo que demuestra el Teorema 3.10. Así, gracias a este teorema, la ecuación de optimalidad, una vez fijada la última coordenada igual a cero, determina de forma única el valor del problema, y se estudiarán las propiedades de v^* a partir de la ecuación (3.18).

LEMA 3.11: *Sea f una función de $X \subseteq \mathbb{R}^t$ en \mathbb{R}^m y sean C_1, \dots, C_N conjuntos cerrados de X tales que la restricción de f a cada C_i sea continua. Si se tiene que:*

$$X = C_1 \cup C_2 \cup \dots \cup C_N,$$

entonces f es continua en X .

DEMOSTRACIÓN: Sea (x_n) una sucesión en X con límite $x \in X$. Para aquellos C_i que contengan infinitos términos de (x_n) , que forman por tanto una sub-sucesión (x_{i_n}) , se cumple que:

$$f(x_{i_n}) \longrightarrow f(x),$$

y al haber una cantidad finita de conjuntos C_i , resulta que $f(x_n) \rightarrow f(x)$, concluyéndose la continuidad de f . Q.E.D.

LEMA 3.12: *El sistema de ecuaciones, en la variable $v \in \mathbb{R}^m$:*

$$v(s_i) = \text{máx} \left\{ g(s_i), \sum_{j=1}^m p(j|i)v(s_j) \right\}, \text{ si } s_i \in S,$$

define v unívocamente como función implícita φ de $P \in T_m \times \dots \times T_m$. La función φ es continua en $T_m \times \dots \times T_m$.

DEMOSTRACIÓN: La unicidad de la función φ es una consecuencia inmediata del Teorema 3.10. A cada $P \in T_m \times \dots \times T_m$ la función φ le asocia el valor del problema en los estados de S , es decir, para una cierta P se tiene:

$$v^* = \begin{pmatrix} \varphi(P) \\ 0 \end{pmatrix}.$$

La función φ coincide en A^d con φ^d . Del Lema 3.5 se dedujo que la función φ era continua cuando se la restringía a A^d . Puesto que los conjuntos A^d son cerrados, en virtud de la Proposición 3.8, y se cumple:

$$T_m \times \cdots \times T_m = \bigcup_{d \in D} A^d,$$

del Lema 3.11 se deduce el resultado.

Q.E.D.

Después de n transiciones, la estimación que se hace de v^* es \hat{v}^n , donde:

$$\hat{v}^n = \begin{pmatrix} \varphi(\hat{P}^{(n)}) \\ 0 \end{pmatrix} = \begin{pmatrix} \varphi^{d_n}(\hat{P}^{(n)}) \\ 0 \end{pmatrix},$$

siempre que $\hat{P}^{(n)} \in T_m \times \cdots \times T_m$. Se recuerda que d_n es una regla óptima para la matriz $\hat{P}^{(n)}$.

De la continuidad de la función φ se deduce el:

TEOREMA 3.13: *Se cumple el siguiente límite:*

$$\hat{v}^n \xrightarrow{\text{cs}} v^*,$$

cuando $n \rightarrow \infty$.

DEMOSTRACIÓN: Es una consecuencia inmediata del Teorema 3.1 de consistencia casi segura de $\hat{P}^{(n)}$ y del Lema 3.12 de continuidad de φ .

Q.E.D.

Dado $P \in T_m \times \cdots \times T_m$, consideramos el vector v de \mathbb{R}^m definido por:

$$v = \varphi(P),$$

es decir, v es el vector formado con las m primeras coordenadas de v^* :

$$v^* = \begin{pmatrix} v \\ 0 \end{pmatrix}.$$

Sabemos, según (3.18), cf. p.136, que para $s_i \in S$:

$$v(s_i) = \max\{g(s_i), \sum_{j=1}^m p(j|i)v(s_j)\}.$$

Definimos los siguientes conjuntos:

$$S_c = \{s_i \in S : g(s_i) < \sum_{j=1}^m p(j|i)v(s_j)\},$$

y S_c es el conjunto de estados de S en los que la decisión óptima es la acción 1, es decir, continuar;

$$S_p = \{s_i \in S : g(s_i) > \sum_{j=1}^m p(j|i)v(s_j)\},$$

y S_p es el conjunto de estados de S en los que la decisión óptima es la acción 0, es decir, parar;

$$S_i = \{s_i \in S : g(s_i) = \sum_{j=1}^m p(j|i)v(s_j)\},$$

y S_i es el conjunto de estados de S en los que tanto parar como continuar es óptimo, es decir, son los estados que llamaremos de indiferencia.

Estos conjuntos dependen tanto de P como de g y en ocasiones esta dependencia se hará explícita en la notación. Si no hay riesgo de confusión, serán denotados simplemente por S_c , S_p o S_i . Nótese que la regla óptima es única si y sólo si $S_i = \emptyset$.

Enunciamos a continuación un resultado sobre la distribución asintótica de \hat{v}^n .

TEOREMA 3.14: *Si el conjunto S_i es vacío y d^* es la regla óptima se cumple que:*

$$\sqrt{n}(\hat{v}^n - v^*) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma(d^*)),$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: Del Teorema 3.9 se deduce, dado que la regla óptima es única, que existe, para casi toda trayectoria, un tamaño muestral a partir del cual $d_n = d^*$, y por tanto:

$$\hat{v}^n = \hat{v}_{d_n}^n = \hat{v}_{d^*}^n,$$

de donde:

$$\sqrt{n}(\hat{v}^n - \hat{v}_{d^*}^n) \xrightarrow{cs} \mathbf{0}, \quad (3.20)$$

cuando $n \rightarrow +\infty$. En el Teorema 3.7 se probó que:

$$\sqrt{n}(\hat{v}_{d^*}^n - v_{d^*}) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma(d^*)), \quad (3.21)$$

cuando $n \rightarrow +\infty$. De (3.20) y (3.21) se deduce el resultado buscado.

Q.E.D.

Nótese que la matriz $\Sigma(d^*)$, que es $(m+1) \times (m+1)$, es una ampliación con ceros de la matriz $l \times l$ del Teorema 3.7: Σ_{d^*} . Recuérdese que el Teorema 3.7 requería que fuese $(d^*)^{-1}(1) \neq \emptyset$. El caso $(d^*)^{-1}(1) = \emptyset$ se resuelve de forma trivial, pues $v = g$. En realidad, la convergencia a cero de $\hat{v}^n - v^*$ es finita en los estados de S_p , puesto que para casi toda trayectoria existe un tamaño muestral a partir del cual $S_p(P) = S_p(\hat{P}^{(n)})$, al ser $S_i = \emptyset$.

OBSERVACIONES:

- La condición impuesta en el enunciado del Teorema 3.14 es, esencialmente, exigir que si $P \in A^E$ entonces:

$$P \in (A^E)^\circ.$$

En el Teorema 3.14 se impone que E sea uni-puntual: $E = \{d^*\}$, y en ese caso:

$$A^{\{d^*\}} = \tilde{A}^{\{d^*\}},$$

que es abierto, según se probó en la Proposición 3.8. Esta condición permite que, con probabilidad uno, para n suficientemente grande, se tenga $E_n = E$. En el caso de no haber regla óptima única, puesto que:

$$P \in \tilde{A}^E,$$

que es abierto, resulta que, con probabilidad uno, para n suficientemente grande, se tiene que $E_n \subseteq E$, pero puede ser que, fijada una regla d^* óptima, existan valores de n arbitrariamente grandes para los cuales $d^* \in E - E_n$, no pudiéndose establecer un resultado del tipo del Teorema 3.14.

- El Teorema 3.14 no es un caso particular del Teorema 3.7 aplicado a la regla d^* , pues en el Teorema 3.14 la regla óptima d^* no es conocida sino que se estima, etapa a etapa, a partir de las observaciones muestrales, y el Teorema 3.7 fija de antemano la regla de la que se quiere estimar el valor.

3.3 Métodos de perturbación

El Teorema 3.14 no deja totalmente resuelto el problema de la distribución asintótica del estimador de v^* . Da una respuesta en el caso de ser $S_i = \emptyset$, es decir, cuando la regla óptima es única, pero no abarca el caso en el que para algún $s_i \in S$ se verifique:

$$g(s_i) = \sum_{j=1}^m p(j|i)v(s_j). \quad (3.22)$$

Resulta intuitivo que una pequeña perturbación de los parámetros del problema, estos parámetros son g y P , puede lograr que se mantengan las desigualdades en las definiciones de los conjuntos S_c y S_p , y que se rompa la igualdad en (3.22), pues en efecto, la igualdad (3.22) parece sensible a una pequeña alteración de los parámetros. El problema perturbado cumpliría entonces que el conjunto de sus estados de indiferencia es vacío y tendría una única regla óptima que sería estimada de forma consistente. Si la perturbación es lo bastante pequeña, se podría esperar que la regla óptima para el problema perturbado lo fuera también para el problema original, para así trasladar la estimación del problema perturbado al original, pudiéndose deducir un teorema central del límite siguiendo los pasos de la demostración del Teorema 3.14. Esta es la motivación del método de la perturbación que desarrollamos a continuación.

Primeramente introducimos nuevas notaciones.

Se define el conjunto S_{ci} , estados de indiferencia o de continuar, como:

$$S_{ci}(g, P) = S_c(g, P) \cup S_i(g, P),$$

donde esta vez sí se ha hecho explícita la dependencia de estos conjuntos respecto de los parámetros g y P . Podemos suponer, sin pérdida de generalidad y por comodidad en la notación, que los estados de $S_{ci}(g, P)$ son los k primeros, es decir:

$$S_{ci}(g, P) = \{s_1, \dots, s_k\}. \quad (3.23)$$

Hacemos de la matriz P la siguiente partición:

$$P = \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix},$$

siendo la matriz Q_1 de tamaño $k \times k$, que corresponde a las transiciones entre estados de S_{ci} . Si $v = \varphi(P) = \varphi(g, P)$ es el vector formado con las m primeras coordenadas de $v^* = v^*(g, P)$, se consideran $v_{ci} \in \mathbb{R}^k$ y $v_p \in \mathbb{R}^{m-k}$ tales que:

$$v = \begin{pmatrix} v_{ci} \\ v_p \end{pmatrix},$$

y de forma análoga, si g es el vector de \mathbb{R}^m de pagos en los estados de S , se definen g_{ci} y g_p tales que:

$$g = \begin{pmatrix} g_{ci} \\ g_p \end{pmatrix}.$$

De la ecuación (3.18) de la página 136 se deduce que en los estados de S_p se verifica:

$$v(s) = g(s),$$

que escrito en forma matricial es:

$$v_p = g_p. \quad (3.24)$$

En los estados s_i de S_{ci} se cumple que:

$$v(s_i) = \sum_{j=1}^m p(j|i)v(s_j),$$

que se escribe en forma matricial, teniendo en cuenta (3.24):

$$v_{ci} = Q_1 v_{ci} + Q_2 g_p,$$

de donde:

$$v_{ci} = (I - Q_1)^{-1} Q_2 g_p.$$

3.3.1 Perturbación de g

En este apartado mantendremos fija la matriz P y alteraremos el vector g , y por tanto escribiremos, por ejemplo, $S_{ci}(g)$ en lugar de $S_{ci}(g, P)$. Se define el vector columna $\delta(g) \in \mathbb{R}^m$ como:

$$\begin{aligned} \delta(g)(s) &= 0, & \text{si } s \in S_{ci}(g), \\ \delta(g)(s) &= 1, & \text{si } s \in S_p(g), \end{aligned}$$

es decir, según el convenio hecho en (3.23):

$$\delta(g) = (0 \cdots 0 \overset{k}{1} \cdots \overset{m-k}{1})'.$$

Se considerará la siguiente perturbación del vector g :

$$g + \varepsilon\delta(g).$$

PROPOSICIÓN 3.15: *Con la notación anterior, existe $\eta > 0$ tal que siempre que ε cumpla: $0 < \varepsilon < \eta$, se tiene que:*

$$S_i(g + \varepsilon\delta(g)) = \emptyset.$$

DEMOSTRACIÓN: Sea $\varepsilon > 0$. Nótese en primer lugar que el problema con matriz P y con pagos $g_\varepsilon = g + \varepsilon\delta(g)$ en los estados de S y pago cero en s_{m+1} es del tipo de los que han sido estudiados anteriormente. En efecto, se trata de una perturbación positiva, por lo que $g + \varepsilon\delta(g) \geq \mathbf{0}$. No se podría garantizar esto en el caso de hacer una perturbación negativa del tipo $g - \varepsilon\delta(g)$.

Se define el vector $v_\varepsilon \in \mathbb{R}^m$ de la siguiente manera: en los estados de $S_{ci}(g)$ viene definido por:

$$\begin{aligned} v_\varepsilon|_{S_{ci}(g)} &= (I - Q_1)^{-1}Q_2(g_p + \varepsilon\mathbf{1}) \\ &= v_{ci} + \varepsilon(I - Q_1)^{-1}Q_2\mathbf{1}; \end{aligned} \quad (3.25)$$

y en los estados de $S_p(g)$ por:

$$\begin{aligned} v_\varepsilon(s) &= g(s) + \varepsilon \\ &= v(s) + \varepsilon. \end{aligned} \quad (3.26)$$

Hay que observar que la matriz $(I - Q_1)^{-1}$ es igual a $\sum Q_1^t$, ya que el autovalor dominante de la matriz positiva Q_1 es menor que uno. Por tanto, las matrices $(I - Q_1)^{-1}$ y Q_2 tienen todos sus términos estrictamente positivos, y se deduce de (3.25) y de (3.26) que:

$$v_\varepsilon > v. \quad (3.27)$$

Comprobemos que:

$$v_\varepsilon = \varphi(g + \varepsilon\delta(g), P) = \varphi(g_\varepsilon, P),$$

es decir, que v_ε es el valor del problema— en los estados de S — con matriz de transición P y con pagos $g_\varepsilon = g + \varepsilon\delta(g)$. Según el Teorema 3.10, basta

con comprobar que para todo $s_i \in S$ se tiene que:

$$v_\varepsilon(s_i) = \text{máx}\{g(s_i) + \varepsilon\delta(g)(s_i), \sum_{j=1}^m p(j|i)v_\varepsilon(s_j)\}. \quad (3.28)$$

Lo probaremos distinguiendo dos casos posibles.

- Supongamos que $s_i \in S_p(g)$.

Se cumple entonces que:

$$g(s_i) > \sum_{j=1}^m p(j|i)v(s_j).$$

Como la aplicación $\alpha \mapsto v_\alpha$ es continua para $\alpha \geq 0$, existe $\eta > 0$ tal que si $\varepsilon < \eta$, entonces para todo $s_i \in S_p(g)$ se cumple que:

$$g(s_i) > \sum_{j=1}^m p(j|i)v_\varepsilon(s_j).$$

Por tanto:

$$g_\varepsilon(s_i) = g(s_i) + \varepsilon > \sum_{j=1}^m p(j|i)v_\varepsilon(s_j). \quad (3.29)$$

En conclusión:

$$\text{máx}\{g_\varepsilon(s_i), \sum_{j=1}^m p(j|i)v_\varepsilon(s_j)\} = g_\varepsilon(s_i) = g(s_i) + \varepsilon,$$

que es igual a $v_\varepsilon(s_i)$ según (3.26), y queda pues comprobada la ecuación (3.28) cuando $s_i \in S_p(g)$.

- Supongamos que $s_i \in S_{ci}(g)$.

Se cumple entonces que:

$$g(s_i) \leq \sum_{j=1}^m p(j|i)v(s_j).$$

Dado que $g_\varepsilon(s_i) = g(s_i)$ y teniendo en cuenta (3.27), resulta que:

$$g_\varepsilon(s_i) < \sum_{j=1}^m p(j|i)v_\varepsilon(s_j), \quad (3.30)$$

y por tanto:

$$\text{máx}\{g_\varepsilon(s_i), \sum_{j=1}^m p(j|i)v_\varepsilon(s_j)\} = \sum_{j=1}^m p(j|i)v_\varepsilon(s_j). \quad (3.31)$$

Por (3.26) se tiene que: $g_p + \varepsilon \mathbf{1} = v_\varepsilon|_{S_p(g)}$, y de (3.25) se deduce que:

$$v_\varepsilon|_{S_{ci}(g)} = (I - Q_1)^{-1}Q_2(v_\varepsilon|_{S_p(g)}),$$

y por tanto:

$$v_\varepsilon|_{S_{ci}(g)} = Q_1 v_\varepsilon|_{S_{ci}(g)} + Q_2 v_\varepsilon|_{S_p(g)},$$

es decir, que si $s_i \in S_{ci}(g)$ entonces:

$$v_\varepsilon(s_i) = \sum_{j=1}^m p(j|i)v_\varepsilon(s_j),$$

concluyéndose, a partir de (3.31), que se cumple la ecuación (3.28) cuando $s_i \in S_{ci}(g)$, como queríamos demostrar.

Ha quedado pues probado que:

$$v_\varepsilon = \varphi(g_\varepsilon, P).$$

De (3.29) y (3.30) se deduce que:

$$S_p(g_\varepsilon) = S_p(g) \quad \text{y} \quad S_c(g_\varepsilon) = S_{ci}(g),$$

de donde, siempre que $0 < \varepsilon < \eta$, se tiene que $S_i(g_\varepsilon) = \emptyset$. Q.E.D.

Nótese que este resultado es cierto independientemente de que S_i sea vacío o no en el problema original.

Tomando ε como en la Proposición 3.15, el problema con parámetros g_ε y P tiene una única regla óptima: d^{**} , que cumple:

$$\begin{aligned} d^{**}(s) &= 1, & \text{si } s \in S_{ci}(g), \\ d^{**}(s) &= 0, & \text{si } s \in S_p(g), \end{aligned} \quad (3.32)$$

y obsérvese que esta regla también es óptima para el problema (g, P) .

Sea $d_{n,\varepsilon}$ una regla óptima del problema $(g_\varepsilon, \hat{P}^{(n)})$, cumpliéndose, según el Teorema 3.9, con probabilidad uno y a partir de un cierto n , que $d_{n,\varepsilon} = d^{**}$.

Se propone como estimación de $v = \varphi(g, P)$, valor del problema (g, P) en los estados de S , el vector $\tilde{v}_\varepsilon^n \in \mathbb{R}^m$ definido por:

$$\tilde{v}_\varepsilon^n = \varphi^{d_{n,\varepsilon}}(g, \hat{P}^{(n)}),$$

que es, en el problema (g, P) , el valor estimado de la regla $d_{n,\varepsilon}$ tras n transiciones. Hay que hacer notar que la estimación de la matriz P no depende de la perturbación del modelo. El siguiente resultado da las propiedades de este nuevo estimador.

TEOREMA 3.16: *El estimador \tilde{v}_ε^n cumple:*

$$\tilde{v}_\varepsilon^n \xrightarrow{\text{CS}} v,$$

y:

$$\sqrt{n}(\tilde{v}_\varepsilon^n - v) \xrightarrow{\text{d}} \mathcal{N}(\mathbf{0}, \Sigma(d^{**})),$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: Ya se ha dicho que la regla d^{**} es óptima para el problema (g, P) , luego:

$$v = \varphi(g, P) = \varphi^{d^{**}}(g, P). \quad (3.33)$$

Puesto que con probabilidad uno y a partir de un cierto tamaño muestral n , se tiene que $d_{n,\varepsilon} = d^{**}$, se deduce de (3.33), del Teorema 3.1, y de la continuidad en la variable P de la función $\varphi^{d^{**}}$, que:

$$\tilde{v}_\varepsilon^n \xrightarrow{\text{CS}} v,$$

cuando $n \rightarrow +\infty$.

Se tiene también, véase Teorema 3.7, que:

$$\sqrt{n}(\varphi^{d^{**}}(g, \hat{P}^{(n)}) - v) \xrightarrow{\text{d}} \mathcal{N}(\mathbf{0}, \Sigma(d^{**})), \quad (3.34)$$

cuando $n \rightarrow +\infty$. Como:

$$\sqrt{n}(\tilde{v}_\varepsilon^n - \varphi^{d^{**}}(g, \hat{P}^{(n)})) \xrightarrow{\text{CS}} \mathbf{0},$$

se concluye de (3.34) que:

$$\sqrt{n}(\tilde{v}_\varepsilon^n - v) \xrightarrow{\text{d}} \mathcal{N}(\mathbf{0}, \Sigma(d^{**})),$$

cuando $n \rightarrow +\infty$.

Q.E.D.

Hay que hacer notar que este teorema no es una aplicación del Teorema 3.7 tomando como regla d la regla d^{**} . En efecto, la regla d^{**} no es conocida y se estima a partir de la muestra. Aunque la regla d^{**} sea óptima para el problema original (g, P) , los procedimientos de estimación sin perturbación pueden verificar que para todo n sea $d^{**} \notin E_n$, es decir que el modelo sin perturbar nunca consideraría la regla d^{**} como estimación de la regla óptima. La perturbación permite obtener un problema (g_ε, P) que tiene una única regla óptima, regla que es también óptima para el problema original, y que, al ser la única regla óptima de (g_ε, P) , es por tanto correctamente estimada con probabilidad uno a partir de un cierto tamaño muestral.

3.3.2 Perturbación de P

En este apartado mantendremos fijo el vector de pagos g y alteraremos la matriz P , y así escribiremos, por ejemplo, $S_{ci}(P)$ en lugar de $S_{ci}(g, P)$.

Se considera una matriz $\delta(P)$ de tamaño $m \times m$:

$$\delta(P) = \begin{pmatrix} \delta_1 & \delta_2 \\ \delta_3 & \delta_4 \end{pmatrix},$$

siendo la matriz $\delta_1 = \delta_1(P)$ de tamaño $k \times k$, que corresponde a los estados de $S_{ci}(P)$, según el convenio de (3.23). Se considerará la perturbación:

$$P + \varepsilon\delta(P).$$

Hay varias formas de la matriz $\delta(P)$ que hacen válidos los resultados sobre perturbación que serán estudiados. Por ejemplo, se puede tomar como δ_1 una matriz cuyas componentes sean todas iguales a uno y tomar como matrices δ_2 , δ_3 y δ_4 matrices nulas. Otra posibilidad es que $\delta(P)$ tenga todas sus componentes iguales a uno. La proposición siguiente es válida bajo unas hipótesis bastante generales: basta con que sea $\delta_1 > \mathbf{0}$ y las demás matrices $\delta_j \geq \mathbf{0}$, y por tanto abarca los dos casos particulares anteriores.

Consideramos valores de $\varepsilon \geq 0$ lo bastante pequeños para que:

$$P + \varepsilon\delta(P) \in T_m \times \overset{m}{\cdots} \times T_m.$$

Enunciamos un resultado del mismo tipo que la Proposición 3.15.

PROPOSICIÓN 3.17: Si $\delta_1 > \mathbf{0}$ y $\delta(P) \geq \mathbf{0}$ entonces existe $\eta > 0$ tal que siempre que ε cumpla: $0 < \varepsilon < \eta$, se tiene que:

$$S_i(P + \varepsilon\delta(P)) = \emptyset.$$

DEMOSTRACIÓN: Recordemos que:

$$P = \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix}, \quad v = \varphi(P),$$

$$v = \begin{pmatrix} v_{ci} \\ v_p \end{pmatrix} \quad \text{y} \quad g = \begin{pmatrix} g_{ci} \\ g_p \end{pmatrix},$$

separando las coordenadas de los vectores v y g según pertenezcan a estados de $S_{ci}(P)$ o de $S_p(P)$.

El problema original es (g, P) y el problema perturbado $(g, P + \varepsilon\delta(P))$, al que se le pueden aplicar todos los resultados anteriores, ya que sigue cumpliendo las mismas propiedades que (g, P) , únicamente teniendo que tomar la precaución de que:

$$P_\varepsilon = P + \varepsilon\delta(P) \in T_m \times \cdots \times T_m.$$

Se define el vector $v_\varepsilon \in \mathbb{R}^m$ de tal forma que si $s \in S_p(P)$ entonces:

$$v_\varepsilon(s) = v(s) = g(s),$$

es decir:

$$v_\varepsilon|_{S_p(P)} = g_p, \quad (3.35)$$

y en los estados de $S_{ci}(P)$ se define v_ε como:

$$\begin{aligned} v_\varepsilon|_{S_{ci}(P)} &= (I - P_{1,\varepsilon})^{-1} P_{2,\varepsilon} v_\varepsilon|_{S_p(P)} \\ &= (I - P_{1,\varepsilon})^{-1} P_{2,\varepsilon} g_p, \end{aligned} \quad (3.36)$$

siendo:

$$P_\varepsilon = \begin{pmatrix} P_{1,\varepsilon} & P_{2,\varepsilon} \\ P_{3,\varepsilon} & P_{4,\varepsilon} \end{pmatrix},$$

es decir que:

$$v_\varepsilon|_{S_{ci}(P)} = (I - (Q_1 + \varepsilon\delta_1))^{-1} (Q_2 + \varepsilon\delta_2) g_p.$$

Se cumple que la matriz $P_{1,\varepsilon} = Q_1 + \varepsilon\delta_1$ es positiva y que su autovalor dominante es menor que uno, ya que $P_\varepsilon \in T_m \times \dots \times T_m$, por lo que existe la inversa de $I - (Q_1 + \varepsilon\delta_1)$, que vale:

$$(I - (Q_1 + \varepsilon\delta_1))^{-1} = I + (Q_1 + \varepsilon\delta_1) + (Q_1 + \varepsilon\delta_1)^2 + \dots \quad (3.37)$$

De la definición de $S_p(P)$ se deduce que si $s \in S_p(P)$ entonces $g(s) > 0$, es decir, que: $g_p > \mathbf{0}$. De (3.37) se concluye, puesto que $\delta_1 > \mathbf{0}$, que:

$$(I - (Q_1 + \varepsilon\delta_1))^{-1} > (I - Q_1)^{-1} > \mathbf{0},$$

y se tiene también, dado que $\delta_2 \geq \mathbf{0}$, que:

$$Q_2 + \varepsilon\delta_2 \geq Q_2 > \mathbf{0},$$

luego:

$$(I - (Q_1 + \varepsilon\delta_1))^{-1}(Q_2 + \varepsilon\delta_2) > (I - Q_1)^{-1}Q_2,$$

y puesto que, según se acaba de ver: $g_p > \mathbf{0}$, se tiene que:

$$v_\varepsilon|_{S_{ci}(P)} = (I - (Q_1 + \varepsilon\delta_1))^{-1}(Q_2 + \varepsilon\delta_2)g_p > (I - Q_1)^{-1}Q_2g_p = v_{ci},$$

y por tanto:

$$v_\varepsilon|_{S_{ci}(P)} > v_{ci}. \quad (3.38)$$

Comprobemos que:

$$v_\varepsilon = \varphi(g, P_\varepsilon),$$

es decir, que v_ε es el valor del problema— en los estados de S — con matriz de transición P_ε y con pagos g . Según el Teorema 3.10, basta con comprobar que para todo $s_i \in S$ se tiene que:

$$v_\varepsilon(s_i) = \text{máx}\{g(s_i), \sum_{j=1}^m p_\varepsilon(j|i)v_\varepsilon(s_j)\}. \quad (3.39)$$

Lo probaremos distinguiendo dos casos posibles.

- Supongamos que $s_i \in S_p(P)$.

Sabemos que por ser $s_i \in S_p(P)$, se cumple que:

$$g(s_i) > \sum_{j=1}^m p(j|i)v(s_j),$$

y existe $\eta > 0$ tal que si $0 < \varepsilon < \eta$, entonces $P_\varepsilon \in T_m \times \cdots \times T_m$ y:

$$g(s_i) > \sum_{j=1}^m (p(j|i) + \varepsilon \delta(j|i)) v_\varepsilon(s_j), \quad (3.40)$$

ya que la función:

$$\alpha \longmapsto (I - (Q_1 + \alpha \delta_1))^{-1} (Q_2 + \alpha \delta_2) g_p,$$

definida en un intervalo de la forma $[0, h]$, es continua, como se deduce del Lema 3.5. Tomando ε de esta forma resulta que, para $s_i \in S_p(P)$:

$$\max\{g(s_i), \sum_{j=1}^m p_\varepsilon(j|i) v_\varepsilon(s_j)\} = g(s_i),$$

y por (3.35) se tiene que $g(s_i) = v_\varepsilon(s_i)$, por lo que:

$$v_\varepsilon(s_i) = \max\{g(s_i), \sum_{j=1}^m p_\varepsilon(j|i) v_\varepsilon(s_j)\},$$

cuando $s_i \in S_p(P)$, como queríamos demostrar.

- Supongamos que $s_i \in S_{ci}(P)$.

Cuando $s_i \in S_{ci}(P)$ se cumple que:

$$g(s_i) \leq \sum_{j=1}^m p(j|i) v(s_j).$$

Como la matriz $\delta(P)$ es tal que $\delta(P) \geq \mathbf{0}$ se tiene que $P_\varepsilon \geq P$ y por tanto:

$$g(s_i) \leq \sum_{j=1}^m p_\varepsilon(j|i) v(s_j).$$

Por otro lado:

$$\begin{aligned} \sum_{j=1}^m p_\varepsilon(j|i) v(s_j) &= \sum_{j:s_j \in S_p(P)} p_\varepsilon(j|i) v(s_j) + \sum_{j:s_j \in S_{ci}(P)} p_\varepsilon(j|i) v(s_j) \\ &= \sum_{j:s_j \in S_p(P)} p_\varepsilon(j|i) v_\varepsilon(s_j) + \sum_{j:s_j \in S_{ci}(P)} p_\varepsilon(j|i) v(s_j) \\ &< \sum_{j:s_j \in S_p(P)} p_\varepsilon(j|i) v_\varepsilon(s_j) + \sum_{j:s_j \in S_{ci}(P)} p_\varepsilon(j|i) v_\varepsilon(s_j), \end{aligned}$$

en virtud de (3.38), resultando que:

$$g(s_i) < \sum_{j=1}^m p_\varepsilon(j|i)v_\varepsilon(s_j). \quad (3.41)$$

Se tiene pues que para $s_i \in S_{ci}(P)$ se cumple que:

$$\text{máx}\{g(s_i), \sum_{j=1}^m p_\varepsilon(j|i)v_\varepsilon(s_j)\} = \sum_{j=1}^m p_\varepsilon(j|i)v_\varepsilon(s_j). \quad (3.42)$$

De (3.36) se deduce que:

$$v_\varepsilon|_{S_{ci}(P)} = P_{1,\varepsilon} v_\varepsilon|_{S_{ci}(P)} + P_{2,\varepsilon} v_\varepsilon|_{S_p(P)},$$

de donde, para $s_i \in S_{ci}(P)$:

$$\sum_{j=1}^m p_\varepsilon(j|i)v_\varepsilon(s_j) = v_\varepsilon(s_i),$$

y de (3.42) se concluye que se cumple (3.39) cuando $s_i \in S_{ci}(P)$, como queríamos demostrar.

Hemos probado pues que:

$$v_\varepsilon = \varphi(g, P_\varepsilon).$$

Se deduce de (3.40) y de (3.41) que:

$$S_p(P_\varepsilon) = S_p(P) \quad \text{y} \quad S_c(P_\varepsilon) = S_{ci}(P),$$

teniendo entonces, para $0 < \varepsilon < \eta$, que $S_i(P_\varepsilon) = \emptyset$.

Q.E.D.

Obsérvese que este resultado no depende de que para el problema original (g, P) se cumpla $S_i(P) \neq \emptyset$.

Tomamos ε tal que se cumpla la Proposición 3.17. El problema (g, P_ε) tiene una única regla óptima d^{**} , que es la misma que la definida en (3.32), cf. p.146. Esta regla “para” cuando en el problema original hay que parar y “continúa” cuando la decisión óptima en el problema original es continuar o bien la decisión es indiferente. La regla d^{**} es por tanto óptima también para el problema (g, P) .

Supongamos que después de observar n transiciones del problema original, con la matriz P , la estimación de máxima verosimilitud de P es $\hat{P}^{(n)}$. La estimación que hacemos de P_ε es:

$$\hat{P}_\varepsilon^{(n)} = \hat{P}^{(n)} + \varepsilon\delta(P),$$

de tal forma que:

$$\hat{P}_\varepsilon^{(n)} \xrightarrow{\text{cs}} P_\varepsilon, \quad (3.43)$$

cuando $n \rightarrow +\infty$. Sea $d_{n,\varepsilon}$ una regla óptima para el problema $(g, \hat{P}_\varepsilon^{(n)})$. Se toma como estimador de v — valor del problema (g, P) en las coordenadas correspondientes a los estados de S — el vector \tilde{v}_ε^n dado por:

$$\tilde{v}_\varepsilon^n = \varphi^{d_{n,\varepsilon}}(g, \hat{P}_\varepsilon^{(n)}),$$

que es el valor de la regla $d_{n,\varepsilon}$ en el problema $(g, \hat{P}_\varepsilon^{(n)})$. El siguiente teorema muestra propiedades de consistencia y distribución límite de este nuevo estimador de v .

TEOREMA 3.18: *El estimador \tilde{v}_ε^n cumple:*

$$\tilde{v}_\varepsilon^n \xrightarrow{\text{cs}} v,$$

y:

$$\sqrt{n}(\tilde{v}_\varepsilon^n - v) \xrightarrow{\text{d}} \mathcal{N}(\mathbf{0}, \Sigma(d^{**})),$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: Puesto que el conjunto $A^{\{d^{**}\}}$ es abierto, véase Proposición 3.8, y $P_\varepsilon \in A^{\{d^{**}\}}$, teniendo en cuenta (3.43) se concluye que con probabilidad uno, a partir de un cierto n , se tiene que $d_{n,\varepsilon} = d^{**}$. Obsérvese que $\hat{P}_\varepsilon^{(n)}$ no es el estimador de máxima verosimilitud de P_ε , pero sí se trata de un estimador consistente casi seguramente.

El resto de la demostración es similar a la del Teorema 3.16. Q.E.D.

La estimación de v se podría haber enfocado de otra manera: considerando el estimador $\tilde{P}_\varepsilon^{(n)}$ de máxima verosimilitud de P_ε y estimando P mediante:

$$\tilde{P}_\varepsilon^{(n)} - \varepsilon\delta(P).$$

Se hubiesen conseguido unos resultados límite— casi seguramente y en distribución— similares. La única diferencia es que esta forma de estimar requiere

que se muestree la matriz perturbada P_ε , mientras que los resultados que se han demostrado se basan únicamente en un muestreo de la matriz original P , no teniendo que cambiar el muestreo que se ha considerado a lo largo de todo el texto.

3.3.3 Comentarios y conclusiones

- Según se hizo notar, el Teorema 3.14 no daba una distribución límite del estimador del valor del problema en todos los casos posibles. El método de la perturbación consigue, en cambio, en todos los casos, una distribución límite de un estimador, que será en general otro que el estudiado en el Teorema 3.14. Nótese que a los problemas perturbados: (g_ε, P) y (g, P_ε) , se les podría aplicar el Teorema 3.14 puesto que verifican que sus conjuntos de indiferencia son vacíos. Sin embargo, el objetivo del método de la perturbación no es hallar un teorema central del límite para los problemas perturbados. El método pretende que la estimación de la regla óptima converja a una regla fija. Como esto no se podía lograr en el caso $S_i \neq \emptyset$, se ha modificado ligeramente el problema (g, P) para obtener un problema con una única regla óptima que además lo fuese también para el problema original. Tomando como estimación de la regla óptima la estimación que se hace en el modelo perturbado, se logra evitar la oscilación de los conjuntos E_n .
- El método de la perturbación nos da una cierta idea de cuántos casos no cubre el método del Teorema 3.14, que correspondería a una perturbación de parámetro $\varepsilon = 0$. En efecto, pensando por ejemplo en una perturbación de P del tipo:

$$p_\varepsilon(j|i) = p(j|i) + \varepsilon,$$

para todos i y j , resulta que: para toda P a la que no se le aplique el Teorema 3.14 existe un cubo abierto de medida positiva con un vértice en P tal que a todos los puntos del cubo sí se les puede aplicar el Teorema 3.14.

- Como se habrá podido observar, los dos métodos de perturbación: tanto el de perturbación del vector g como el de perturbación de la matriz P , llegan a resultados análogos. Pese a la mayor sencillez en los cálculos y en las demostraciones del método de perturbación del vector de pagos g , es más recomendable utilizar el método de perturbación de

la matriz P . En efecto, el vector g se alteraba sumándole un múltiplo del vector $\delta(g)$, que valía uno en los estados de $S_p(g)$ y cero en los demás. Este vector de perturbación depende de la matriz P que es desconocida y por tanto no se conoce en realidad el vector $\delta(g)$. Es más, si tomamos como estimación de $S_p(g, P)$ el conjunto $S_p(g, \hat{P}^{(n)})$ no tiene por qué cumplirse que, con probabilidad uno a partir de un cierto tamaño muestral, se verifique:

$$S_p(g, \hat{P}^{(n)}) = S_p(g, P),$$

y no se podría estimar directamente de forma consistente el vector $\delta(g)$. Sí es cierto, en cambio, que se podría tomar como estimación del conjunto $S_p(g, P)$ el conjunto $S_p(g_\varepsilon(\hat{P}^{(n)}), \hat{P}^{(n)})$, verificándose para casi toda trayectoria y a partir de un n que:

$$S_p(g_\varepsilon(\hat{P}^{(n)}), \hat{P}^{(n)}) = S_p(g, P).$$

Este procedimiento añade la complicación de que la perturbación varía con los valores de n , siendo además aleatoria. Hay una vía más sencilla para aplicar de otra forma el método expuesto.

Ya se observó, antes de la Proposición 3.17, que hay muchas matrices $\delta(P)$ para las cuales el método de perturbación de P es válido, entre las cuales está la matriz $\delta(P)$ de tamaño $m \times m$ que tiene todas sus componentes iguales a uno. Esta matriz no depende de P y, al contrario que con $\delta(g)$, se puede aplicar la perturbación a la matriz P sin necesitar ninguna información previa sobre P ni tener que hacer la estimación de $S_p(g, P)$. Por todo esto, para la aplicación de la perturbación, lo más recomendable parece ser una alteración, para todos i y j , del tipo: $p_\varepsilon(j|i) = p(j|i) + \varepsilon$.

CAPÍTULO 4

Método Directo en Parada Óptima

4.1 Introducción

4.1.1 Objetivos del estudio

En el capítulo anterior se abordó un problema de parada óptima en el que las probabilidades de transición de la cadena de Markov eran desconocidas. El objetivo último del estudio realizado era el de estimar una regla óptima, así como el valor del problema de decisión. El método que se empleó para dar solución a esta cuestión, que se puede designar mediante la expresión “resolver y estimar”, fue el de estimar primeramente la matriz de transición de la cadena de Markov, para después resolver el problema de parada óptima cuya matriz de transición fuese la matriz estimación. Se tomó la solución del problema estimado como estimación de la solución del problema. Con esta técnica se llegó a resultados de convergencia casi segura y de convergencia en distribución para las estimaciones del valor de las reglas de decisión y del valor del problema. Se lograba también estimar consistentemente una regla óptima. Esta técnica de estimación tiene, desde el punto de vista práctico, algunos inconvenientes. Uno de ellos es la gran cantidad de operaciones que los procedimientos de estimación llevan implícitas, como el cálculo de inversas de matrices. Otro inconveniente es que se están estimando muchos más parámetros de los que presumiblemente sería necesario estimar para

alcanzar el objetivo propuesto: dar una aproximación de una regla óptima y del valor del problema.

En este nuevo estudio, se ha tratado el mismo problema de estimación, planteando los mismos objetivos, pero dentro de la línea que marcaría la expresión: “estimar sin resolver”. El procedimiento empleado es directo en el sentido de que los estadísticos se proponen con el único fin de que estimen las cantidades de interés. No se produce pues una situación de sobre-estimación, aunque esta técnica no proporciona un conocimiento tan profundo del problema como la utilizada en el capítulo anterior. Además, las operaciones que requiere llevar a cabo este método se reducen prácticamente al cálculo de promedios y a la comparación de cantidades, llegando a la misma clase de resultados que con el procedimiento ya estudiado: consistencia casi segura de la estimación de una regla óptima, y teoremas de convergencia casi segura y en distribución para los estimadores del valor de las reglas y del valor del problema.

El orden seguido para realizar este estudio es el siguiente.

En primer lugar, se procederá a plantear el problema y a definir la notación que será utilizada.

En una segunda parte, se analiza el problema de la estimación del valor de una regla de decisión. Para estudiar las propiedades límite de la estimación propuesta se tratarán dos métodos distintos. El primer método, basado en la Teoría de Martingalas, nos proporcionará un resultado de convergencia casi segura y un resultado de convergencia en distribución. El segundo método, que parte del carácter markoviano del problema, también resuelve las cuestiones de la convergencia casi segura y en ley de los estimadores.

En la tercera sección, se aborda la estimación de una regla óptima. Esta estimación se hace de una manera sensiblemente diferente a la forma en que se hizo en la anterior resolución del problema de parada óptima, y en este punto se apreciará una diferencia esencial entre los métodos de “resolver y estimar” y de “estimar sin resolver”.

En la cuarta sección se procede a la estimación del valor del problema. Se prueba que el estimador considerado en un principio es consistente con probabilidad uno, pero que no verifica un teorema central del límite a no ser que la regla óptima sea única. Esta pega ya había surgido en el análisis del Capítulo 3. Los métodos de perturbación que allí se desarrollaron son igualmente aplicables y permiten demostrar un teorema central del límite con total generalidad. Se estudia, de todas maneras, un nuevo método de perturbación que resulta más operativo y más adecuado al tipo de estimación que se está llevando a cabo. También para esta nueva técnica de pertur-

bación se prueba un resultado de convergencia en ley, con independencia de la unicidad de la regla óptima del problema original.

Finalmente se hace, mediante una simulación, la comparación de este método de estimación y del de máxima verosimilitud.

4.1.2 Notación y definición del modelo

Cuando se pruebe que una sucesión $(X_n)_{n \geq 1}$ converge en ley a una variable aleatoria normal de media cero, la notación empleada será:

$$X_n \xrightarrow{d} \mathcal{N}(0, \cdot).$$

Al escribir expresiones que involucren esperanzas condicionadas, se omitirá que dichas expresiones se verifican casi seguramente, y esto no supondrá pérdida de rigor.

Se plantea el modelo en los mismos términos que en el capítulo anterior. Sobre un espacio de probabilidad $(\Omega, \mathcal{F}, \Pr)$ se considera una cadena de Markov homogénea con espacio de estados finito:

$$\bar{S} = \{s_1, \dots, s_m, s_{m+1}\}.$$

Los estados de:

$$S = \{s_1, \dots, s_m\},$$

son los estados de transición de la cadena, y el estado s_{m+1} es absorbente. Se toma $m > 1$. La matriz de transición entre los estados de S es:

$$P = (p(s'|s))_{s, s' \in S}.$$

Las probabilidades de transición entre los estados transitorios son desconocidas.

En el Capítulo 3 habíamos impuesto que:

$$\forall (s, s') \in S \times S, p(s'|s) > 0 \quad \text{y} \quad \forall s \in S, \sum_{s' \in S} p(s'|s) < 1,$$

lo que se escribía:

$$P \in T_m \times \overset{m}{\cdots} \times T_m.$$

Esta hipótesis ya no será necesaria.

Recordemos parte de la notación del Capítulo 3. La matriz de transición de la cadena de Markov asociada al problema es:

$$\bar{P} = \begin{pmatrix} P & T \\ 0 \cdots 0 & 1 \end{pmatrix}.$$

En cada estado se pueden tomar dos acciones: parar y recibir un pago según el estado en el que se encuentre el sistema, o bien continuar y efectuar una transición más, con arreglo a la cadena de Markov con matriz de transición \bar{P} . El pago que se recibe en el estado absorbente es igual a cero: $g(s_{m+1}) = 0$, y los pagos en los demás estados vienen dados por el vector m -dimensional g definido de la siguiente manera:

$$g = (g(s))_{s \in S}.$$

Se supone que $g \geq \mathbf{0}$.

Los conjuntos de acciones son, para $s \in S$:

$$A_s = \{0, 1\} \quad \text{y} \quad A_{s_{m+1}} = \{0\},$$

donde la acción 0 significa “parar” y la acción 1 significa “continuar”. Las probabilidades de transición entre los estados de \bar{S} , según la acción tomada, son, cuando $s \in S$ y $s' \in \bar{S}$:

$$p(s'|s, 1) = p(s'|s), \quad p(s_{m+1}|s, 0) = 1 \quad \text{y} \quad p(s_{m+1}|s_{m+1}, 0) = 1,$$

y las funciones de pago son, cuando $s \in S$:

$$r(s, 1) = 0, \quad r(s, 0) = g(s), \quad \text{y} \quad r(s_{m+1}, 0) = 0.$$

4.2 Estimación del valor de una regla de decisión

4.2.1 Introducción y definiciones

Definamos en primer lugar el tipo de muestreo que se va a utilizar. Suponemos que observamos las evoluciones de la cadena de Markov que tiene como matriz de transición la matriz \bar{P} . Como esta cadena tiene un único estado absorbente: s_{m+1} , el sistema acaba llegando a ese estado absorbente. Una vez en s_{m+1} , volvemos a observar las evoluciones de la cadena partiendo de otro estado de \bar{S} . En otras palabras, fijamos de antemano unas cantidades positivas: q_1, \dots, q_{m+1} , tales que:

$$q_1 + \dots + q_{m+1} = 1,$$

de tal forma que observamos las transiciones de la cadena de Markov con matriz asociada:

$$\tilde{P} = \begin{pmatrix} P & T \\ q_1 \cdots q_m & q_{m+1} \end{pmatrix}.$$

Considerar esta matriz equivale a decir que cada vez que la cadena original llega al estado s_{m+1} la reinicializamos en un estado de \bar{S} elegido con arreglo a las probabilidades q_1, \dots, q_{m+1} .

Nótese que la cadena que se observa: la que tiene matriz de transición \tilde{P} , es irreducible y aperiódica independientemente de la condición:

$$P \in T_m \times \cdots \times T_m.$$

En efecto, al ser las probabilidades q_1, \dots, q_m positivas, la cadena resulta ser irreducible, y al tomar $q_{m+1} > 0$, es aperiódica.

Partimos de un estado inicial $s^{(1)}$ y la sucesión aleatoria de estados observada es:

$$s^{(1)}, s^{(2)}, s^{(3)}, \dots \quad (4.1)$$

Fijemos una regla de decisión $d \in D$. El valor de d en los estados tales que $d(s) = 0$, es decir, en los estados en los que la acción tomada es detenerse, es conocido y vale:

$$v_d(s) = g(s),$$

y no necesita pues ser estimado. La estimación es necesaria en los estados de S tales que $d(s) = 1$. Por tanto, supondremos que la regla d fijada verifica:

$$d^{-1}(1) \neq \emptyset.$$

DEFINICIÓN: La sucesión de variables aleatorias: $(K_r^d)_{r \geq 0}$, viene dada de forma recurrente:

- $K_0^d = 0$;
- $K_r^d = \min\{n > K_{r-1}^d : d(s^{(n)}) = 0\}$, para $r \geq 1$.

Hay que observar que, en efecto, las variables aleatorias definidas dependen de la regla $d \in D$ que se haya fijado. La interpretación de la sucesión definida es la siguiente. La variable K_1^d representa la primera observación de la sucesión (4.1) en la que la decisión tomada por la regla d es detenerse. El valor K_2^d es el número de la primera observación posterior a la K_1^d -ésima

en la que la regla d se para. Puesto que la cadena de Markov que induce la regla d en el conjunto \bar{S} tiene como estados absorbentes aquéllos en los que:

$$d(s) = 0,$$

resulta que la sucesión $(K_r^d)_{r \geq 1}$ representa los instantes de absorción de la cadena inducida por d que se observan en la muestra de la cadena con matriz de transición \tilde{P} .

Se hace el convenio de que si $\{n > K_{r-1}^d : d(s^{(n)}) = 0\} = \emptyset$, entonces:

$$K_r^d = K_{r+1}^d = K_{r+2}^d = \dots = +\infty.$$

Hay que hacer notar que la sucesión $(K_r^d)_{r \geq 0}$ es estrictamente creciente y finita con probabilidad uno y, sin pérdida de generalidad, vamos a suponer que sólo observamos trayectorias muestrales $\omega \in \Omega$ tales que:

$$\forall r \geq 0, K_r^d(\omega) < +\infty.$$

Nos referiremos a las observaciones:

$$1, \dots, K_1^d$$

de (4.1), como las observaciones del primer tramo, que son las observaciones hechas hasta el momento de la primera absorción, es decir, hasta la primera vez que se detiene la regla d . Las observaciones:

$$K_1^d + 1, \dots, K_2^d,$$

forman el segundo tramo, y así sucesivamente.

Al aplicar la regla d , el sistema ha evolucionado a través de los estados del primer tramo hasta detenerse en la observación K_1^d -ésima, recibándose un pago igual a:

$$g(s^{(K_1^d)}).$$

Una vez obtenido ese primer pago, la cadena vuelve a transitar de estado en estado durante el segundo tramo, al final del cual se recibe un pago:

$$g(s^{(K_2^d)}),$$

etc..

La forma que se propone para estimar el valor $v_d(s)$ es la siguiente. Los tramos que se observan resultan de aplicar la política estacionaria d^∞ . Por

tanto, si el estado s pertenece, por ejemplo, al i -ésimo tramo, entonces se puede interpretar la porción del i -ésimo tramo que va de la observación del estado s a la absorción por el estado $s^{(K_i^d)}$ como las transiciones que resultan de aplicar la regla d tomando como estado inicial el estado s , siendo pues el pago $g(s^{(K_i^d)})$ una estimación del valor $v_d(s)$. Esta observación sugiere dos formas distintas de estimar $v_d(s)$, aunque razonaremos que una de ellas debe ser descartada.

DEFINICIONES: Fijados $d \in D$, $s \in S$ tal que $d(s) = 1$, y un entero $r \geq 1$, consideramos las siguientes variables aleatorias:

•

$$M_d(s, r) = \sum_{i=K_{r-1}^d+1}^{i=K_r^d} I_{\{s^{(i)}=s\}},$$

que es igual al número de veces que se observa el estado s en el r -ésimo tramo;

•

$$I_d(s, r) = \begin{cases} 0, & \text{si } M_d(s, r) = 0, \\ 1, & \text{si } M_d(s, r) > 0, \end{cases}$$

que indica si el estado s ha aparecido en el r -ésimo tramo;

•

$$N_d(s, r) = \sum_{i=1}^r I_d(s, i),$$

que es el número de tramos, de entre los r primeros, en los que se ha observado s .

Una primera forma de estimar $v_d(s)$ sería tener en cuenta cada vez que en la sucesión de estados observada aparece el estado s y promediar los pagos de los estados finales de cada tramo por cada observación del estado s . Esta estimación sería pues:

$$\frac{\sum_{i=1}^r M_d(s, i)g(s^{(K_i^d)})}{\sum_{i=1}^r M_d(s, i)}, \quad (4.2)$$

cuando se hayan observado r tramos. De esta forma, cada aparición del estado s se ve reflejada añadiendo un sumando más.

Veamos mediante un ejemplo sencillo que este estimador no resulta adecuado. Supongamos que la sucesión de estados observada es:

$$s_1, s_1, s_1, s_1, s_2, s_1, s_3, \quad (4.3)$$

donde:

$$d(s_1) = 1, \quad d(s_2) = 0 \quad \text{y} \quad d(s_3) = 0.$$

La muestra (4.3) tiene pues dos tramos: el primero formado por las cinco primeras observaciones y el segundo por las dos últimas observaciones. Con la notación introducida anteriormente resulta que:

$$M_d(s_1, 1) = 4, \quad s^{(K_1^d)} = s_2, \quad M_d(s_1, 2) = 1 \quad \text{y} \quad s^{(K_2^d)} = s_3,$$

y la estimación definida en (4.2) es, para $s = s_1$:

$$\frac{4g(s_2) + g(s_3)}{5}.$$

En la estimación, el peso relativo del término $g(s_2)$ es igual a cuatro y el peso relativo del término $g(s_3)$ es uno, lo cual quiere decir que la estimación de $v_d(s_1)$ va a resultar más cercana a $g(s_2)$ que a $g(s_3)$. Esto sería equivalente a decir que, partiendo del estado s_1 y aplicando la regla de decisión d , la probabilidad de que la absorción por los estados en los que la decisión es parar se produzca por el estado s_2 es mucho mayor que la probabilidad de que dicha absorción se produzca por el estado s_3 . Pero resulta claro que de ninguna manera la muestra (4.3) permite inferir esa afirmación. Lo más que se puede deducir a la vista de (4.3) es que la probabilidad:

$$p(s_1|s_1)$$

es alta, pero no se pueden sacar conclusiones sobre las probabilidades de absorción por los estados en los que la regla d para. Es por esto que del primer tramo lo único que realmente nos interesa es saber si se ha observado el estado s_1 o no, independientemente de cuántas veces antes de la absorción vuelva a observarse el estado s_1 . De esta forma, el estimador que se propondría sería:

$$\frac{g(s_2) + g(s_3)}{2},$$

tomando un sumando por cada tramo en el que haya aparecido el estado s_1 .

Pasamos a definir, de forma general, el estimador de $v_d(s)$. Suponemos que se han observado los r primeros tramos. Se define:

$$\hat{v}_d^{(r)}(s) = \frac{\sum_{i=1}^r I_d(s, i)g(s^{(K_i^d)})}{N_d(s, r)}, \quad (4.4)$$

siempre que $N_d(s, r) > 0$. Si $N_d(s, r) = 0$ entonces la definición de $\hat{v}_d^{(r)}(s)$ es cualquiera. Hay que observar que, puesto que la cadena que se está muestreando es irreducible, para casi todo $\omega \in \Omega$ existe un tamaño muestral a partir del cual $N_d(s, r) > 0$, y por tanto, con probabilidad uno, tomando r lo bastante grande, la definición de $\hat{v}_d^{(r)}(s)$ es la dada por (4.4).

La estimación (4.4) consiste en hacer la media aritmética de los pagos recibidos en los estados finales de los tramos en los que se ha observado el estado s . De un tramo sólo nos interesa saber si se ha observado el estado s y no cuántas veces se ha observado s en ese tramo.

Hay que hacer notar que $\hat{v}_d^{(r)}(s)$ es el estimador conseguido tras observar los r primeros tramos, y necesita pues una cantidad aleatoria de observaciones:

$$K_r^d,$$

de la muestra (4.1), cf. p.161.

Se puede definir también un estimador para un tamaño muestral fijo igual a n . Dado $n \geq 1$, se define la variable aleatoria r_n como:

$$r_n = \text{máx}\{r \geq 0 : n \geq K_r^d\},$$

y r_n es igual al número de tramos completos que se han observado en la muestra de tamaño n . La definición del estimador de $v_d(s)$ para un tamaño muestral n es:

$$\tilde{v}_d^{(n)}(s) = \hat{v}_d^{(r_n)}(s), \quad (4.5)$$

siempre que $r_n > 0$. Si $r_n = 0$, es decir, si no se ha observado ningún tramo completo, la estimación es cualquiera. Nótese que r_n depende de d , aunque no haremos explícita esta dependencia.

Es importante definir un estimador para cada tamaño muestral fijo por lo siguiente: según se observó, el número de tramos que contiene una muestra depende de la regla d que se esté considerando. Cuando tengamos que comparar las estimaciones de varias reglas y establecer resultados límite para esas comparaciones, necesitaremos que las estimaciones se expresen a través del tamaño muestral n , que es su característica común, mientras que el número de tramos puede ser distinto para cada una de las estimaciones.

4.2.2 Convergencia del estimador. Primer método

Planteamiento

Vamos a utilizar resultados de martingalas para abordar el estudio de la convergencia de la estimación:

$$(\hat{v}_d^{(r)}(s))_{r \geq 1},$$

fijados una regla $d \in D$ y un estado $s \in d^{-1}(1)$.

La muestra es la sucesión de variables aleatorias con valores en \bar{S} :

$$s^{(1)}, s^{(2)}, s^{(3)}, \dots,$$

y se define una filtración de sub- σ -álgebras de \mathcal{F} que va a representar la información que se tiene en cada etapa del muestreo. En concreto:

$$\mathcal{A}_0 = \{\emptyset, \Omega\} \quad \text{y} \quad \mathcal{A}_n = \sigma(s^{(1)}, \dots, s^{(n)}), \quad \text{para } n \geq 1,$$

y \mathcal{A}_n es la información disponible hasta la n -ésima muestra. Se cumple que cada variable aleatoria K_j^d , siendo $j \geq 0$, es un tiempo de parada respecto de la filtración $\{\mathcal{A}_n\}_{n \geq 0}$.

Se define, para cada $j \geq 0$, la sub- σ -álgebra \mathcal{F}_j como la σ -álgebra de sucesos anteriores al tiempo de parada K_j^d , es decir:

$$\mathcal{F}_0 = \{\emptyset, \Omega\} \quad \text{y} \quad \mathcal{F}_j = \sigma(s^{(1)}, \dots, s^{(K_j^d)}), \quad \text{para } j \geq 1,$$

y \mathcal{F}_j es la información disponible en los j primeros tramos.

Convergencia casi segura

Queremos estudiar la convergencia casi segura de:

$$\hat{v}_d^{(r)}(s) = \frac{\sum_{i=1}^r I_d(s, i) g(s^{(K_i^d)})}{N_d(s, r)},$$

y lo haremos analizando en primer lugar el comportamiento asintótico del denominador. La sucesión, en la variable $r \geq 1$, de denominadores:

$$N_d(s, r) = I_d(s, 1) + \dots + I_d(s, r),$$

está adaptada a la filtración:

$$\{\mathcal{F}_r\}_{r \geq 1},$$

ya que $N_d(s, r)$ es el número de tramos, de entre los r primeros, en los que se ha observado el estado s , y es pues \mathcal{F}_r -medible. Al ser $0 \leq N_d(s, r) \leq r$ se tiene que el denominador es una variable aleatoria integrable. Lo que ocurre es que la sucesión de denominadores no forma una martingala, o dicho en otros términos, la sucesión de término:

$$I_d(s, r), \tag{4.6}$$

para $r \geq 1$, no es una diferencia de martingalas. Si queremos aplicar métodos para martingalas, o para diferencias de martingalas, debemos convertir la sucesión de término general (4.6) en una diferencia de martingalas restándole su esperanza condicionada respecto del elemento \mathcal{F}_{r-1} de la filtración.

LEMA 4.1: *La sucesión de variables aleatorias $(W_r)_{r \geq 1}$ definida por:*

$$W_r = I_d(s, r) - \mathbf{E}(I_d(s, r) | \mathcal{F}_{r-1}),$$

es una diferencia de martingalas para la filtración:

$$\{\mathcal{F}_r\}_{r \geq 0}.$$

Se verifica que:

$$\frac{W_1 + \dots + W_r}{r} \xrightarrow{\text{cs}} 0,$$

cuando $r \rightarrow +\infty$.

DEMOSTRACIÓN: La comprobación de la primera parte del enunciado es inmediata.

El resultado de convergencia está basado en el siguiente teorema (cf. IBARROLA et al., p.407):

Dada una sucesión $(X_r)_{r \geq 1}$ de variables aleatorias adaptada a la filtración $\{\mathcal{F}_r\}_{r \geq 0}$ tal que:

$$\mathbf{E}(X_r | \mathcal{F}_{r-1}) = 0, \tag{4.7}$$

para todo $r \geq 1$, y si existe una sucesión $(C_r)_{r \geq 1}$ de números positivos que crezca a $+\infty$ que cumple:

$$\sum_{r=1}^{\infty} \mathbf{E}X_r^2 / C_r^2 < +\infty, \tag{4.8}$$

entonces:

$$\frac{X_1 + \cdots + X_r}{C_r}$$

converge a cero casi seguramente.

Adaptamos este resultado tomando $X_r = W_r$, siendo (4.7) la condición de diferencia de martingalas, y haciendo $C_r = r$, verificándose (4.8) puesto que las variables W_r están uniformemente acotadas en r , cumpliéndose pues el resultado de convergencia buscado. Q.E.D.

Recordemos que la cantidad en estudio es $N_d(s, r)$. Al haber introducido la diferencia de martingalas hemos descompuesto $N_d(s, r)/r$ en dos sumandos:

$$\frac{1}{r}N_d(s, r) = \frac{1}{r} \sum_{i=1}^r I_d(s, i) = \frac{1}{r} \sum_{i=1}^r W_i + \frac{1}{r} \sum_{i=1}^r \mathbf{E}(I_d(s, i) | \mathcal{F}_{i-1}). \quad (4.9)$$

En el Lema 4.1 hemos probado que el primer sumando del último miembro de (4.9) converge casi seguramente a cero, y por tanto el límite casi seguro de $N_d(s, r)/r$ es el mismo que el del segundo sumando del último miembro de (4.9). Debemos pues estudiar ahora el límite de:

$$\frac{1}{r} \sum_{i=1}^r \mathbf{E}(I_d(s, i) | \mathcal{F}_{i-1}).$$

Se verifica el siguiente:

LEMA 4.2: *Existe una función $q_d(\cdot)$ definida sobre $d^{-1}(0)$ tal que, para todo $i > 1$:*

$$\mathbf{E}(I_d(s, i) | \mathcal{F}_{i-1}) = q_d(s^{(K_{i-1}^d)}).$$

Existe una distribución de probabilidad \mathbf{p} sobre $d^{-1}(0)$, que es la distribución estacionaria de una cadena de Markov irreducible y aperiódica con conjunto de estados $d^{-1}(0)$, tal que se cumple, cuando $r \rightarrow +\infty$:

$$\frac{1}{r} \sum_{i=1}^r \mathbf{E}(I_d(s, i) | \mathcal{F}_{i-1}) \xrightarrow{\text{cs}} \mathbf{E}_{\mathbf{p}} q_d,$$

donde $\mathbf{E}_{\mathbf{p}}$ significa que la toma de esperanzas es respecto a \mathbf{p} .

DEMOSTRACIÓN: La variable $I_d(s, i)$ es el indicador de que en el i -ésimo tramo se haya visitado el estado s . Por tanto:

$$\mathbf{E}(I_d(s, i) | \mathcal{F}_{i-1})$$

es la probabilidad, condicionada por los $i - 1$ primeros tramos, de que se observe el estado s en el tramo i -ésimo. Del carácter markoviano de la muestra observada:

$$s^{(1)}, s^{(2)}, \dots, \quad (4.10)$$

se deduce que dicha probabilidad condicionada es función únicamente del último estado del tramo $(i - 1)$ -ésimo, de donde, para $i \geq 2$:

$$\mathbf{E}(I_d(s, i) | \mathcal{F}_{i-1}) = q_d(s^{(K_{i-1}^d)}), \quad (4.11)$$

y la función q_d no depende de i por la homogeneidad de la cadena observada. Ha quedado pues probada la primera parte del enunciado.

Consideremos la sucesión de variables aleatorias que definen los estados finales de cada tramo:

$$s^{(K_1^d)}, s^{(K_2^d)}, s^{(K_3^d)}, \dots \quad (4.12)$$

Estas variables aleatorias toman valores en $d^{-1}(0)$, y es claro que se trata de una cadena de Markov homogénea. Se deduce del hecho de que la cadena con matriz de transición \tilde{P} es irreducible y aperiódica que la cadena (4.12) también es irreducible y aperiódica. Sea \mathbf{p} su distribución estacionaria.

Hay que observar que aunque el primer estado de la muestra (4.10) sea fijo, el primer estado de (4.12) sigue una cierta distribución inicial.

Para la cadena de Markov (4.12) se cumple (cf. BILLINGSLEY, p.6, Teorema 1.1), haciendo:

$$\varphi(x_1, x_2) = q_d(x_1),$$

en la notación de BILLINGSLEY, el siguiente resultado:

$$\frac{1}{r} \sum_{i=1}^r q_d(s^{(K_i^d)}) \xrightarrow{\text{cs}} \mathbf{E}_{\mathbf{p}} q_d, \quad (4.13)$$

cuando $r \rightarrow +\infty$, puesto que la función q_d es acotada. Este resultado es cierto independientemente de la distribución inicial de la cadena (4.12).

Por tanto:

$$\frac{1}{r} \sum_{i=1}^r \mathbf{E}(I_d(s, i) | \mathcal{F}_{i-1}) = \frac{1}{r} \mathbf{E} I_d(s, 1) + \frac{r-1}{r} \cdot \frac{1}{r-1} \sum_{i=1}^{r-1} q_d(s^{(K_i^d)}), \quad (4.14)$$

donde se ha aplicado (4.11). De (4.13) y (4.14) se concluye que:

$$\frac{1}{r} \sum_{i=1}^r \mathbf{E}(I_d(s, i) | \mathcal{F}_{i-1}) \xrightarrow{\text{cs}} \mathbf{E}_{\mathbf{p}} q_d,$$

cuando $r \rightarrow +\infty$.

Q.E.D.

Ya hemos estudiado el límite de cada uno de los términos de la descomposición (4.9).

PROPOSICIÓN 4.3: *Se verifica el siguiente límite:*

$$\frac{1}{r} N_d(s, r) \xrightarrow{\text{cs}} \mathbf{E}_{\mathbf{p}} q_d,$$

cuando $r \rightarrow +\infty$.

DEMOSTRACIÓN: Es una consecuencia inmediata de la expresión (4.9) y de los lemas 4.1 y 4.2. Q.E.D.

Para determinar el límite de $\hat{v}_d^{(r)}(s)$ debemos analizar el comportamiento de:

$$\frac{1}{r} \sum_{i=1}^r I_d(s, i) g(s^{(K_i^d)}). \quad (4.15)$$

Hay que observar que el término $g(s^{(K_i^d)})$ se puede escribir de la forma:

$$g(s^{(K_i^d)}) = \sum_{d(s')=0} I_{\{s^{(K_i^d)}=s'\}} g(s'),$$

distinguiendo cuál es el estado final del i -ésimo tramo. El sumatorio con subíndice $d(s') = 0$ significa que se hace la suma para todos los estados $s' \in \bar{S}$ tales que $d(s') = 0$. Por tanto, la expresión (4.15) se escribe:

$$\sum_{d(s')=0} g(s') \left(\frac{1}{r} \sum_{i=1}^r I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) \right). \quad (4.16)$$

Al aplicar la política estacionaria d^∞ se obtiene una cadena de Markov homogénea que tiene como estados recurrentes los estados de $d^{-1}(0)$ y como estados de transición los estados de $d^{-1}(1)$. Fijados dos estados s y s' tales que:

$$d(s) = 1 \quad \text{y} \quad d(s') = 0,$$

definimos la probabilidad:

$$\pi_{ss'}^d,$$

como la probabilidad, para la cadena de Markov inducida por d , de que, partiendo del estado s , la absorción por los estados recurrentes se haga por el estado s' .

PROPOSICIÓN 4.4: *Se cumple el siguiente límite:*

$$\frac{1}{r} \sum_{i=1}^r I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) \xrightarrow{\text{cs}} \pi_{ss'}^d \cdot \mathbf{E}_{\mathbf{p}} q_d,$$

cuando $r \rightarrow +\infty$.

DEMOSTRACIÓN: El método de demostración que se seguirá es el mismo que el que se empleó en el Lema 4.2. Hay que observar que la variable aleatoria:

$$I_{\{s^{(K_i^d)}=s'\}} I_d(s, i)$$

es \mathcal{F}_i -medible y está acotada por uno. Definiendo:

$$Y_i = I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) - \mathbf{E}(I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) | \mathcal{F}_{i-1}),$$

para $i \geq 1$, resulta que la sucesión $(Y_i)_{i \geq 1}$ es una diferencia de martingalas uniformemente acotada. Aplicando el teorema de ley fuerte de los grandes números para diferencias de martingalas enunciado en la demostración del Lema 4.1, se obtiene que:

$$\frac{1}{r} \sum_{i=1}^r \left(I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) - \mathbf{E}(I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) | \mathcal{F}_{i-1}) \right) \xrightarrow{\text{cs}} 0,$$

cuando $r \rightarrow +\infty$. Puesto que:

$$\frac{1}{r} \sum_{i=1}^r I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) = \frac{1}{r} \sum_{i=1}^r Y_i + \frac{1}{r} \sum_{i=1}^r \mathbf{E}(I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) | \mathcal{F}_{i-1}), \quad (4.17)$$

resulta que para demostrar esta proposición, debemos estudiar el comportamiento asintótico del segundo sumando del segundo miembro de (4.17).

Se tiene que:

$$\mathbf{E}(I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) | \mathcal{F}_{i-1})$$

es la probabilidad, condicionada por los $i-1$ primeros tramos, de que en el i -ésimo tramo se pase por el estado s y la absorción por los estados de $d^{-1}(0)$ tenga lugar por el estado s' , hablando en términos de la cadena de Markov inducida por d . Se cumple, para $i \geq 2$:

$$\mathbf{E}(I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) | \mathcal{F}_{i-1}) = q_d(s^{(K_{i-1}^d)}) \pi_{ss'}^d.$$

Esta igualdad resulta muy intuitiva ya que, partiendo del estado $s^{(K_{i-1}^d)}$, que es una información conocida puesto que estamos condicionando por \mathcal{F}_{i-1} , la probabilidad de que antes de que se produzca la absorción por los estados de $d^{-1}(0)$ la cadena pase por el estado s es igual a $q_d(s^{(K_{i-1}^d)})$, como se justificó en el Lema 4.2; una vez en el estado s , la probabilidad de que el i -ésimo tramo termine en el estado s' es igual a la probabilidad de que, partiendo de s , la absorción por los estados de $d^{-1}(0)$ se haga por s' , que es: $\pi_{ss'}^d$. Para hacer totalmente rigurosa esta demostración hay que condicionar por el primer instante en el que se llega al estado s antes de terminar el i -ésimo tramo.

Así:

$$\begin{aligned} & \frac{1}{r} \sum_{i=1}^r \mathbf{E}(I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) | \mathcal{F}_{i-1}) = \\ & = \frac{1}{r} \mathbf{E}(I_{\{s^{(K_1^d)}=s'\}} I_d(s, 1)) + \pi_{ss'}^d \frac{r-1}{r} \cdot \frac{1}{r-1} \sum_{i=1}^{r-1} q_d(s^{(K_i^d)}), \end{aligned}$$

cuyo límite, en virtud del Lema 4.2, es igual a $\pi_{ss'}^d \cdot \mathbf{E}_{\mathbf{P}} q_d$, deduciéndose el resultado buscado. Q.E.D.

Combinando la Proposición 4.3 y la Proposición 4.4 se obtiene el siguiente resultado:

TEOREMA 4.5: *Fijados una regla $d \in D$ y un estado $s \in S$ para el que $d(s) = 1$, el estimador del valor de la regla d en s verifica:*

$$\hat{v}_d^{(r)}(s) \xrightarrow{\text{cs}} v_d(s),$$

cuando $r \rightarrow +\infty$.

DEMOSTRACIÓN: El estimador $\hat{v}_d^{(r)}(s)$ se puede poner como el cociente de las cantidades:

$$\frac{1}{r} \sum_{i=1}^r I_d(s, i) g(s^{(K_i^d)}) \tag{4.18}$$

y

$$\frac{1}{r} N_d(s, r). \tag{4.19}$$

El numerador, como se observó en (4.16), se escribe:

$$\sum_{d(s')=0} g(s') \left(\frac{1}{r} \sum_{i=1}^r I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) \right),$$

y aplicando la Proposición 4.4, el límite de esta expresión, cuando $r \rightarrow +\infty$, es:

$$\sum_{d(s')=0} \pi_{ss'}^d g(s') \cdot \mathbf{E}_{\mathbf{P}} q_d,$$

que resulta ser igual a:

$$v_d(s) \cdot \mathbf{E}_{\mathbf{P}} q_d.$$

Del denominador (4.19) se probó en la Proposición 4.3 que convergía con probabilidad uno a $\mathbf{E}_{\mathbf{P}} q_d$, que es una cantidad estrictamente positiva.

Resulta pues que el cociente de (4.18) y (4.19) converge a $v_d(s)$ casi seguramente. Q.E.D.

Resumamos en pocas palabras cómo se ha demostrado esta convergencia. Numerador y denominador del estimador se han separado en dos partes: una relativa a una diferencia de martingalas, y un resto. La parte de la diferencia de martingalas converge a cero, con lo cual el límite casi seguro es igual al límite del término del resto. Se ha probado que el resto es un proceso relativo a una cadena de Markov anidada en la cadena de Markov original, estableciéndose así su límite.

Convergencia en ley

Una vez probada la convergencia casi segura del estimador por el método de las martingalas, vamos a aplicar esta misma técnica para tratar de obtener un resultado de convergencia en distribución.

Consideramos, bajo las mismas hipótesis, la sucesión:

$$\frac{\sum_{i=1}^r I_d(s, i) g(s^{(K_i^d)})}{\sqrt{r} N_d(s, r)},$$

para $r \geq 1$, e investigamos su comportamiento en ley. Esta última expresión se puede escribir:

$$\frac{\frac{1}{\sqrt{r}} \sum_{i=1}^r I_d(s, i) g(s^{(K_i^d)})}{N_d(s, r)/r}, \tag{4.20}$$

y vamos a estudiar el límite en distribución de (4.20) analizando el comportamiento asintótico de su numerador. Nos basaremos en el siguiente teorema (cf. Corolario 3.1, HALL et al., p.58):

TEOREMA 4.6: *Se considera $(S_{ri}, \mathcal{F}_{ri}, 1 \leq i \leq r, r \geq 1)$ un arreglo triangular que sea, para cada r fijo, una martingala de media cero y de cuadrado integrable, con diferencia de martingalas asociada:*

$$(X_{ri}, 1 \leq i \leq r, r \geq 1).$$

Supongamos que se verifican las condiciones:

- si $r \geq 1$ y $1 \leq i \leq r$:

$$\mathcal{F}_{ri} \subseteq \mathcal{F}_{r+1,i}; \quad (4.21)$$

- propiedad de Lindeberg condicional: para todo $\varepsilon > 0$ se cumple:

$$\sum_{i=1}^r \mathbf{E}(X_{ri}^2 I_{\{|X_{ri}| > \varepsilon\}} | \mathcal{F}_{r,i-1}) \xrightarrow{\mathbf{P}} 0, \quad (4.22)$$

cuando $r \rightarrow +\infty$;

- existe una variable aleatoria finita casi seguramente: ξ^2 , tal que las varianzas condicionales convergen en probabilidad a ξ^2 , es decir:

$$\sum_{i=1}^r \mathbf{E}(X_{ri}^2 | \mathcal{F}_{r,i-1}) \xrightarrow{\mathbf{P}} \xi^2, \quad (4.23)$$

cuando $r \rightarrow +\infty$.

Se cumple entonces que:

$$S_{rr} = \sum_{i=1}^r X_{ri},$$

converge en distribución, cuando $r \rightarrow +\infty$, a una variable aleatoria con función característica, en la variable $t \in \mathbb{R}$:

$$\mathbf{E} \exp\left\{-\frac{1}{2}\xi^2 t^2\right\}.$$

La sucesión de variables aleatorias $(Z_i)_{i \geq 1}$, que está adaptada a la filtración $\{\mathcal{F}_i\}_{i \geq 1}$ de sub- σ -álgebras de \mathcal{F} , viene definida por:

$$Z_i = I_d(s, i)g(s^{(K_i^d)}).$$

Se tiene entonces que el numerador de (4.20) se escribe:

$$\frac{1}{\sqrt{r}}(Z_1 + \cdots + Z_r).$$

Para hacer aparecer una diferencia de martingalas se considera la descomposición:

$$Z_i = (Z_i - \mathbf{E}(Z_i|\mathcal{F}_{i-1})) + \mathbf{E}(Z_i|\mathcal{F}_{i-1}),$$

siendo la sucesión, para $i \geq 1$, de término:

$$Z_i - \mathbf{E}(Z_i|\mathcal{F}_{i-1}), \quad (4.24)$$

una diferencia de martingalas de media cero adaptada a la filtración $\{\mathcal{F}_i\}_{i \geq 0}$.

No vamos a aplicar el Teorema 4.6 directamente a esta diferencia de martingalas. El motivo es el siguiente. Utilizando el Teorema 4.6, se podría llegar a determinar la distribución límite de (4.24), pero lo que ocurre es que aún tendríamos que analizar la sucesión de término:

$$\mathbf{E}(Z_i|\mathcal{F}_{i-1}),$$

para la cual, de nuevo, podríamos obtener un teorema central del límite. Nos encontraríamos así con dos sucesiones asintóticamente normales, siendo sin embargo de interés la distribución límite de su suma, y no sus distribuciones límite marginales. Es por este motivo que debemos considerar todos los términos simultáneamente para poder determinar la distribución asintótica del error de estimación.

El siguiente resultado es obvio. Se probó en el Lema 4.1 que la sucesión:

$$W_i = I_d(s, i) - \mathbf{E}(I_d(s, i)|\mathcal{F}_{i-1}),$$

definida para $i \geq 1$, es una diferencia de martingalas adaptada a la filtración $\{\mathcal{F}_i\}_{i \geq 0}$ de sub- σ -álgebras de \mathcal{F} . Si multiplicamos la sucesión $(W_i)_{i \geq 1}$ por una constante, conseguimos de nuevo una diferencia de martingalas. Al restar dos diferencias de martingalas de media cero adaptadas a una misma filtración, se obtiene una nueva diferencia de martingalas con esas mismas propiedades. En particular, la sucesión de término:

$$Z_i - \mathbf{E}(Z_i|\mathcal{F}_{i-1}) - v_d(s) \cdot W_i,$$

para $i \geq 1$, es una diferencia de martingalas de media cero para la filtración de sub- σ -álgebras de $\mathcal{F}: \{\mathcal{F}_i\}_{i \geq 0}$. Con vistas a aplicar el Teorema 4.6 a esta diferencia de martingalas, definimos la siguiente sucesión de índice doble:

$$X_{ri} = \frac{1}{\sqrt{r}} (Z_i - \mathbf{E}(Z_i | \mathcal{F}_{i-1}) - v_d(s) \cdot W_i),$$

para $1 \leq i \leq r$ y $r \geq 1$. El siguiente resultado prueba además que X_{ri} admite una expresión mucho más sencilla, a partir de la diferencia entre el pago al final del i -ésimo tramo y el valor de d en s , cuando el estado s se ha observado en el i -ésimo tramo.

PROPOSICIÓN 4.7: *Se cumple, para $1 \leq i \leq r$ y $r \geq 1$, que:*

$$X_{ri} = \frac{1}{\sqrt{r}} \left(I_d(s, i) (g(s^{(K_i^d)}) - v_d(s)) \right),$$

y que:

$$\sum_{i=1}^r X_{ri} \xrightarrow{d} \mathcal{N}(0, \cdot),$$

cuando $r \rightarrow +\infty$.

DEMOSTRACIÓN: La variable Z_i se escribe de la siguiente manera:

$$Z_i = I_d(s, i) g(s^{(K_i^d)}) = \sum_{d(s')=0} g(s') I_{\{s^{(K_i^d)}=s'\}} I_d(s, i),$$

y por tanto:

$$\mathbf{E}(Z_i | \mathcal{F}_{i-1}) = \sum_{d(s')=0} g(s') \mathbf{E}(I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) | \mathcal{F}_{i-1}).$$

Se probó en la demostración de la Proposición 4.4 que, cuando $i \geq 2$:

$$\mathbf{E}(I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) | \mathcal{F}_{i-1}) = q_d(s^{(K_{i-1}^d)}) \pi_{ss'}^d,$$

luego, para $i \geq 2$:

$$\begin{aligned} \mathbf{E}(Z_i | \mathcal{F}_{i-1}) &= \sum_{d(s')=0} \pi_{ss'}^d g(s') q_d(s^{(K_{i-1}^d)}) \\ &= v_d(s) q_d(s^{(K_{i-1}^d)}). \end{aligned}$$

Por otro lado, se demostró en el Lema 4.2 que:

$$W_i = I_d(s, i) - q_d(s^{(K_{i-1}^d)}),$$

cuando $i \geq 2$. Se cumple pues que:

$$Z_i - \mathbf{E}(Z_i | \mathcal{F}_{i-1}) - v_d(s) \cdot W_i$$

es igual a:

$$I_d(s, i)g(s^{(K_i^d)}) - v_d(s)q_d(s^{(K_{i-1}^d)}) - v_d(s)(I_d(s, i) - q_d(s^{(K_{i-1}^d)})),$$

para $i \geq 2$, de donde:

$$Z_i - \mathbf{E}(Z_i | \mathcal{F}_{i-1}) - v_d(s) \cdot W_i = I_d(s, i)(g(s^{(K_i^d)}) - v_d(s)).$$

Cuando $i = 1$, la expresión es la misma aunque no es válida la demostración que se ha hecho. El caso particular $i = 1$ se trata aparte, obteniendo el mismo resultado.

Se ha probado pues la primera parte del enunciado. Obsérvese que la conclusión que se obtiene es que la sucesión de términos:

$$I_d(s, i)(g(s^{(K_i^d)}) - v_d(s)),$$

siendo $i \geq 1$, es una diferencia de martingalas de media cero para la filtración $\{\mathcal{F}_i\}_{i \geq 1}$.

Apliquemos ahora el Teorema 4.6 a la diferencia de martingalas:

$$X_{ri} = \frac{1}{\sqrt{r}} \left(I_d(s, i)(g(s^{(K_i^d)}) - v_d(s)) \right),$$

definida para $1 \leq i \leq r$ y $r \geq 1$.

En el enunciado del Teorema 4.6 tomamos:

$$\mathcal{F}_{ri} = \mathcal{F}_i,$$

de tal forma que se verifica la condición (4.21), cf. p.174.

Al cumplirse que el valor de la regla d en s : $v_d(s)$, es un promedio de los valores $g(s')$, siendo éstos mayores o iguales que cero, se tiene que:

$$|g(s^{(K_i^d)}) - v_d(s)| \leq \max_{s' \in \bar{S}} g(s'),$$

y por tanto:

$$I_d(s, i)|g(s^{(K_i^d)}) - v_d(s)| \leq \max_{s' \in \bar{S}} g(s'),$$

para todo $i \geq 1$. Resulta pues que, fijado $\varepsilon > 0$ y tomando r suficientemente grande, el conjunto:

$$\{|X_{ri}| > \varepsilon\} = \{I_d(s, i)|g(s^{(K_i^d)}) - v_d(s)| > \varepsilon\sqrt{r}\},$$

es vacío para $1 \leq i \leq r$, verificándose pues la propiedad de Lindeberg condicional (4.22), cf. p.174.

Falta por comprobar la convergencia (4.23), cf. p.174, para lo cual se define la varianza condicional, siendo $r \geq 1$:

$$V_r^2 = \sum_{i=1}^r \mathbf{E}(X_{ri}^2 | \mathcal{F}_{i-1}) = \frac{1}{r} \sum_{i=1}^r \mathbf{E}(I_d(s, i)(g(s^{(K_i^d)}) - v_d(s))^2 | \mathcal{F}_{i-1}).$$

Fijemos un entero $i \geq 1$. Se cumple que:

$$\mathbf{E}(I_d(s, i)(g(s^{(K_i^d)}) - v_d(s))^2 | \mathcal{F}_{i-1})$$

es igual a:

$$\sum_{d(s')=0} (g(s') - v_d(s))^2 \mathbf{E}(I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) | \mathcal{F}_{i-1}),$$

que a su vez es igual a:

$$\sum_{d(s')=0} (g(s') - v_d(s))^2 \pi_{ss'}^d \mathbf{E}(I_d(s, i) | \mathcal{F}_{i-1}).$$

En conclusión:

$$V_r^2 = \left(\sum_{d(s')=0} \pi_{ss'}^d (g(s') - v_d(s))^2 \right) \cdot \frac{1}{r} \sum_{i=1}^r \mathbf{E}(I_d(s, i) | \mathcal{F}_{i-1}),$$

y aplicando el Lema 4.2, se tiene que:

$$V_r^2 \xrightarrow{\text{cs}} \left(\sum_{d(s')=0} \pi_{ss'}^d (g(s') - v_d(s))^2 \right) \mathbf{E}_{\mathbf{P}} q_d,$$

cuando $r \rightarrow +\infty$. Si definimos:

$$V_{d,s}(g) = \sum_{d(s')=0} \pi_{ss'}^d (g(s') - v_d(s))^2,$$

que es igual a la varianza de la variable aleatoria g respecto de las probabilidades de absorción, es decir, es la varianza de la variable aleatoria que toma los valores $g(s')$ con probabilidades $\pi_{ss'}^d$, resulta que:

$$V_r^2 \xrightarrow{\text{cs}} V_{d,s}(g)\mathbf{E}_{\mathbf{P}}q_d,$$

cuando $r \rightarrow +\infty$. Se ha probado que la sucesión de varianzas condicionales converge casi seguramente, luego también en probabilidad, a una constante finita, es decir, la variable ξ^2 es degenerada, de donde, en virtud del Teorema 4.6:

$$\sum_{i=1}^r X_{ri}$$

converge en distribución a una variable aleatoria normal uni-dimensional de media cero y de varianza:

$$V_{d,s}(g)\mathbf{E}_{\mathbf{P}}q_d.$$

Q.E.D.

El resultado siguiente es el teorema central del límite para los estimadores $(\hat{v}_d^{(r)}(s))_{r \geq 1}$.

TEOREMA 4.8: *Dados una regla $d \in D$ y un estado s tal que $d(s) = 1$, se cumple que:*

$$\sqrt{r}(\hat{v}_d^{(r)}(s) - v_d(s)) \xrightarrow{d} \mathcal{N}(0, \cdot),$$

cuando $r \rightarrow +\infty$.

DEMOSTRACIÓN: Vamos a manipular la expresión de $\sqrt{r}(\hat{v}_d^{(r)}(s) - v_d(s))$ hasta hacer aparecer la sucesión (X_{ri}) de la Proposición 4.7.

Se tiene, en primer lugar, que:

$$\frac{1}{\sqrt{r}} \left(\sum_{i=1}^r Z_i - v_d(s)N_d(s, r) \right) \tag{4.25}$$

es igual a:

$$\frac{1}{\sqrt{r}} \sum_{i=1}^r (Z_i - \mathbf{E}(Z_i|\mathcal{F}_{i-1}) + \mathbf{E}(Z_i|\mathcal{F}_{i-1}) - v_d(s)I_d(s, i)),$$

y dado que:

$$\mathbf{E}(Z_i|\mathcal{F}_{i-1}) = v_d(s)\mathbf{E}(I_d(s, i)|\mathcal{F}_{i-1}),$$

resulta que (4.25) se escribe:

$$\frac{1}{\sqrt{r}} \sum_{i=1}^r (Z_i - \mathbf{E}(Z_i | \mathcal{F}_{i-1}) - v_d(s) \cdot W_i) = \sum_{i=1}^r X_{ri}.$$

Se cumple que:

$$\begin{aligned} \sqrt{r}(\hat{v}_d^{(r)}(s) - v_d(s)) &= \frac{r}{N_d(s, r)} \cdot \frac{1}{\sqrt{r}} \left(\sum_{i=1}^r Z_i - v_d(s) N_d(s, r) \right) \\ &= \frac{r}{N_d(s, r)} \sum_{i=1}^r X_{ri}. \end{aligned}$$

En la Proposición 4.3 se probó que:

$$\frac{N_d(s, r)}{r}$$

converge casi seguramente a $\mathbf{E}_{\mathbf{p}} q_d$, que es una constante positiva. De aplicar el resultado de la Proposición 4.7, se deduce que:

$$\sqrt{r}(\hat{v}_d^{(r)}(s) - v_d(s))$$

tiende en distribución, cuando $r \rightarrow +\infty$, a una variable aleatoria normal de media cero y de varianza:

$$V_{d,s}(g)(\mathbf{E}_{\mathbf{p}} q_d)^{-1}.$$

Q.E.D.

Así pues, el método de las martingalas ha dado buenos resultados para el análisis del estimador del valor de una regla. Cabe preguntarse por qué este método de las martingalas ha permitido alcanzar los objetivos propuestos, cuando su cabida en este problema, que es de naturaleza markoviana, no parece del todo evidente. Lo primero que hay que observar es que las diferencias de martingalas no han surgido de manera natural sino que se han introducido de forma artificial: restando a una variable aleatoria su esperanza condicionada respecto de toda la información disponible hasta la anterior etapa del muestreo. El efecto que ha tenido esta operación es que las cantidades en estudio se han descompuesto como suma de dos términos: una diferencia de martingalas y una esperanza condicionada. La diferencia de martingalas se estudiaba utilizando la Teoría de Martingalas. La esperanza

condicionada, debido al carácter markoviano del problema, se ha podido expresar siempre como función de las probabilidades de transición de cadenas contenidas en el proceso original, y es por eso que el que haya aparecido este sumando de más no ha resultado un inconveniente.

En la demostración de la convergencia casi segura, se probó que el término de la diferencia de martingalas convergía casi seguramente a cero, siendo pues el comportamiento límite dominante el del término residual de esperanza condicionada. Para la convergencia en ley, de haberse tratado de aplicar la misma técnica, el término de diferencia de martingalas no hubiera tendido esta vez a cero sino que hubiese tenido también límite en distribución. En este caso, lo que ha ocurrido es que poniendo el término residual de la esperanza condicionada en combinación con el denominador aleatorio de $\hat{v}_d^{(r)}(s)$, se ha obtenido una nueva diferencia de martingalas, expresándose el error de estimación normalizado como un cociente cuyo numerador es una diferencia de martingalas convergente en distribución.

Para el estudio del estimador del valor de una regla proponemos un segundo método. Este método se basa únicamente en la naturaleza markoviana del problema de parada óptima. Si bien las demostraciones utilizando la técnica de martingalas son, en general, ciertamente elegantes y contundentes, las demostraciones que vamos a hacer por el método markoviano son, en algún sentido, más directas.

4.2.3 Convergencia del estimador. Segundo método

Planteamiento

Para estudiar este segundo método debemos definir una nueva cadena de Markov. Esta cadena está asociada a una regla de decisión d y a un estado s con $d(s) = 1$, y la denotaremos por $C^d(s)$.

Su conjunto de estados es un subconjunto de:

$$d^{-1}(0) \times \{0, 1\}.$$

Decimos que la cadena $C^d(s)$ está, en la etapa i -ésima, en el estado $(s', 0)$ cuando el estado final del i -ésimo tramo es s' y a lo largo de ese tramo la cadena original no ha pasado por el estado s . La cadena $C^d(s)$ está en el estado $(s', 1)$ en la etapa i -ésima cuando el estado final del i -ésimo tramo es s' , y a lo largo de ese tramo la cadena original ha pasado por el estado s .

Es claro, según la definición dada, que el proceso $C^d(s)$ es una cadena de Markov homogénea. La probabilidad de transición del estado (s', α) al

estado (s'', β) , que escribiremos:

$$p(s'', \beta | s', \alpha),$$

no depende de α , puesto que para saber lo que ocurre en el i -ésimo tramo, la única información que necesitamos de la historia del proceso es el estado final del $(i - 1)$ -ésimo tramo, y por tanto:

$$p(s'', \beta | s', \alpha) = p(s'', \beta | s').$$

Hay que hacer dos observaciones. Pese a que el estado inicial $s^{(1)}$ de la cadena de Markov que se muestra es fijo, el primer estado de la cadena $C^d(s)$ es aleatorio, dependiendo de cómo sea el primer tramo que se observa, por lo que la cadena $C^d(s)$ tiene asociada una distribución de probabilidad inicial, aunque todos los resultados que van a demostrarse son independientes de esa probabilidad inicial. La segunda observación se refiere a que se ha dicho que el conjunto de estados de $C^d(s)$ es un subconjunto de:

$$d^{-1}(0) \times \{0, 1\}.$$

Analicemos dos ejemplos.

Si el estado s_1 es tal que:

$$p(s_2 | s_1) = 1, \quad d(s_1) = 1 \quad \text{y} \quad d(s_2) = 0,$$

entonces la cadena $C^d(s_1)$ no pasará por ningún estado de la forma $(s', 1)$ salvo el $(s_2, 1)$. Se eliminan estos estados superfluos para procurar que la cadena $C^d(s_1)$ sea irreducible.

Para dar otro ejemplo, supongamos que:

$$P \in T_m \times \cdots \times T_m.$$

Entonces el conjunto de estados de $C^d(s)$ es todo $d^{-1}(0) \times \{0, 1\}$, sin posibilidad de reducción de estados. De ahora en adelante supondremos que la cadena $C^d(s)$ es irreducible y aperiódica.

El resultado siguiente es cierto en general en una cadena finita, irreducible y aperiódica.

TEOREMA 4.9: *Se considera una cadena de Markov finita, homogénea, irreducible y aperiódica, con K estados y con probabilidades estacionarias:*

$$\pi_1, \dots, \pi_K.$$

Se observa una muestra de tamaño n de la cadena, y las frecuencias muestrales son:

$$\hat{p}_1^{(n)}, \dots, \hat{p}_K^{(n)}.$$

Se verifican los siguientes límites:

$$\hat{p}_i^{(n)} \xrightarrow{\text{cs}} \pi_i, \text{ para } i \in \{1, \dots, K\};$$

$$\sqrt{n}(\hat{p}_i^{(n)} - \pi_i)_{1 \leq i \leq K} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma),$$

cuando $n \rightarrow +\infty$, independientemente de la distribución inicial de la cadena.

DEMOSTRACIÓN: La demostración del límite casi seguro es consecuencia del Teorema 1.1 de BILLINGSLEY, p.6, que ya ha sido aplicado varias veces.

El resultado del límite en distribución, así como la expresión de Σ , se encuentran en el Capítulo IV, §43, p.213, de ROMANOVSKY. Q.E.D.

Convergencia casi segura

Basando las demostraciones en los resultados del Teorema 4.9, enunciamos de nuevo el Teorema 4.5.

TEOREMA 4.10: *Fijados una regla $d \in D$ y un estado $s \in S$ para el que $d(s) = 1$, el estimador del valor de la regla d en s verifica:*

$$\hat{v}_d^{(r)}(s) \xrightarrow{\text{cs}} v_d(s),$$

cuando $r \rightarrow +\infty$.

DEMOSTRACIÓN: El estimador de la cantidad $v_d(s)$ tras la observación de r tramos es:

$$\hat{v}_d^{(r)}(s) = \frac{\sum_{i=1}^r I_d(s, i)g(s^{(K_i^d)})}{N_d(s, r)}.$$

Se tiene que el denominador de esta expresión es:

$$N_d(s, r) = \sum_{i=1}^r I_d(s, r),$$

que es el número de tramos, de entre los r primeros, en los que se ha observado el estado s . En términos de la cadena $C^d(s)$, es el número de veces que, en las r primeras etapas, la cadena ha visitado estados de la forma:

$$(s', 1),$$

siendo $d(s') = 0$. Utilizando la notación de las frecuencias introducida en el enunciado del Teorema 4.9, resulta que:

$$N_d(s, r) = r \sum_{d(s')=0} \hat{p}_{(s',1)}^{(r)},$$

y por tanto, según el Teorema 4.9, cuando $r \rightarrow +\infty$:

$$\frac{N_d(s, r)}{r} \xrightarrow{\text{cs}} \sum_{d(s')=0} \pi_{(s',1)}, \quad (4.26)$$

que, por la ergodicidad de la cadena $C^d(s)$, es una cantidad positiva. El sumatorio:

$$\sum_{d(s')=0} \pi_{(s',1)}$$

se extiende sólo a los estados s' tales que $(s', 1)$ está en el conjunto de estados de $C^d(s)$, aunque esto no se hará explícito por no recargar la notación. Obsérvese que en la Proposición 4.3 se probó que este límite era igual a $\mathbf{E}_{\mathbf{p}}q_d$, es decir:

$$\sum_{d(s')=0} \pi_{(s',1)} = \mathbf{E}_{\mathbf{p}}q_d.$$

Por otro lado:

$$\sum_{i=1}^r I_d(s, i)g(s^{(K_i^d)}) = \sum_{d(s')=0} g(s') \sum_{i=1}^r I_{\{s^{(K_i^d)}=s'\}} I_d(s, i),$$

y $I_{\{s^{(K_i^d)}=s'\}} I_d(s, i)$ es el indicador de que la cadena $C^d(s)$ está en el estado $(s', 1)$ en la etapa i -ésima, por tanto:

$$\sum_{i=1}^r I_d(s, i)g(s^{(K_i^d)}) = r \sum_{d(s')=0} g(s') \hat{p}_{(s',1)}^{(r)},$$

luego:

$$\frac{1}{r} \sum_{i=1}^r I_d(s, i) g(s^{(K_i^d)}) \xrightarrow{\text{cs}} \sum_{d(s')=0} g(s') \pi_{(s',1)}, \quad (4.27)$$

cuando $r \rightarrow +\infty$.

Dividiendo la expresión (4.27) por la (4.26) resulta que:

$$\hat{v}_d^{(r)}(s) \xrightarrow{\text{cs}} \sum_{d(s')=0} g(s') \left(\pi_{(s',1)} / \sum \pi_{(s'',1)} \right).$$

El cociente:

$$\frac{\pi_{(s',1)}}{\sum_{d(s'')=0} \pi_{(s'',1)}}$$

es la probabilidad estacionaria condicionada de que un tramo que pasa por el estado s acabe en el estado s' . Es igual pues a la probabilidad de que, en la cadena inducida por d , partiendo de s , la absorción por los estados de $d^{-1}(0)$ se haga por s' , es decir:

$$\frac{\pi_{(s',1)}}{\sum_{d(s'')=0} \pi_{(s'',1)}} = \pi_{ss'}^d, \quad (4.28)$$

de donde:

$$\hat{v}_d^{(r)}(s) \xrightarrow{\text{cs}} \sum_{d(s')=0} g(s') \pi_{ss'}^d = v_d(s),$$

cuando $r \rightarrow +\infty$.

Q.E.D

Hay que hacer notar que los estimadores, gracias a considerar la cadena $C^d(s)$, adoptan, a partir de las frecuencias observadas, expresiones sencillas.

Este nuevo método de la cadena de Markov $C^d(s)$ también da respuesta al problema de la distribución límite del estimador del valor de d .

Convergencia en ley

El resultado de convergencia en ley, análogo al Teorema 4.8, es el siguiente.

TEOREMA 4.11: *Dados una regla $d \in D$ y un estado s tal que $d(s) = 1$, se cumple que:*

$$\sqrt{r}(\hat{v}_d^{(r)}(s) - v_d(s)) \xrightarrow{d} \mathcal{N}(0, \cdot),$$

cuando $r \rightarrow +\infty$.

DEMOSTRACIÓN: Vamos a escribir el término $\hat{v}_d^{(r)}(s) - v_d(s)$ de forma adecuada para la demostración:

$$\begin{aligned} \hat{v}_d^{(r)}(s) - v_d(s) &= \frac{1}{N_d(s, r)} \left(\sum_{d(s')=0} g(s') \sum_{i=1}^r I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) \right) - v_d(s) \\ &= \frac{r}{N_d(s, r)} \left(\sum_{d(s')=0} g(s') \hat{p}_{(s',1)}^{(r)} \right) - v_d(s). \end{aligned}$$

Ahora surgen ciertas incomodidades derivadas del hecho de que el denominador de $\hat{v}_d^{(r)}(s)$ es aleatorio. Se tiene que:

$$\hat{v}_d^{(r)}(s) - v_d(s) = \frac{r}{N_d(s, r)} \left(\sum_{d(s')=0} g(s') \hat{p}_{(s',1)}^{(r)} - \frac{N_d(s, r)}{r} v_d(s) \right). \quad (4.29)$$

Tal y como se probó en la demostración del Teorema 4.10:

$$\frac{N_d(s, r)}{r} = \sum_{d(s')=0} \hat{p}_{(s',1)}^{(r)} \xrightarrow{cs} \sum_{d(s')=0} \pi_{(s',1)},$$

y expresamos $\frac{N_d(s, r)}{r} v_d(s)$ de la forma:

$$\left(\frac{N_d(s, r)}{r} - \sum_{d(s')=0} \pi_{(s',1)} \right) v_d(s) + \sum_{d(s')=0} \pi_{(s',1)} v_d(s). \quad (4.30)$$

Si recordamos la igualdad (4.28) a la que se llegó en el Teorema 4.10:

$$\frac{\pi_{(s',1)}}{\sum_{d(s'')=0} \pi_{(s'',1)}} = \pi_{ss'}^d,$$

y teniendo en cuenta que:

$$v_d(s) = \sum_{d(s')=0} \pi_{ss'}^d g(s'),$$

se cumple la identidad fundamental:

$$\sum_{d(s')=0} \pi_{(s',1)} g(s') = \sum_{d(s')=0} \pi_{(s',1)} v_d(s).$$

Podemos pues escribir (4.30) como:

$$v_d(s) \sum_{d(s')=0} (\hat{p}_{(s',1)}^{(r)} - \pi_{(s',1)}) + \sum_{d(s')=0} \pi_{(s',1)} g(s').$$

A partir de esta última expresión y de (4.29), se tiene, tras agrupar términos, que:

$$\hat{v}_d^{(r)}(s) - v_d(s) = \frac{r}{N_d(s, r)} \sum_{d(s')=0} (g(s') - v_d(s)) (\hat{p}_{(s',1)}^{(r)} - \pi_{(s',1)}). \quad (4.31)$$

Aquí se puede constatar la gran utilidad de haber introducido la cadena de Markov $C^d(s)$ ya que la expresión del error de estimación es muy sencilla cuando está escrita relativa a la cadena $C^d(s)$. Se tiene que:

$$\sqrt{r}(\hat{v}_d^{(r)}(s) - v_d(s)) = \frac{r}{N_d(s, r)} \sum_{d(s')=0} (g(s') - v_d(s)) \sqrt{r}(\hat{p}_{(s',1)}^{(r)} - \pi_{(s',1)}).$$

Del Teorema 4.9 se deduce que las variables:

$$\sqrt{r}(\hat{p}_{(s',1)}^{(r)} - \pi_{(s',1)})$$

tienen distribución límite conjunta normal, por lo que:

$$\sum_{d(s')=0} (g(s') - v_d(s)) \sqrt{r}(\hat{p}_{(s',1)}^{(r)} - \pi_{(s',1)})$$

es asintóticamente normal. Puesto que:

$$\frac{N_d(s, r)}{r}$$

converge casi seguramente a una constante positiva, resulta que, cuando el tamaño $r \rightarrow +\infty$:

$$\sqrt{r}(\hat{v}_d^{(r)}(s) - v_d(s))$$

tiende en distribución a una variable aleatoria normal uni-dimensional.

Q.E.D.

Es importante comentar con cierto detalle la expresión (4.31):

$$\hat{v}_d^{(r)}(s) - v_d(s) = \frac{r}{N_d(s, r)} \sum_{d(s')=0} (g(s') - v_d(s))(\hat{p}_{(s',1)}^{(r)} - \pi_{(s',1)}).$$

La peculiaridad del estimador $\hat{v}_d^{(r)}(s)$ es que su denominador:

$$N_d(s, r),$$

es aleatorio. Un estimador de tipo media aritmética, llamémosle “clásico”, tendría como denominador el tamaño muestral: r , siendo el desarrollo estudiado más sencillo, evitando por ejemplo el paso dado en (4.29). La consecuencia de la aleatoriedad del denominador es que aparece un término de más en la expresión (4.31), que es:

$$-v_d(s) \sum_{d(s')=0} (\hat{p}_{(s',1)}^{(r)} - \pi_{(s',1)}),$$

con lo cual, el error de estimación se pone como la suma del error que cometería el estimador “clásico” más un segundo término. En el método de las martingalas habíamos llegado a una conclusión similar.

Nótese que el Teorema 4.11 proporciona la distribución límite de la estimación para un estado en concreto: s . Si queremos obtener la distribución conjunta, para todos los estados s tales que $d(s) = 1$, parece que la idea que habría que utilizar es la de considerar una cadena de Markov que tuviera como espacio de estados un subconjunto de:

$$d^{-1}(0) \times \{0, 1\} \times \overset{|d^{-1}(1)|}{\dots} \times \{0, 1\},$$

para tener la información de todos los estados por los que ha pasado la cadena a lo largo del i -ésimo tramo.

4.2.4 Cambio de espacio muestral

Recordemos que el estimador $\hat{v}_d^{(r)}(s)$ está basado en la observación de los r primeros tramos de la regla d . En (4.5), cf. p.165, habíamos definido un estimador para un tamaño muestral fijo: $\tilde{v}_d^{(n)}(s)$, salvando así el inconveniente de que el parámetro r variaba de una regla a otra, aun estando los estimadores basados en una misma muestra.

Vamos a probar ahora los resultados límite para el estimador $\hat{v}_d^{(n)}(s)$. La demostración de la convergencia casi segura es inmediata. En cambio, el estudio de la distribución asintótica es bastante delicado y requiere volver a emplear resultados de martingalas.

Convergencia casi segura

El siguiente teorema demuestra la convergencia casi segura.

TEOREMA 4.12: *Fijados una regla $d \in D$ y un estado $s \in S$ que cumpla $d(s) = 1$, se tiene:*

$$\tilde{v}_d^{(n)}(s) \xrightarrow{\text{cs}} v_d(s),$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: Para toda trayectoria $\omega \in \Omega$, la sucesión $(r_n(\omega))_{n \geq 1}$ tiende a $+\infty$, y puesto que con probabilidad uno:

$$\hat{v}_d^{(r)}(s)$$

converge a $v_d(s)$, resulta que:

$$\tilde{v}_d^{(n)}(s) = \hat{v}_d^{(r_n)}(s),$$

también tiende casi seguramente a $v_d(s)$.

Q.E.D.

Convergencia en ley

Obsérvese que la sucesión $(\hat{v}_d^{(r)}(s))_{r \geq 1}$ es aleatoria y que también es aleatoria la sucesión $(r_n)_{n \geq 1}$, con lo cual $(\tilde{v}_d^{(n)}(s))_{n \geq 1}$ no es una sucesión aleatoria sub-sucesión de $(\hat{v}_d^{(r)}(s))_{r \geq 1}$, sino que es una sucesión aleatoria sub-sucesión aleatoria de $(\hat{v}_d^{(r)}(s))_{r \geq 1}$, y es por eso que el resultado de convergencia en ley no se puede demostrar siguiendo la técnica del Teorema 4.12. Necesitamos un nuevo teorema central del límite para diferencias de martingalas (cf. Teorema 3.2, HALL et al., p.58), que es el Teorema 1.19, cf. p.50.

TEOREMA 4.13: *Se considera $(S_{ni}, \mathcal{F}_{ni}, 1 \leq i \leq n, n \geq 1)$ un arreglo triangular que sea, para cada n fijo, una martingala de media cero y de cuadrado integrable, con diferencia de martingalas asociada:*

$$(X_{ni}, 1 \leq i \leq n, n \geq 1).$$

Supongamos que se verifican las condiciones:

- si $n \geq 1$ y $1 \leq i \leq n$:

$$\mathcal{F}_{ni} \subseteq \mathcal{F}_{n+1,i}; \quad (4.32)$$

- cuando n tiende a $+\infty$:

$$\max_{1 \leq i \leq n} |X_{ni}| \xrightarrow{\text{P}} 0; \quad (4.33)$$

- se cumple:

$$\sup_{n \geq 1} \mathbf{E}(\max_{1 \leq i \leq n} X_{ni}^2) < +\infty; \quad (4.34)$$

- existe una variable aleatoria finita casi seguramente: ξ^2 , tal que:

$$\sum_{i=1}^n X_{ni}^2 \xrightarrow{\text{P}} \xi^2, \quad (4.35)$$

cuando $n \rightarrow +\infty$.

Se cumple entonces que:

$$S_{nn} = \sum_{i=1}^n X_{ni},$$

converge en distribución, cuando $n \rightarrow +\infty$, a una variable aleatoria con función característica, en la variable $t \in \mathbb{R}$:

$$\mathbf{E} \exp\left\{-\frac{1}{2}\xi^2 t^2\right\}.$$

Hagamos el convenio, que no supone pérdida de generalidad, de que el primer estado que se observa en la muestra es s_{m+1} , es decir:

$$s^{(1)} = s_{m+1},$$

de tal forma que el primer tramo está formado por la primera observación. Esto no afecta al resto de las estimaciones ya que toda regla $d \in D$ verifica que $d(s_{m+1}) = 0$. Hacemos esta hipótesis por comodidad puesto que así evitamos que las variables r_n puedan tomar el valor cero, siendo $r_1 = 1$. El resultado de convergencia en distribución es el siguiente.

TEOREMA 4.14: *Fijados una regla $d \in D$ y un estado $s \in S$ que cumpla $d(s) = 1$, se tiene:*

$$\sqrt{n}(\tilde{v}_d^{(n)}(s) - v_d(s)) \xrightarrow{d} \mathcal{N}(0, \cdot),$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: En la Proposición 4.7 se demostró que la sucesión:

$$I_d(s, i)(g(s^{(K_i^d)}) - v_d(s)),$$

para $i \geq 1$, es una diferencia de martingalas de media cero para la filtración $\{\mathcal{F}_i\}_{i \geq 1}$, y por tanto, la sucesión:

$$\sum_{i=1}^n I_d(s, i)(g(s^{(K_i^d)}) - v_d(s)),$$

para $n \geq 1$, es una martingala de media cero.

Por otro lado, la sucesión de variables aleatorias: $(r_n)_{n \geq 1}$, es monótona no decreciente, puesto que r_n es el número de tramos completos observados en la muestra de tamaño n . Cada variable r_n es un tiempo de parada para la filtración $\{\mathcal{F}_i\}_{i \geq 1}$ de sub- σ -álgebras de \mathcal{F} , pues $\{r_n \leq i\}$ es el conjunto de trayectorias tales que los i primeros tramos tienen longitud mayor o igual que n , es decir:

$$\{r_n \leq i\} = \{K_i^d \geq n\} \in \mathcal{F}_i,$$

ya que la variable K_i^d es \mathcal{F}_i -medible. Cada tiempo de parada r_n está acotado:

$$\forall \omega \in \Omega, r_n(\omega) \leq n.$$

Podemos por tanto aplicar el Teorema de Muestreo Opcional (cf. ASH, Teorema 7.7.3, p.304) para concluir que:

$$\sum_{i=1}^{r_n} I_d(s, i)(g(s^{(K_i^d)}) - v_d(s)), \quad (4.36)$$

para $n \geq 1$, es una martingala de media cero para la filtración:

$$\{\mathcal{F}_{r_n}\}_{n \geq 1}.$$

La σ -álgebra \mathcal{F}_{r_n} contiene, para una muestra de tamaño n , toda la información de los tramos completos contenidos en esa muestra, pero no abarca las

observaciones desde la $(K_{r_n}^d + 1)$ -ésima hasta la n -ésima. Por comodidad en la notación, vamos a definir:

$$r_0 = 0.$$

Para cada $n \geq 1$ y cada i tal que $1 \leq i \leq n$, se define la diferencia de martingalas:

$$X_{ni} = \frac{1}{\sqrt{n}} \sum_{j=r_{i-1}+1}^{j=r_i} I_d(s, j)(g(s^{(K_j^d)}) - v_d(s)), \quad (4.37)$$

obtenida a partir de la martingala (4.36). Nótese que al tomar una observación más en la muestra, a lo sumo se puede completar un tramo de más. Resulta entonces que $r_i - r_{i-1}$ sólo puede valer cero o uno, con lo cual el sumatorio de (4.37) o bien es igual a cero o bien consta de un único término: el correspondiente a r_i , deduciéndose de esto que:

$$|X_{ni}| \leq \frac{1}{\sqrt{n}} \max_{s' \in \bar{S}} g(s'). \quad (4.38)$$

También se cumple que:

$$\sum_{i=1}^n \sum_{j=r_{i-1}+1}^{j=r_i} I_d(s, j)(g(s^{(K_j^d)}) - v_d(s)) = \sum_{i=1}^{r_n} I_d(s, i)(g(s^{(K_i^d)}) - v_d(s)), \quad (4.39)$$

y $n - r_n$ de los n sumandos del miembro izquierdo son nulos— los correspondientes a muestras en las que no se completa tramo—, y los r_n sumandos que no son necesariamente nulos son los mismos que los r_n sumandos del miembro derecho.

Vamos a comprobar que la sucesión (X_{ni}) verifica el Teorema 4.13.

Se toma $\mathcal{F}_{ni} = \mathcal{F}_{r_i}$, cumpliéndose así (4.32), cf. p.190. Las propiedades (4.33) y (4.34) son consecuencias inmediatas de la acotación (4.38). Hay que estudiar ahora la convergencia en probabilidad de:

$$\frac{1}{n} \sum_{i=1}^n \left(\sum_{j=r_{i-1}+1}^{j=r_i} I_d(s, j)(g(s^{(K_j^d)}) - v_d(s)) \right)^2,$$

que, por lo observado en (4.39), es igual a:

$$\frac{1}{n} \sum_{i=1}^{r_n} I_d(s, i)(g(s^{(K_i^d)}) - v_d(s))^2,$$

y se puede escribir:

$$\sum_{d(s')=0} (g(s') - v_d(s))^2 \frac{1}{n} \sum_{i=1}^{r_n} I_{\{s^{(K_i^d)}=s'\}} I_d(s, i),$$

que es igual a:

$$\sum_{d(s')=0} (g(s') - v_d(s))^2 \frac{r_n}{n} \cdot \frac{1}{r_n} \sum_{i=1}^{r_n} I_{\{s^{(K_i^d)}=s'\}} I_d(s, i). \quad (4.40)$$

Dado que r_n es el número de tramos que se han observado en la muestra de tamaño n , resulta que r_n es el número de veces que la muestra de tamaño n ha visitado estados de $d^{-1}(0)$, es decir:

$$r_n = \sum_{i=1}^n I_{\{d(s^{(i)})=0\}},$$

y por tanto:

$$\frac{r_n}{n}$$

converge casi seguramente a la probabilidad estacionaria de la cadena con matriz de transición \tilde{P} del conjunto $d^{-1}(0)$, que denotaremos por:

$$\tilde{P}\{d^{-1}(0)\},$$

siendo pues el límite de r_n/n positivo. Por otra parte, cuando $n \rightarrow +\infty$:

$$\frac{1}{r_n} \sum_{i=1}^{r_n} I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) \xrightarrow{\text{cs}} \pi_{ss'}^d \cdot \mathbf{E}_{\mathbf{p}} q_d,$$

en virtud de la Proposición 4.4. Nótese que se ha podido emplear la técnica del Teorema 4.12 puesto que se trata del límite casi seguro de una sub-sucesión aleatoria de una sucesión aleatoria. Esta última convergencia también se hubiese podido probar observando que:

$$\frac{1}{r_n} \sum_{i=1}^{r_n} I_{\{s^{(K_i^d)}=s'\}} I_d(s, i) = \hat{p}_{(s',1)}^{(r_n)},$$

en la notación del Teorema 4.10.

En conclusión, se verifica la condición (4.35), cf. p.190, siendo la convergencia casi segura a la variable degenerada: ξ^2 , igual a:

$$\tilde{P}\{d^{-1}(0)\}(\mathbf{E}_{\mathbf{p}q_d}) \sum_{d(s')=0} \pi_{ss'}^d (g(s') - v_d(s))^2 = \tilde{P}\{d^{-1}(0)\}(\mathbf{E}_{\mathbf{p}q_d}) V_{d,s}(g),$$

y por tanto:

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{r_n} I_d(s, i) (g(s^{(K_i^d)}) - v_d(s)) \xrightarrow{d} \mathcal{N}(0, \cdot),$$

cuando $n \rightarrow +\infty$.

En la demostración del Teorema 4.8 se probó que:

$$\sqrt{r}(\hat{v}_d^{(r)}(s) - v_d(s)) = \frac{r}{N_d(s, r)} \cdot \frac{1}{\sqrt{r}} \sum_{i=1}^r I_d(s, i) (g(s^{(K_i^d)}) - v_d(s)),$$

luego se tiene que:

$$\sqrt{r_n}(\tilde{v}_d^{(n)}(s) - v_d(s)) = \frac{r_n}{N_d(s, r_n)} \cdot \frac{1}{\sqrt{r_n}} \sum_{i=1}^{r_n} I_d(s, i) (g(s^{(K_i^d)}) - v_d(s)),$$

y por tanto:

$$\sqrt{n}(\tilde{v}_d^{(n)}(s) - v_d(s)) = \frac{r_n}{N_d(s, r_n)} \cdot \frac{n}{r_n} \cdot \frac{1}{\sqrt{n}} \sum_{i=1}^{r_n} I_d(s, i) (g(s^{(K_i^d)}) - v_d(s)).$$

Se deduce de la Proposición 4.3 y de lo dicho anteriormente que:

$$\frac{r_n}{N_d(s, r_n)} \cdot \frac{n}{r_n} \xrightarrow{cs} (\tilde{P}\{d^{-1}(0)\} \mathbf{E}_{\mathbf{p}q_d})^{-1},$$

cuando $n \rightarrow +\infty$. Del límite en distribución demostrado para la diferencia de martingalas se concluye que:

$$\left(\sqrt{n}(\tilde{v}_d^{(n)}(s) - v_d(s)) \right)_{n \geq 1}$$

converge en ley a una variable aleatoria normal de media cero y de varianza:

$$(\tilde{P}\{d^{-1}(0)\} \mathbf{E}_{\mathbf{p}q_d})^{-1} V_{d,s}(g).$$

Q.E.D.

Hay que subrayar las dificultades que ha entrañado esta demostración. Hasta ahora, el problema estaba planteado en términos del tamaño r , es decir, del número de tramos completos observados. Al fijar el tamaño muestral y no el número de tramos, ha habido que hacer una nueva demostración de la convergencia en ley. Se ha hecho utilizando una sucesión de tiempos de parada que transformaba el problema de tamaño r — número de tramos— en el problema de tamaño n — número de observaciones—. La propiedad de martingala que se tenía en el problema de tamaño r se ha podido conservar para el problema de tamaño n gracias al Teorema de Muestreo Opcional.

Resumamos desde un punto de vista conceptual lo que se ha hecho hasta el momento. Se ha planteado un problema de estimación para un tamaño muestral n . Para adaptar el modelo a los estimadores que se han considerado, se ha cambiado de espacio muestral, pasando a un tamaño r . En ese nuevo espacio muestral se han construido martingalas y cadenas de Markov que han resuelto el problema propuesto. Para volver al problema de tamaño n ha habido que realizar la operación inversa para de nuevo cambiar de espacio muestral. Los cambios de espacio muestral se hacen a través de sucesiones de tiempos de parada que son: $(K_i^d)_{i \geq 1}$ para el cambio directo, y $(r_n)_{n \geq 1}$ para el cambio recíproco. Así pues, estas dos sucesiones de tiempos de parada que permiten hacer cambios inversos de espacios muestrales, están formadas, en cierto modo, por tiempos de parada inversos unos de otros. Esta propiedad de ser los tiempos de parada inversos viene reflejada en la relación:

$$r_{K_i^d} = i,$$

con lo cual K_i^d es inverso por la derecha de r_n , aunque no es inverso por la izquierda, pudiéndose cumplir que:

$$\mathcal{F}_{r_n} \subset \mathcal{A}_n.$$

En el espacio muestral relativo al problema de tamaño r habíamos construido una martingala. Al cambiar de espacio muestral, aplicando el tiempo de parada inverso, el Teorema de Muestreo Opcional nos garantiza que en el otro problema volvemos a obtener otra martingala, es decir: los dos espacios muestrales son transformados el uno del otro a través de los tiempos de parada inversos, y las martingalas que se consideran en los dos espacios muestrales son transformadas la una de la otra a través de los tiempos de parada inversos, conservándose la propiedad de martingala por el Teorema de Muestreo Opcional.

Estas últimas observaciones son verdaderamente cruciales y muestran la dificultad que tiene el cambio de espacio muestral y lo interesante de la aplicación del Teorema de Muestreo Opcional. En efecto, hemos pasado de tener varios teoremas centrales del límite en espacios muestrales distintos a tener varios teoremas centrales del límite en un espacio muestral común.

Recordemos que sólo necesitamos hacer estimaciones del valor $v_d(s)$ cuando $d(s) = 1$, ya que si $d(s) = 0$ entonces el valor $v_d(s) = g(s)$ es conocido. Por uniformidad en la notación, definiremos:

$$\tilde{v}_d^{(n)}(s) = g(s),$$

cuando $d(s) = 0$.

Para una muestra de tamaño n , disponemos pues, para cada regla de decisión $d \in D$ y cada estado $s \in \bar{S}$, de un estimador $\tilde{v}_d^{(n)}(s)$ que es consistente casi seguramente y asintóticamente normal, lo que nos va a permitir estimar una regla óptima y el valor del problema de decisión.

4.3 Estimación de una regla óptima

4.3.1 Estudio de un ejemplo

En esta sección vamos a tratar el problema de la estimación de una regla óptima, una vez resuelta la estimación de cada una de las reglas de decisión. Pero antes de proponer la estimación de la regla óptima, es interesante hacer ciertas observaciones. En la sección anterior se ha presentado un método directo para la estimación del valor de las reglas de decisión. En el estudio realizado en el Capítulo 3 se utilizaba el procedimiento de “resolver y estimar”, en el cual se estimaba la matriz P por máxima verosimilitud, obteniendo después el estimador de máxima verosimilitud del valor de las reglas de decisión. Es conveniente, en este momento, mostrar mediante un ejemplo sencillo cómo estos dos métodos dan lugar a estimadores distintos.

EJEMPLO: Consideremos un problema de parada óptima con conjunto de estados:

$$\bar{S} = \{s_1, s_2, s_3, s_4\},$$

siendo los estados de transición: s_1 , s_2 y s_3 , y el estado absorbente: s_4 . Los pagos en cada estado son:

$$g(s_1) = 1, \quad g(s_2) = 2, \quad g(s_3) = 5 \quad \text{y} \quad g(s_4) = 0.$$

Supongamos que se observa una muestra, de tamaño $n = 5$, que es la siguiente:

$$s_2, s_3, s_1, s_2, s_4.$$

Cualquier regla de decisión d verifica: $d(s_4) = 0$, por lo que para este problema existen ocho reglas de decisión distintas, pero puesto que el pago máximo se consigue en el estado s_3 , es suficiente considerar únicamente aquellas reglas que paran en s_3 , es decir, tales que:

$$d(s_3) = 0,$$

y no tendremos en cuenta más que cuatro de las ocho reglas de decisión, que vendrán pues caracterizadas por las acciones que toman en s_1 y s_2 , estando las decisiones en los demás estados prefijadas. Estas cuatro reglas que consideraremos son:

- d_1 , que cumple: $d_1(s_1) = 0$ y $d_1(s_2) = 0$;
- d_2 , que cumple: $d_2(s_1) = 1$ y $d_2(s_2) = 1$;
- d_3 , que cumple: $d_3(s_1) = 1$ y $d_3(s_2) = 0$;
- d_4 , que cumple: $d_4(s_1) = 0$ y $d_4(s_2) = 1$.

Utilizando el método de la máxima verosimilitud se tiene que:

$$\hat{P}^{(5)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0,5 \\ 1 & 0 & 0 \end{pmatrix}.$$

Puesto que de las cuatro reglas d_1 , d_2 , d_3 y d_4 , sabemos que paran en s_3 y s_4 , resulta que sus valores en esos dos estados son 5 y 0, respectivamente. Por tanto, sólo estimaremos sus valores en s_1 y s_2 , denotando ese vector bi-dimensional de estimaciones por \hat{u}_{d_i} . La estimación máximo-verosímil es:

$$\hat{u}_{d_1} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \hat{u}_{d_2} = \begin{pmatrix} 2,5 \\ 2,5 \end{pmatrix}, \quad \hat{u}_{d_3} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \text{y} \quad \hat{u}_{d_4} = \begin{pmatrix} 1 \\ 2,5 \end{pmatrix}.$$

Con el método directo que ha sido estudiado, teniendo en cuenta el número de tramos observados para cada regla de decisión, las estimaciones que se obtienen, que escribiremos \hat{v}_{d_i} , son:

$$\hat{v}_{d_1} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \hat{v}_{d_2} = \begin{pmatrix} 0 \\ 2, 5 \end{pmatrix}, \quad \hat{v}_{d_3} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \text{y} \quad \hat{v}_{d_4} = \begin{pmatrix} 1 \\ 2, 5 \end{pmatrix}.$$

Nótese que la estimación del valor de la regla d_2 varía según se emplee un método u otro, con lo cual la nueva técnica introducida no es el método anterior presentado bajo otra apariencia, sino que es efectivamente un método distinto.

A la hora de estimar la regla óptima con el método de la máxima verosimilitud, se tomaba como estimación aquella regla cuya estimación \hat{u}_{d_i} dominase— coordenada a coordenada— a todas las demás estimaciones. En este ejemplo:

$$\hat{u}_{d_2} \geq \hat{u}_{d_i}, \text{ para todo } i,$$

y se tomaría d_2 como estimación de la regla óptima. En cambio, con el método directo, no existe ninguna estimación \hat{v}_{d_i} que supere a las demás, y por tanto deberemos modificar la definición de la estimación de la regla óptima.

Utilizando el método de la máxima verosimilitud siempre se garantiza que existe una estimación del valor de las reglas que domina a todas las demás. Esto es porque los estimadores del valor de las reglas son, de hecho, valores verdaderos de esas reglas de decisión: son el valor de esas reglas cuando la matriz de transición es $\hat{P}^{(5)}$. Es decir, lo que hace el método de máxima verosimilitud no es, en realidad, más que resolver un problema de parada óptima: el que tiene matriz $\hat{P}^{(5)}$, y la solución de ese problema hereda todas las propiedades de las soluciones de problemas de parada óptima, entre otras, la existencia de una regla de decisión que mejora a todas las demás, que era la que en nuestro contexto tomábamos como estimación de la regla óptima.

No ocurre lo mismo con las estimaciones que propone el método directo ya que no existe ninguna matriz de transición Q tal que al resolver el problema con matriz de transición Q obtengamos los valores \hat{v}_{d_i} para las reglas de decisión d_i .

Es preciso insistir, porque aquí radica la diferencia esencial entre los métodos de máxima verosimilitud y directo: los estimadores por el primer método son los valores de un cierto problema de decisión concreto, mientras que los estimadores por el segundo método no se corresponden con los valores

de ningún problema de parada óptima, y en ese sentido, pierden buenas propiedades con respecto al método de máxima verosimilitud, lo que por otra parte era de esperar, al tratarse de un método de estimación menos exhaustivo.

4.3.2 Convergencia de la estimación

Procedamos ahora a la estimación de una regla óptima. Dada una regla de decisión $d \in D$ y un estado $s \in \bar{S}$, se define:

$$\Delta_d(s) = \max_{\delta \in D} v_\delta(s) - v_d(s),$$

y para una muestra de tamaño n , se define la variable aleatoria:

$$\Delta_d^{(n)}(s) = \max_{\delta \in D} \tilde{v}_\delta^{(n)}(s) - \tilde{v}_d^{(n)}(s),$$

que es una cantidad que sólo depende de las observaciones muestrales y no de datos desconocidos. Definimos también:

$$\Delta_d = \sum_{s \in \bar{S}} \Delta_d(s) \quad \text{y} \quad \Delta_d^{(n)} = \sum_{s \in \bar{S}} \Delta_d^{(n)}(s).$$

Se verifica el siguiente resultado.

LEMA 4.15: *Para toda regla $d \in D$ y todo estado $s \in \bar{S}$ las cantidades:*

$$\Delta_d(s), \quad \Delta_d, \quad \Delta_d^{(n)}(s) \quad \text{y} \quad \Delta_d^{(n)}$$

son mayores o iguales que cero.

Una regla $d \in D$ es óptima si y sólo si $\Delta_d = 0$.

DEMOSTRACIÓN: La primera parte del enunciado es una consecuencia directa de las definiciones.

Una regla $d \in D$ verifica $\Delta_d = 0$ si y sólo si para cada $s \in \bar{S}$ se cumple:

$$\Delta_d(s) = 0,$$

o lo que es equivalente:

$$\forall s \in \bar{S}, \quad v_d(s) = \max_{\delta \in D} v_\delta(s),$$

que es lo mismo que decir que d es óptima.

Q.E.D.

Obsérvese que otras definiciones de Δ_d que harían igualmente válido el Lema 4.15 son:

$$\Delta_d = \sum_{s \in \bar{S}} (\Delta_d(s))^2 \quad \text{y} \quad \Delta_d = \max_{s \in \bar{S}} \Delta_d(s).$$

Necesitaremos también el siguiente resultado previo.

LEMA 4.16: *Se verifican los siguientes límites:*

$$\max_{d \in D} |\Delta_d^{(n)} - \Delta_d| \xrightarrow{\text{cs}} 0 \quad \text{y} \quad \min_{d \in D} \Delta_d^{(n)} \xrightarrow{\text{cs}} 0,$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: Es una consecuencia inmediata de la consistencia casi segura de los estimadores $\tilde{v}_d^{(n)}(s)$ y de la propiedad:

$$\min_{d \in D} \Delta_d = 0.$$

Q.E.D.

Dada una muestra de tamaño n , se define la estimación de una regla óptima como una regla $d_n \in D$ que verifique:

$$\Delta_{d_n}^{(n)} = \min_{d \in D} \Delta_d^{(n)}, \quad (4.41)$$

cumpliendo esta definición el siguiente resultado de consistencia.

TEOREMA 4.17: *Para casi toda trayectoria $\omega \in \Omega$ existe un tamaño muestral a partir del cual la regla d_n es óptima para el problema de decisión.*

DEMOSTRACIÓN: Apoyándonos en el Lema 4.16, para todo ω en un conjunto de probabilidad uno, se cumple:

$$\max_{d \in D} |\Delta_d^{(n)} - \Delta_d| \longrightarrow 0,$$

y también:

$$\min_{d \in D} \Delta_d^{(n)} \longrightarrow 0,$$

cuando $n \rightarrow +\infty$.

Sea D^* el conjunto de reglas óptimas del problema de decisión, y para evitar situaciones triviales supondremos que $D^* \subset D$. Sea:

$$\varepsilon = \min_{d \in D - D^*} \Delta_d. \quad (4.42)$$

Del Lema 4.15 se deduce que $\varepsilon > 0$.

Existe un tamaño muestral N tal que si $n \geq N$ entonces:

$$\max_{d \in D} |\Delta_d^{(n)} - \Delta_d| < \varepsilon/2 \quad \text{y} \quad \min_{d \in D} \Delta_d^{(n)} < \varepsilon/2,$$

luego para una regla d_n que verifique (4.41), se tiene:

$$|\Delta_{d_n}^{(n)} - \Delta_{d_n}| < \varepsilon/2 \quad \text{y} \quad \Delta_{d_n}^{(n)} < \varepsilon/2,$$

de donde:

$$\Delta_{d_n} < \varepsilon.$$

De la definición de ε dada en (4.42) se deduce que $d_n \in D^*$, siendo pues, para tamaños muestrales $n \geq N$, la regla d_n óptima. Q.E.D.

En el ejemplo que se planteó al principio de esta sección, la estimación de una regla óptima es la regla d_3 , que no es la misma regla que se obtendría por el método de máxima verosimilitud, que sería d_2 .

4.4 Estimación del valor del problema

4.4.1 Definición de los estimadores

Una vez tratadas las estimaciones de una regla óptima y del valor de las reglas de decisión, abordamos el objetivo principal de este estudio, que es la estimación del valor del problema de decisión: v^* .

El valor del problema verifica:

$$v^*(s) = v_{d^*}(s), \quad (4.43)$$

para una regla $d^* \in D^*$, y también:

$$v^*(s) = \max_{d \in D} v_d(s), \quad (4.44)$$

para todo $s \in \bar{S}$. Cada una de las ecuaciones (4.43) y (4.44) sugiere una forma de estimar el valor del problema.

Para una muestra de tamaño n , la estimación de la regla óptima es d_n . Una primera posibilidad para estimar el valor del problema, inspirada en la expresión (4.43), es tomar como estimador de v^* el valor estimado de la regla d_n , es decir:

$$\hat{v}^{(n)}(s) = \tilde{v}_{d_n}^{(n)}(s), \quad (4.45)$$

para cada $s \in \bar{S}$. La segunda opción, basada en la igualdad (4.44), es la estimación:

$$\hat{w}^{(n)}(s) = \max_{d \in D} \tilde{v}_d^{(n)}(s), \quad (4.46)$$

siendo $s \in \bar{S}$. Estos dos estimadores no son necesariamente iguales. Para el ejemplo de la sección anterior, se tiene:

$$\hat{v}^{(5)}(s_2) = 2 \quad \text{y} \quad \hat{w}^{(5)}(s_2) = 2, 5.$$

Hay que hacer notar, además, que ninguno de estos dos estimadores coincide forzosamente con el estimador de máxima verosimilitud, tratándose pues de dos nuevos estimadores. En el ejemplo considerado, se tiene que el estimador máximo-verosímil del valor del problema en s_1 es 2, 5 mientras que:

$$\hat{v}^{(5)}(s_1) = \hat{w}^{(5)}(s_1) = 2.$$

4.4.2 Convergencia de la estimación

Los dos estimadores propuestos son consistentes casi seguramente.

TEOREMA 4.18: *Para cada estado $s \in \bar{S}$ se cumple:*

$$\hat{v}^{(n)}(s) \xrightarrow{\text{cs}} v^*(s) \quad \text{y} \quad \hat{w}^{(n)}(s) \xrightarrow{\text{cs}} v^*(s),$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: Probemos en primer lugar que el estimador (4.46) es consistente. Puesto que:

$$\hat{w}^{(n)}(s) = \max_{d \in D} \tilde{v}_d^{(n)}(s),$$

y, con probabilidad uno, cuando $n \rightarrow +\infty$ se tiene que:

$$\tilde{v}_d^{(n)}(s) \longrightarrow v_d(s),$$

resulta que:

$$\hat{w}^{(n)}(s) \longrightarrow \max_{d \in D} v_d(s) = v^*(s).$$

Para probar la convergencia del estimador definido en (4.45), recordemos que:

$$\Delta_d^{(n)}(s) = \max_{\delta \in D} \tilde{v}_\delta^{(n)}(s) - \tilde{v}_d^{(n)}(s) = \hat{w}^{(n)}(s) - \tilde{v}_d^{(n)}(s),$$

y por tanto:

$$\Delta_{d_n}^{(n)}(s) = \hat{w}^{(n)}(s) - \hat{v}^{(n)}(s),$$

que es, según el Lema 4.15, una cantidad mayor o igual que cero. De aplicar el segundo resultado del Lema 4.16 resulta que:

$$\min_{d \in D} \Delta_d^{(n)} = \Delta_{d_n}^{(n)} = \sum_{s \in \bar{S}} (\hat{w}^{(n)}(s) - \hat{v}^{(n)}(s)) \xrightarrow{\text{cs}} 0,$$

cuando $n \rightarrow +\infty$, de donde, para todo $s \in \bar{S}$: $\hat{w}^{(n)}(s) - \hat{v}^{(n)}(s) \xrightarrow{\text{cs}} 0$, concluyéndose el resultado del enunciado. Q.E.D.

Llegado el momento de estudiar la distribución límite del estimador del valor del problema:

$$(\hat{v}^{(n)}(s))_{n \geq 1},$$

surgen las mismas dificultades que con el método de máxima verosimilitud, y se garantiza que este estimador es asintóticamente normal solamente cuando la regla óptima es única.

El estimador $(\hat{w}^{(n)}(s))_{n \geq 1}$ no verifica un teorema central del límite y es porque su definición no está asociada a ninguna regla de decisión, puesto que está basada en máximos de cantidades. En cambio, el estimador $(\hat{v}^{(n)}(s))_{n \geq 1}$ sí está asociado a una regla de decisión: la regla d_n , por lo que se le podrán aplicar los resultados límite del estimador del valor de una regla.

TEOREMA 4.19: *Si la regla óptima es única, entonces para cada $s \in \bar{S}$:*

$$\sqrt{n}(\hat{v}^{(n)}(s) - v^*(s)) \xrightarrow{d} \mathcal{N}(0, \cdot),$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: Sea $d^* \in D$ la única regla óptima. Del Teorema 4.17 se deduce que para casi todo $\omega \in \Omega$ existe un tamaño muestral a partir del cual:

$$d_n = d^*,$$

y por tanto:

$$\hat{v}^{(n)}(s) = \tilde{v}_{d_n}^{(n)}(s) = \tilde{v}_{d^*}^{(n)}(s),$$

para n suficientemente grande. En conclusión:

$$\sqrt{n}(\hat{v}^{(n)}(s) - \tilde{v}_{d^*}^{(n)}(s)) \xrightarrow{\text{cs}} 0, \quad (4.47)$$

cuando $n \rightarrow +\infty$. Por el Teorema 4.14, se tiene que:

$$\sqrt{n}(\tilde{v}_{d^*}^{(n)}(s) - v_{d^*}(s)) = \sqrt{n}(\tilde{v}_{d^*}^{(n)}(s) - v^*(s)) \xrightarrow{d} \mathcal{N}(0, \cdot), \quad (4.48)$$

cuando $n \rightarrow +\infty$. De los límites (4.47) y (4.48) se concluye que:

$$\sqrt{n}(\hat{v}^{(n)}(s) - v^*(s))$$

es asintóticamente normal.

La distribución límite es una normal de media cero y de varianza igual a cero si $d^*(s) = 0$, e igual a:

$$(\tilde{P}\{(d^*)^{-1}(0)\}\mathbf{E}_{\mathbf{p}^*}q_{d^*})^{-1}V_{d^*,s}(g)$$

si $d^*(s) = 1$.

Q.E.D.

Obsérvese que este teorema no es una aplicación del Teorema 4.14 tomando como regla d la regla d^* . En efecto, en el Teorema 4.14 la regla de decisión estaba fijada de antemano, pero en este último resultado la regla d^* no es conocida sino que se va estimando etapa a etapa.

En el caso de no ser la regla óptima única, no se puede establecer un teorema central del límite ya que la estimación $(d_n)_{n \geq 1}$ no identifica necesariamente una única regla óptima con probabilidad uno, pudiendo alternar la sucesión $(d_n)_{n \geq 1}$ sus valores entre varias de las reglas óptimas. Es por esta razón que en el análisis del problema de parada óptima por máxima verosimilitud se introdujeron métodos de perturbación, que vuelven pues a ser necesarios en este nuevo estudio del problema de parada óptima.

4.5 Métodos de perturbación

4.5.1 Introducción

El objetivo de un método de perturbación es introducir una pequeña alteración en los parámetros del problema para conseguir que el problema

perturbado tenga una única regla óptima. Si la perturbación es lo bastante pequeña, es de esperar que esa única regla óptima para el problema modificado lo sea también para el problema original. De poderse lograr esto, gracias al problema perturbado conseguiríamos estimar de forma consistente su única regla óptima, sirviéndonos de ella para luego estimar el valor del problema de decisión original, obteniendo un nuevo estimador que sería consistente casi seguramente y asintóticamente normal, salvando así la dificultad que plantea la multiplicidad de las reglas óptimas.

Trataremos tres métodos de perturbación distintos. Los dos primeros— perturbación del vector de pagos g y perturbación de la matriz P — ya fueron estudiados cuando se analizó el método de la máxima verosimilitud, y son igualmente aplicables en el contexto de los estimadores que estamos considerando. El tercer método consiste en alterar el vector g , aunque de forma distinta a la ya tratada. Es igualmente aplicable para la técnica de máxima verosimilitud y parece ser el más operativo de los tres métodos presentados.

Recordamos la definición de los siguientes conjuntos:

$$S_c = \{s \in S : g(s) < \sum_{s' \in S} p(s'|s)v^*(s')\},$$

y S_c es el conjunto de estados de S en los que la decisión óptima es la acción 1, es decir, continuar;

$$S_p = \{s \in S : g(s) > \sum_{s' \in S} p(s'|s)v^*(s')\},$$

y S_p es el conjunto de estados de S en los que la decisión óptima es la acción 0, es decir, parar;

$$S_i = \{s \in S : g(s) = \sum_{s' \in S} p(s'|s)v^*(s')\},$$

y S_i es el conjunto de estados de S en los que tanto parar como continuar es óptimo, es decir, son los estados que llamaremos de indiferencia, por ser indiferente la acción que se tome. Estos conjuntos dependen de los parámetros g y P , y los escribiremos:

$$S_c(g, P), \quad S_p(g, P) \quad \text{y} \quad S_i(g, P).$$

Definimos también:

$$S_{ci}(g, P) = S_c(g, P) \cup S_i(g, P),$$

que son los estados en los que la acción óptima es continuar o indiferente. La regla óptima de un problema (g, P) será única si y sólo si:

$$S_i(g, P) = \emptyset.$$

Hay que hacer notar que estamos excluyendo de todas estas consideraciones el estado s_{m+1} ya que, en ese estado, la única acción posible es parar recibiendo un pago igual a cero. Hagamos el convenio, que no supone ninguna pérdida de generalidad, de que los estados de $S_{ci}(g, P)$ son los k primeros, es decir:

$$S_{ci}(g, P) = \{s_1, \dots, s_k\} \quad y \quad S_p(g, P) = \{s_{k+1}, \dots, s_m\}. \quad (4.49)$$

4.5.2 Primer método de perturbación

Suponemos que:

$$P \in T_m \times \dots \times T_m.$$

Se define el vector columna $\delta(g) \in \mathbb{R}^m$ como:

$$\begin{aligned} \delta(g)(s) &= 0, & \text{si } s \in S_{ci}(g, P), \\ \delta(g)(s) &= 1, & \text{si } s \in S_p(g, P), \end{aligned}$$

es decir, según el convenio hecho en (4.49):

$$\delta(g) = (0 \overset{k}{\dots} 0 \ 1 \overset{m-k}{\dots} 1)'$$

Se considerará la perturbación del vector g :

$$g_\varepsilon = g + \varepsilon \delta(g),$$

siendo $\varepsilon > 0$. Enunciamos la siguiente proposición, que no es más que la Proposición 3.15, cf. p.144, del método de máxima verosimilitud.

PROPOSICIÓN 4.20: *Con la notación anterior, existe $\eta > 0$ tal que siempre que ε cumpla: $0 < \varepsilon < \eta$, se tiene que:*

$$S_i(g_\varepsilon, P) = \emptyset,$$

verificándose:

$$S_c(g_\varepsilon, P) = S_{ci}(g, P) \quad y \quad S_p(g_\varepsilon, P) = S_p(g, P).$$

Si $\varepsilon > 0$ es como en la Proposición 4.20, el problema (g_ε, P) tiene una única regla óptima d^{**} dada por:

$$\begin{aligned} d^{**}(s) &= 1, & \text{si } s \in S_{ci}(g, P), \\ d^{**}(s) &= 0, & \text{si } s \in S_p(g, P), \end{aligned} \quad (4.50)$$

que también es óptima para el problema original (g, P) .

Tomada una muestra de tamaño n del problema con matriz P , y considerando g_ε como vector de pagos, sea $d_{n,\varepsilon}$ la estimación de la regla óptima del problema (g_ε, P) . Se define un nuevo estimador del valor del problema original (g, P) : $\hat{v}_\varepsilon^{(n)}$, mediante:

$$\hat{v}_\varepsilon^{(n)}(s) = \tilde{v}_{d_{n,\varepsilon}}^{(n)}(s),$$

para cada estado $s \in \bar{S}$. El estimador $\hat{v}_\varepsilon^{(n)}$ es el valor estimado en el problema (g, P) de la regla $d_{n,\varepsilon}$. Este nuevo estimador sí verifica un teorema central del límite.

TEOREMA 4.21: *Tomando ε como en la Proposición 4.20, se cumple, para cada $s \in \bar{S}$:*

$$\hat{v}_\varepsilon^{(n)}(s) \xrightarrow{\text{cs}} v^*(s) \quad \text{y} \quad \sqrt{n}(\hat{v}_\varepsilon^{(n)}(s) - v^*(s)) \xrightarrow{\text{d}} \mathcal{N}(0, \cdot),$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: Según el Teorema 4.17, con probabilidad uno, a partir de un cierto tamaño muestral:

$$d_{n,\varepsilon} = d^{**},$$

es decir que:

$$\hat{v}_\varepsilon^{(n)}(s) = \tilde{v}_{d^{**}}^{(n)}(s).$$

Puesto que la regla d^{**} es óptima para el problema (g, P) , de los teoremas 4.12 y 4.14 se deduce que:

$$\tilde{v}_{d^{**}}^{(n)}(s) \xrightarrow{\text{cs}} v^*(s) \quad \text{y} \quad \sqrt{n}(\tilde{v}_{d^{**}}^{(n)}(s) - v^*(s)) \xrightarrow{\text{d}} \mathcal{N}(0, \cdot),$$

y se concluye ya el resultado del enunciado, razonando como en el Teorema 4.19. La varianza de la distribución límite es la misma que la obtenida en el Teorema 4.19, pero para la regla d^{**} . Q.E.D.

Obsérvese que en este teorema se ha adaptado, de alguna manera, el estimador $\hat{v}^{(n)}$ al problema perturbado. No se puede hacer lo mismo con el estimador $\hat{w}^{(n)}$ al no estar éste basado en ninguna regla concreta, sino en máximos de cantidades.

Este método de perturbación resulta muy cómodo porque las alteraciones no afectan al modelo de probabilidad: la matriz P no se perturba, y sólo afectan al vector de pagos. El inconveniente que tiene esta técnica es que el vector de perturbación $\delta(g)$ depende de datos que son en realidad desconocidos, como son los estados en los que la decisión óptima es parar.

La segunda técnica de perturbación, que se propone a continuación, también ha sido ya estudiada para la resolución del problema de parada óptima por el método de la máxima verosimilitud.

4.5.3 Segundo método de perturbación

Volvemos a suponer que:

$$P \in T_m \times \cdots \times T_m.$$

En este segundo método, no se modifica el vector de pagos g y se altera la matriz P , que es la formada por las probabilidades de transición de la cadena de Markov entre los estados de S .

Manteniendo el convenio (4.49), se considera una matriz $\delta(P)$ de tamaño $m \times m$:

$$\delta(P) = \begin{pmatrix} \delta_1 & \delta_2 \\ \delta_3 & \delta_4 \end{pmatrix},$$

siendo la matriz $\delta_1 = \delta_1(P)$ de tamaño $k \times k$, que corresponde a los estados de $S_{ci}(g, P)$. Se estudiará la perturbación:

$$P_\varepsilon = P + \varepsilon\delta(P).$$

Tomamos valores de $\varepsilon \geq 0$ lo bastante pequeños para que:

$$P_\varepsilon \in T_m \times \cdots \times T_m.$$

Enunciamos el resultado dado en la Proposición 3.17, cf. p.149.

PROPOSICIÓN 4.22: Si $\delta_1 > \mathbf{0}$ y $\delta(P) \geq \mathbf{0}$ entonces existe $\eta > 0$ tal que siempre que ε cumpla: $0 < \varepsilon < \eta$, se tiene que:

$$S_i(g, P_\varepsilon) = \emptyset,$$

verificándose:

$$S_c(g, P_\varepsilon) = S_{ci}(g, P) \quad \text{y} \quad S_p(g, P_\varepsilon) = S_p(g, P).$$

La regla d^{**} definida en (4.50) es óptima tanto para el problema perturbado (g, P_ε) como para el problema original (g, P) . Tomada una muestra de tamaño n de la cadena de Markov con matriz alterada P_ε , sea $d_{n,\varepsilon}$ la estimación de la regla óptima. El nuevo estimador del valor del problema que se propone es:

$$\hat{v}_\varepsilon^{(n)}(s) = \tilde{v}_{d_{n,\varepsilon}}^{(n)}(s),$$

para cada $s \in \bar{S}$, que es el valor estimado, para la muestra de la matriz P , de la regla de decisión $d_{n,\varepsilon}$. Este estimador basado en la perturbación cumple el siguiente resultado.

TEOREMA 4.23: Tomando ε como en la Proposición 4.22, se cumple, para cada $s \in \bar{S}$:

$$\hat{v}_\varepsilon^{(n)}(s) \xrightarrow{\text{cs}} v^*(s) \quad \text{y} \quad \sqrt{n}(\hat{v}_\varepsilon^{(n)}(s) - v^*(s)) \xrightarrow{\text{d}} \mathcal{N}(0, \cdot),$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: Es la misma que la del Teorema 4.21.

Q.E.D.

A diferencia del primer método de perturbación, en éste la perturbación no depende de datos desconocidos ya que hay una gran libertad para escoger la matriz $\delta(P)$, pudiéndose tomar como $\delta(P)$ una matriz con todos sus términos iguales a uno.

El inconveniente de esta técnica es otro. Obsérvese que para llevar a cabo la nueva estimación son necesarios dos muestreos diferentes: el de la matriz P y el de la matriz P_ε , con lo cual este método de perturbación resulta costoso. Cuando se analizó el problema de parada óptima desde la perspectiva de la estimación máximo-verosímil no surgió este inconveniente. Esto es porque

la matriz P se estimaba, mediante la matriz $\hat{P}^{(n)}$, y no era necesario hacer un muestreo de la matriz P_ε , puesto que la podíamos estimar, gracias a la estimación de la matriz original, de la forma:

$$\hat{P}^{(n)} + \varepsilon\delta(P),$$

estableciéndose el resultado análogo al Teorema 4.23 con un único muestreo. En este nuevo estudio, al no estimar la matriz de transición P — es precisamente lo que se pretende evitar—, la única forma de obtener información del problema (g, P_ε) es muestreándolo, siendo pues, para utilizar esta técnica de perturbación, imprescindibles los muestreos de los problemas original y perturbado.

En conclusión, este método no depende de datos desconocidos pero, en contrapartida, resulta muy costoso.

4.5.4 Tercer método de perturbación

Ya no es necesaria la hipótesis:

$$P \in T_m \times \dots \times T_m.$$

Este tercer método de perturbación intenta ser aplicable: es decir, no depende de datos desconocidos; y también operativo: está basado en un único muestreo. En este método se altera, al igual que en el primero, el vector de pagos g . Anteriormente, se fijaba un parámetro ε y se aumentaban algunas componentes del vector g en una cantidad ε , mientras que las otras se conservaban iguales. La técnica que ahora se va a explicar aumenta el valor de g en todos los estados de S , pero estos aumentos son de diferente magnitud en cada estado.

El vector de pagos perturbado será denotado por h_ε , conservando la notación g_ε para la primera perturbación. Se define, para $\varepsilon > 0$:

$$h_\varepsilon(s_i) = g(s_i) + \varepsilon^i, \text{ para } 1 \leq i \leq m,$$

y $h_\varepsilon(s_{m+1}) = g(s_{m+1}) = 0$. Esta perturbación aumenta el pago del primer estado de S en una cantidad ε , en el segundo estado el aumento es de ε^2, \dots , en el m -ésimo de ε^m . Hay que hacer notar que la perturbación en un estado es un infinitésimo de la perturbación en el estado anterior.

El problema perturbado (h_ε, P) es del mismo tipo que el problema original (g, P) , a saber: que el vector de pagos es mayor o igual que $\mathbf{0}$, y que

el pago en el estado absorbente s_{m+1} es nulo. Tiene además una única regla óptima, como prueba el siguiente resultado.

PROPOSICIÓN 4.24: *Existe un número $\eta > 0$ tal que si $0 < \varepsilon < \eta$, entonces el problema perturbado (h_ε, P) tiene una única regla óptima, que es también óptima para el problema original (g, P) .*

DEMOSTRACIÓN: Dada una regla $d \in D$, sea v_d su valor en el problema (g, P) y $v_{d,\varepsilon}$ su valor en el problema perturbado. Analicemos la relación que existe entre v_d y $v_{d,\varepsilon}$:

- $v_{d,\varepsilon}(s_{m+1}) = v_d(s_{m+1}) = 0$;
- si $s_i \in S \cap d^{-1}(0)$, es decir, si $d(s_i) = 0$ y $i \neq m + 1$, entonces:

$$v_{d,\varepsilon}(s_i) = v_d(s_i) + \varepsilon^i; \quad (4.51)$$

- si $d(s) = 1$, entonces el estado s es de transición para la cadena de Markov inducida por d , y los estados recurrentes son los de $d^{-1}(0)$. Sea $\pi_{ss'}^d$, para un estado s' tal que $d(s') = 0$, la probabilidad de que, partiendo de s , la absorción por $d^{-1}(0)$ se produzca por s' . Se cumple que:

$$\begin{aligned} v_{d,\varepsilon}(s) &= \sum_{s_i \in S \cap d^{-1}(0)} \pi_{ss_i}^d h_\varepsilon(s_i) \\ &= \sum_{s_i \in S \cap d^{-1}(0)} \pi_{ss_i}^d (g(s_i) + \varepsilon^i) \\ &= \sum_{s_i \in S \cap d^{-1}(0)} \pi_{ss_i}^d g(s_i) + \sum_{s_i \in S \cap d^{-1}(0)} \pi_{ss_i}^d \varepsilon^i, \end{aligned}$$

luego:

$$v_{d,\varepsilon}(s) = v_d(s) + \sum_{s_i \in S \cap d^{-1}(0)} \pi_{ss_i}^d \varepsilon^i. \quad (4.52)$$

Fijemos ahora dos reglas de decisión d y δ distintas. Si son distintas es porque existe $s_j \in S$ tal que la acción que toman d y δ en s_j es distinta, es decir: una de las dos reglas para y la otra continúa. Supongamos que:

$$d(s_j) = 1 \quad \text{y} \quad \delta(s_j) = 0.$$

Aplicando las fórmulas (4.51) y (4.52) se tiene que:

$$v_{d,\varepsilon}(s_j) = v_d(s_j) + \sum_{s_i \in S \cap d^{-1}(0)} \pi_{s_j s_i}^d \varepsilon^i \quad \text{y} \quad v_{\delta,\varepsilon}(s_j) = v_\delta(s_j) + \varepsilon^j,$$

llegando a la expresión:

$$(v_{\delta,\varepsilon}(s_j) - v_{d,\varepsilon}(s_j)) = (v_\delta(s_j) - v_d(s_j)) + \varepsilon^j - \sum_{s_i \in S \cap d^{-1}(0)} \pi_{s_j s_i}^d \varepsilon^i. \quad (4.53)$$

Si las dos cantidades $v_\delta(s_j)$ y $v_d(s_j)$ son distintas, entonces tomando $\varepsilon > 0$ lo bastante pequeño, también lo son las cantidades $v_{\delta,\varepsilon}(s_j)$ y $v_{d,\varepsilon}(s_j)$. Si, en cambio, $v_\delta(s_j)$ y $v_d(s_j)$ son iguales, entonces (4.53) se escribe:

$$v_{\delta,\varepsilon}(s_j) - v_{d,\varepsilon}(s_j) = \varepsilon^j - \sum_{s_i \in S \cap d^{-1}(0)} \pi_{s_j s_i}^d \varepsilon^i.$$

El segundo miembro de esta igualdad es un polinomio en la variable ε que no es el polinomio nulo ya que el índice j no es tal que $d(s_j) = 0$, y por tanto, en el sumatorio no aparece la potencia ε^j , de donde, dicho polinomio no se anula en ningún punto de un intervalo de la forma $(0, \alpha)$, no coincidiendo los valores de d y δ en s_j cuando ε pertenece a dicho intervalo.

En conclusión, dadas dos reglas de decisión d y δ distintas, tomando ε en un intervalo de la forma $(0, \alpha)$, resulta que las reglas d y δ no tienen el mismo valor en el problema perturbado (h_ε, P) , es decir:

$$v_{\delta,\varepsilon} \neq v_{d,\varepsilon}.$$

Repitiendo este procedimiento para cada par de reglas distintas, se concluye que tomando ε dentro de un intervalo $(0, \xi)$, para algún $\xi > 0$, no existen dos reglas distintas con el mismo valor en (h_ε, P) y, en particular, la regla óptima del problema (h_ε, P) es única, como queríamos demostrar.

Probemos ahora que tomando un parámetro de perturbación lo bastante pequeño, la regla óptima del problema (h_ε, P) es también óptima para el problema original (g, P) . Sea D^* el conjunto de reglas óptimas de (g, P) y fijemos $d \in D^*$. Tomamos $\delta \in D$, una regla que no esté en D^* . Existe pues un estado $s \in S$ tal que:

$$v^*(s) = v_d(s) > v_\delta(s),$$

es decir:

$$v_d(s) - v_\delta(s) > 0. \quad (4.54)$$

Independientemente de las acciones que ambas reglas tomen en el estado s , se tiene que:

$$(v_{d,\varepsilon}(s) - v_{\delta,\varepsilon}(s)) = (v_d(s) - v_\delta(s)) + Q(\varepsilon),$$

donde Q es un polinomio en la variable ε que tiene su término independiente igual a cero, tendiendo por tanto el polinomio a cero cuando ε tiende a cero. Podemos pues escoger $\varepsilon > 0$ lo bastante pequeño para que:

$$v_{d,\varepsilon}(s) - v_{\delta,\varepsilon}(s) > 0,$$

como se deduce de (4.54), no pudiendo ser pues la regla δ óptima para el problema (h_ε, P) . Denotemos por D_ε^* el conjunto de reglas óptimas de (h_ε, P) . Lo que hemos demostrado no es que la regla d esté en D_ε^* , sino que hemos justificado que δ no está en D_ε^* , luego debemos buscar las reglas óptimas para el problema perturbado en D^* , es decir: existe un intervalo de la forma $(0, \eta)$, con $\eta > 0$, tal que para cada $\varepsilon \in (0, \eta)$:

$$D_\varepsilon^* \subseteq D^*.$$

Por lo probado anteriormente, tomando $0 < \eta < \xi$, resulta además que D_ε^* es uni-puntual, quedando demostrado el resultado. Q.E.D.

Observada una muestra de tamaño n y tomando como vector de pagos h_ε , verificando las condiciones de la Proposición 4.24, sea $d_{n,\varepsilon}$ la estimación de la regla óptima del problema (h_ε, P) . Se define otro estimador del valor del problema original (g, P) : $\hat{v}_\varepsilon^{(n)}$, a través de:

$$\hat{v}_\varepsilon^{(n)}(s) = \tilde{v}_{d_{n,\varepsilon}}^{(n)}(s),$$

para cada estado $s \in \bar{S}$. El valor $\hat{v}_\varepsilon^{(n)}$ es la estimación, en el problema (g, P) , del valor de la regla estimada óptima en el problema (h_ε, P) .

TEOREMA 4.25: *Tomando ε como en la Proposición 4.24, se cumple, para cada $s \in \bar{S}$:*

$$\hat{v}_\varepsilon^{(n)}(s) \xrightarrow{\text{cs}} v^*(s) \quad \text{y} \quad \sqrt{n}(\hat{v}_\varepsilon^{(n)}(s) - v^*(s)) \xrightarrow{\text{d}} \mathcal{N}(0, \cdot),$$

cuando $n \rightarrow +\infty$.

DEMOSTRACIÓN: Según el Teorema 4.17, para casi todo $\omega \in \Omega$ existe un tamaño muestral a partir del cual:

$$d_{n,\varepsilon} = d^{**},$$

siendo:

$$D_\varepsilon^* = \{d^{**}\},$$

es decir que:

$$\hat{v}_\varepsilon^{(n)}(s) = \tilde{v}_{d^{**}}^{(n)}(s).$$

Puesto que la regla d^{**} es óptima para el problema (g, P) , de los teoremas 4.12 y 4.14 se deduce que:

$$\tilde{v}_{d^{**}}^{(n)}(s) \xrightarrow{\text{cs}} v^*(s) \quad \text{y} \quad \sqrt{n}(\tilde{v}_{d^{**}}^{(n)}(s) - v^*(s)) \xrightarrow{\text{d}} \mathcal{N}(0, \cdot),$$

y se concluye el resultado enunciado para el estimador $(\hat{v}_\varepsilon^{(n)}(s))_{n \geq 1}$. La varianza asintótica tiene la misma expresión que la del Teorema 4.21.

Q.E.D.

En resumen, este tercer método de perturbación parece ser el más recomendable a la hora de una aplicación práctica. No requiere ningún conocimiento previo sobre el problema, como el primer método, puesto que altera los pagos de todos los estados de S y, a diferencia del segundo método de perturbación expuesto, requiere un único muestreo. Nótese que esta nueva técnica de perturbación es igualmente válida para la estimación por máxima verosimilitud.

Este método, además, es aplicable bajo hipótesis menos restrictivas que los dos primeros presentados. La hipótesis:

$$P \in T_m \times \overset{m}{\cdots} \times T_m,$$

necesaria para los dos primeros métodos de perturbación, no es necesaria para esta tercera técnica.

En cambio, la regla óptima para el problema perturbado no se determina de forma tan sencilla como en los dos otros procedimientos de perturbación. En los dos primeros métodos, la regla óptima del problema perturbado es:

$$\begin{aligned} d^{**}(s) &= 1, & \text{si } s \in S_{ci}(g, P), \\ d^{**}(s) &= 0, & \text{si } s \in S_p(g, P), \end{aligned}$$

mientras que para la tercera técnica no existe una expresión general para la única regla de D_ε^* .

4.6 Conclusiones

Han quedado cubiertos los objetivos que se plantearon en la introducción puesto que se ha logrado estimar consistentemente una regla óptima, y se han propuesto estimadores del valor del problema de decisión consistentes y asintóticamente normales, independientemente de la unicidad de la regla óptima. Todo ello se ha hecho sin estimar la matriz de transición de la cadena de Markov, empleando un método directo que podríamos calificar como de: “estimar sin resolver”.

En el desarrollo de este estudio, en mi opinión, ha habido tres puntos fundamentales que resumen la esencia del trabajo.

El primero es el haber utilizado la técnica de martingalas para analizar el comportamiento de los estimadores. Resulta más natural, y en cierto modo más sencillo, el método markoviano de estudio de la cadena $C^d(s)$, aunque éste requiere hipótesis algo más restrictivas: irreductibilidad y aperiodicidad de $C^d(s)$. Las martingalas no han surgido de forma natural, pero han resuelto de forma muy elegante el estudio de los estimadores propuestos.

El segundo aspecto esencial es darse cuenta de que al proponer los estimadores basados, no en la muestra de estados, sino en la muestra de tramos, se ha producido, como ya se ha comentado, un cambio de espacio muestral. Formalmente, no ha habido tal cambio de espacio muestral puesto que todos los desarrollos se han basado en un único espacio: (Ω, \mathcal{F}) , pero desde un punto de vista conceptual, es irrefutable que el problema ha cambiado de espacio puesto que se ha cambiado de muestreo. Es de nuevo gracias a las martingalas que dicho cambio de espacio muestral se ha podido deshacer para volver al problema originalmente planteado, lo que no parece viable con el método de la cadena $C^d(s)$.

El tercer punto importante es la introducción del método de perturbación por potencias de ε . Esta perturbación supera a las otras dos propuestas en generalidad y operatividad. Con esta técnica ha quedado zanjado el estudio de la estimación del valor del problema.

Es interesante comparar este método con la técnica de máxima verosimilitud de: “resolver y estimar”. Lo primero que hay que señalar es que las demostraciones de los teoremas han resultado más difíciles en este estudio que en el anterior. Esto es porque el método de la máxima verosimilitud, gracias a los resultados de BILLINGSLEY, tenía, en el conocimiento del comportamiento asintótico de la estimación de la matriz de transición, una base

teórica muy sólida, a partir de la cual se analizaban las demás propiedades. En este método, precisamente por tratarse de un método directo, ha habido que hacer demostraciones específicas.

Se ha producido una ganancia en generalidad al suprimir la hipótesis:

$$P \in T_m \times \cdots \times T_m.$$

Esta suposición fue necesaria en la técnica de máxima verosimilitud para poder garantizar la convergencia en distribución de la estimación de la matriz de transición, para que dicha matriz variase en un conjunto abierto. No era, en cambio, necesaria para analizar la convergencia casi segura. En el método directo no ha hecho falta esta hipótesis. El motivo es que los teoremas centrales del límite han sido probados por técnicas de martingalas, y las condiciones suficientes para dichos teoremas están basadas en convergencias en probabilidad, que luego han resultado ser casi seguras, no necesitando para probar esas convergencias, como se ha observado, que los parámetros varíen en un conjunto abierto.

En conclusión, con hipótesis menos restrictivas y con menos base teórica, el método directo ha llegado a los mismos resultados, aunque con más esfuerzo, que el método de máxima verosimilitud. Puede parecer paradójico que con suposiciones más débiles se alcancen resultados similares. En realidad, esto no tiene nada de contradictorio. Todos los resultados de BILLINGSLEY, con sus hipótesis correspondientes, han actuado más bien como un lastre, precisamente porque no era necesario estimar la matriz de transición P . Al estimar la matriz de transición, se introdujeron hipótesis que no eran esenciales, y dichas hipótesis se arrastraron debido al “defecto inherente” al método de “resolver y estimar”: la sobre-estimación. Suprimida, al aplicar el método directo, la sobre-estimación, logramos liberarnos de hipótesis que de hecho eran superfluas.

4.7 Simulaciones

4.7.1 Introducción

A continuación se presenta una simulación del problema de parada óptima. Se ha tomado el valor $m = 7$ para el número de estados de transición de la cadena de Markov. Las probabilidades de transición entre los estados

transitorios de la cadena de Markov son:

$$P = \begin{pmatrix} 0,13 & 0,15 & 0,13 & 0,30 & 0,14 & 0,05 & 0,05 \\ 0,10 & 0,10 & 0,18 & 0,13 & 0,10 & 0,02 & 0,27 \\ 0,03 & 0,06 & 0,01 & 0,11 & 0,17 & 0,32 & 0,10 \\ 0,37 & 0,22 & 0,11 & 0,10 & 0,03 & 0,01 & 0,01 \\ 0,02 & 0,02 & 0,02 & 0,04 & 0,15 & 0,37 & 0,08 \\ 0,15 & 0,08 & 0,12 & 0,10 & 0,05 & 0,10 & 0,05 \\ 0,15 & 0,12 & 0,10 & 0,20 & 0,05 & 0,08 & 0,05 \end{pmatrix},$$

y el vector de pagos es:

$$g = (1,7 \quad 2,6 \quad 4,2 \quad 3,2 \quad 4,5 \quad 5,2 \quad 6,2)'$$

Se ha estimado este problema de parada óptima por los dos métodos estudiados: el método de máxima verosimilitud expuesto en el Capítulo 3 y el método directo desarrollado en el Capítulo 4. Se ha pretendido ilustrar distintas facetas de estos métodos, además de establecer comparaciones entre ellos.

En concreto, para cada uno de los dos procedimientos se ha estudiado la estimación del valor de una regla de decisión concreta, así como la normalidad asintótica de dicha estimación. La regla considerada ha sido la regla d dada por:

$$\begin{aligned} d(s_1) &= 1, & d(s_2) &= 1, & d(s_3) &= 1, & d(s_4) &= 0, \\ d(s_5) &= 1, & d(s_6) &= 0, & d(s_7) &= 1, & d(s_8) &= 0, \end{aligned}$$

cuyo valor es:

$$v_d = (2,8751 \quad 2,5164 \quad 2,9996 \quad 3,2 \quad 2,8326 \quad 5,2 \quad 2,3482 \quad 0)'$$

También se ha analizado la estimación del valor del problema, estudiándose la normalidad asintótica de la estimación. El valor del problema es:

$$v^* = (3,9268 \quad 4,2492 \quad 4,2 \quad 3,443 \quad 4,5 \quad 5,2 \quad 6,2 \quad 0)'$$

Asimismo, se ha tratado la cuestión de la velocidad de convergencia, estimando la probabilidad de que la regla que los algoritmos estiman como regla óptima sea efectivamente óptima. La única regla óptima del problema considerado es la regla d^* :

$$\begin{aligned} d^*(s_1) &= 1, & d^*(s_2) &= 1, & d^*(s_3) &= 0, & d^*(s_4) &= 1, \\ d^*(s_5) &= 0, & d^*(s_6) &= 0, & d^*(s_7) &= 0, & d^*(s_8) &= 0. \end{aligned}$$

Para hacer esta simulación se han generado doscientas muestras de dos mil transiciones de la cadena de Markov.

4.7.2 Simulación del método máximo-verosímil

Estimación del valor de la regla d

Cuando se han observado n transiciones de la cadena de Markov para la simulación k -ésima, con $1 \leq k \leq 200$ y siendo $1 \leq n \leq 2000$, se dispone de una estimación $\hat{v}_d^n(s)$ del valor de la regla d en el estado s . Puesto que conocemos el verdadero valor de la regla d , podemos medir la bondad de la estimación \hat{v}_d^n mediante la divergencia:

$$D_k(d, n) = \max_{s \in S} |\hat{v}_d^n(s) - v_d(s)|. \quad (4.55)$$

Se repite esta operación para cada una de las doscientas simulaciones que se hacen, y se estima la divergencia esperada de la estimación en la etapa n -ésima mediante el promedio de las doscientas cantidades del tipo de (4.55):

$$\hat{D}(d, n) = \frac{1}{200} \sum_{k=1}^{200} D_k(d, n).$$

La figura 1, cf. p.223, es la gráfica de la función $n \mapsto \hat{D}(d, n)$, que mide la velocidad a la que se aproxima la estimación del valor de la regla d a su verdadero valor.

Normalidad asintótica de la estimación

Se quiere analizar desde un punto de vista empírico la normalidad asintótica de la variable:

$$\sqrt{n}(\hat{v}_d^n(s) - v_d(s)).$$

Con las muestras obtenidas, disponemos de doscientas realizaciones de la citada variable para $n = 2000$. En la figura 2, cf. p.223, se muestra el histograma de frecuencias de las doscientas realizaciones de la variable tomando como estado s el estado tercero. La media de la muestra de tamaño doscientos es igual a $-0,009$ y su desviación típica es $6,7$. Se hacen los contrastes de hipótesis de Kolmogorov-Smirnov y de la χ^2 suponiendo, en la hipótesis nula, que la muestra proviene de una distribución normal de media cero. Los p -valores de los contrastes son $0,88$ y $0,75$, respectivamente.

En conclusión, se observa una convergencia rápida de la estimación del valor de la regla d y un buen ajuste, para una muestra de tamaño $n = 2000$, a una distribución normal.

Estimación de la regla óptima

Fijado un tamaño muestral n , y para cada una de las doscientas simulaciones, se cuenta el número de veces que la estimación d_n de la regla óptima resulta ser óptima para el problema original. Dividiendo el número de fallos en la etapa n por doscientos se obtiene la estimación de la probabilidad de fallo de la estimación de la regla óptima para cada tamaño muestral n , siendo n tal que $1 \leq n \leq 2000$. La figura 3, cf. p.224, representa, en función del tamaño de la muestra, la probabilidad empírica de estimar de forma incorrecta la regla óptima.

Estimación del valor del problema

Para estimar el valor del problema de decisión se procede igual que para la estimación del valor de una regla de decisión. Para $1 \leq n \leq 2000$ y para $1 \leq k \leq 200$ se evalúa:

$$D_k(n) = \max_{s \in S} |\hat{v}^n(s) - v^*(s)|,$$

y en la figura 4, cf. p.224, se representa la gráfica de:

$$n \mapsto \frac{1}{200} \sum_{k=1}^{200} D_k(n),$$

que refleja la velocidad de la convergencia de las estimaciones del valor del problema.

Normalidad asintótica de la estimación

Igual que anteriormente, se estudia la normalidad asintótica de:

$$\sqrt{n}(\hat{v}^n(s) - v^*(s)), \tag{4.56}$$

para el segundo estado. En la figura 5, cf. p.225, se muestra el histograma de frecuencias de las doscientas realizaciones de la variable (4.56). La media muestral es $-0,34$ y la desviación típica $6,15$. Los p -valores de los contrastes de Kolmogorov-Smirnov y de la χ^2 , poniendo como distribución en la hipótesis nula una normal de media cero, son: $0,84$ y $0,38$, respectivamente.

Se observa, al igual que en la estimación del valor de una regla de decisión, una convergencia rápida al verdadero valor del problema y una buena aproximación a la distribución normal.

4.7.3 Simulación del método directo

Estimación del valor de la regla d

Se supone fijada la misma regla de decisión d que en la sección anterior. Observadas n transiciones de la cadena de Markov, para la simulación k -ésima, con $1 \leq k \leq 200$ y siendo $1 \leq n \leq 2000$, se dispone de una estimación $\tilde{v}_d^{(n)}(s)$ del valor de la regla d en el estado s . Se calcula la divergencia:

$$\max_{s \in S} |\tilde{v}_d^{(n)}(s) - v_d(s)|,$$

y se hace el promedio de estas divergencias para las doscientas simulaciones. La figura 6, cf. p.225, representa esta gráfica en la variable n .

Normalidad asintótica de la estimación

Se quiere estudiar la normalidad asintótica de la variable:

$$\sqrt{n}(\tilde{v}_d^{(n)}(s) - v_d(s)).$$

Al igual que para el método de la máxima verosimilitud, tenemos una muestra de tamaño doscientos de dicha variable, siendo $n = 2000$. La figura 7, cf. p.226, es el histograma de frecuencias de la muestra, fijándonos en el primer estado. La media de la muestra de tamaño doscientos es $-0,3$ y la desviación típica es $6,2$. Se hacen los contrastes de hipótesis de Kolmogorov-Smirnov y de la χ^2 suponiendo, en la hipótesis nula, que la muestra proviene de una distribución normal de media cero. Los p -valores de los contrastes son $0,93$ y $0,72$, respectivamente.

También con el método directo de estimación se observa una convergencia rápida de la estimación del valor de la regla d y un ajuste preciso, para una muestra de tamaño $n = 2000$, a una distribución normal.

Estimación de la regla óptima

La velocidad de convergencia a una regla óptima se estima de la misma forma que se hizo en el método de máxima verosimilitud. En la gráfica 8, cf. p.226, se muestran los resultados de esta estimación.

Estimación del valor del problema

Para la estimación del valor del problema de decisión se calcula la divergencia:

$$\max_{s \in S} |\hat{v}^{(n)}(s) - v^*(s)|,$$

y la figura 9, cf. p.227, es la gráfica de la media de las doscientas divergencias observadas.

Normalidad asintótica de la estimación

Se procede igual que en los otros estudios de normalidad asintótica. La figura 10, cf. p.227, muestra el histograma del error de estimación normalizado para el primer estado.

La media muestral es $-0,34$ y la desviación típica $6,05$. Los p -valores de los contrastes de Kolmogorov-Smirnov y de la χ^2 , poniendo como distribución en la hipótesis nula una normal de media cero, son: $0,73$ y $0,78$, respectivamente.

Se comprueba una convergencia rápida al verdadero valor del problema y un buen ajuste a la normalidad.

4.7.4 Comparación de los dos métodos

Puesto que disponemos de dos métodos distintos para estimar un mismo problema, resulta interesante su comparación. Para poderse llevar a cabo ésta de forma más precisa, en las simulaciones que se han hecho para ambos métodos se han utilizado las mismas muestras.

En las gráficas comparativas, la línea continua corresponderá al método directo y la línea discontinua al método de máxima verosimilitud.

En cuanto a la estimación del valor de una regla, se comparan ambos procedimientos en la figura 11, cf. p.228. Se observa que, para tamaños muestrales pequeños, la estimación que proporciona el método directo es mucho más precisa que la del método de máxima verosimilitud. Para tamaños muestrales grandes ambos procedimientos tienen una eficiencia similar.

Se sacan las mismas conclusiones de la figura 12, cf. p.228, que compara la bondad de la estimación del valor del problema para ambos métodos.

Para la velocidad de convergencia a una regla óptima, véase la figura 13, cf. p.229, se observa también que hay una menor probabilidad de fallo empleando el método directo para tamaños muestrales pequeños, mientras que el

comportamiento del método máximo-verosímil es mejor para tamaños muestrales grandes.

Se ha podido comprobar que la convergencia a la distribución normal es más rápida para los estimadores del método directo que para los de máxima verosimilitud.

En conclusión, para la estimación del valor del problema ambos procedimientos son igualmente precisos cuando la muestra es grande, pero el método directo supera netamente al de máxima verosimilitud para tamaños muestrales pequeños, que en este caso serían los inferiores a quinientas observaciones. Uno de los motivos es que la definición de los estimadores de máxima verosimilitud requiere el cálculo de la inversa de la matriz:

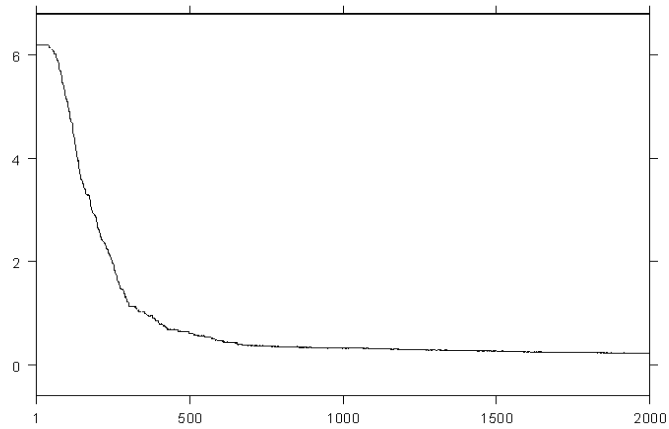
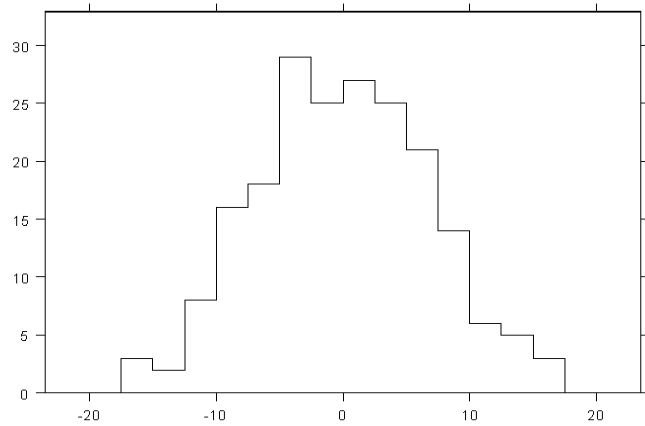
$$I - \hat{P}_d^{(n)},$$

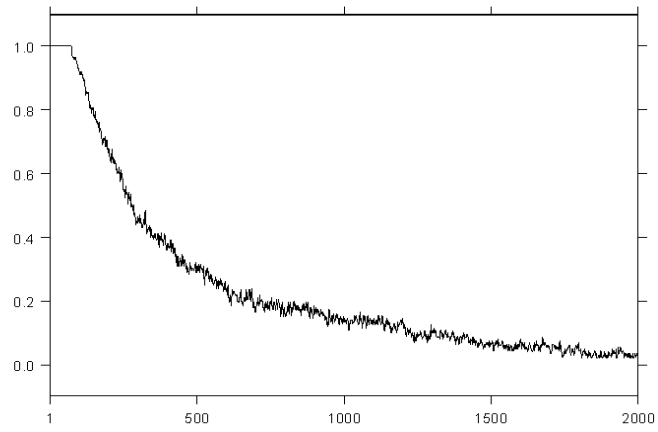
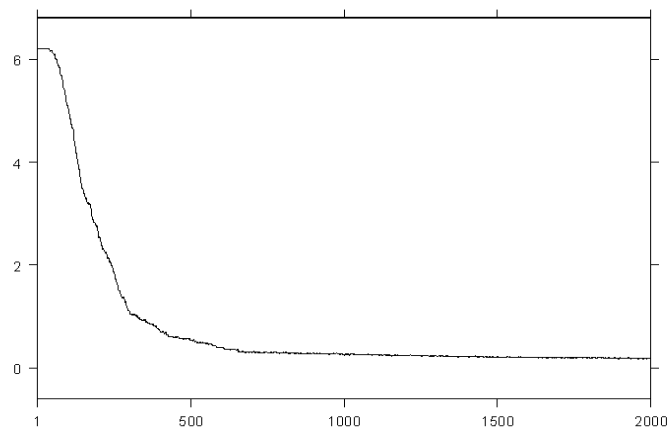
y esta inversa no existe mientras la matriz $\hat{P}^{(n)}$ no tenga “suficientes” componentes estrictamente positivas. Es decir, para poder comenzar la estimación por máxima verosimilitud se necesita haber observado ya unas ciertas transiciones entre estados de \bar{S} , lo que retrasa la aplicación eficiente del método de máxima verosimilitud. En cambio, el método directo es menos exigente a la hora de comenzar a hacer estimaciones, de ahí que para una misma muestra se obtengan, cuando aún se han observado pocas transiciones, errores de estimación que decrecen más rápidamente para el método directo que para el de máxima verosimilitud.

Aunque, en apariencia, ambas estimaciones parezcan equivalentes para tamaños muestrales grandes, el tiempo de computación de ordenador que precisa la estimación por el método de máxima verosimilitud es mucho mayor que el que requiere el método directo. En concreto, para un tamaño de problema $m = 7$ el procedimiento de máxima verosimilitud ha resultado ser treinta y dos veces más lento que el procedimiento directo. Podemos pues afirmar, a la vista de los resultados de las simulaciones, que a la hora de una aplicación práctica es ventajoso utilizar el método directo, ya que proporciona estimaciones mejores para tamaños muestrales pequeños y es computacionalmente más eficiente que el procedimiento de máxima verosimilitud, alcanzando, para tamaños muestrales grandes, niveles de precisión similares.

Estos resultados empíricos confirman la tesis sostenida en la introducción de esta memoria, a saber: que es mejor emplear, para hacer una estimación, técnicas específicas para esa estimación en concreto que estimar el modelo completo.

4.7.5 Apéndice de figuras

Figura 1. Método máximo-verosímil. Divergencia de v_d .Figura 2. Método máximo-verosímil. Normalidad de \hat{v}_d^n .

Figura 3. Método máximo-verosímil. Estimación de d^* .Figura 4. Método máximo-verosímil. Divergencia de v^* .

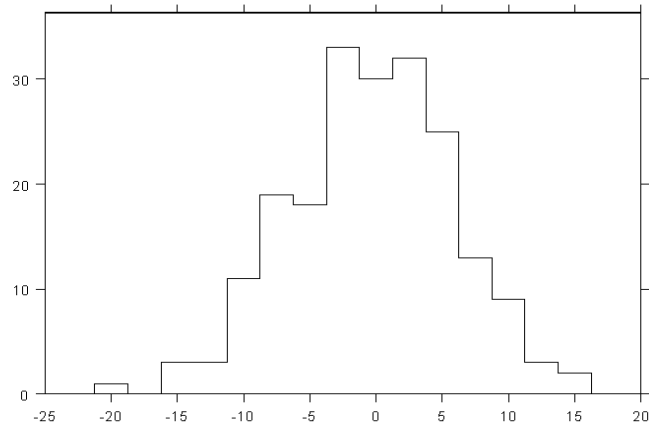


Figura 5. Método máximo-verosímil. Normalidad de \hat{v}^n .

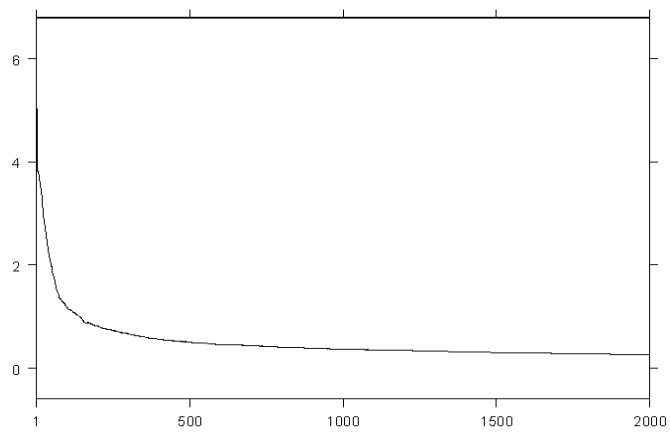


Figura 6. Método directo. Divergencia de v_d .

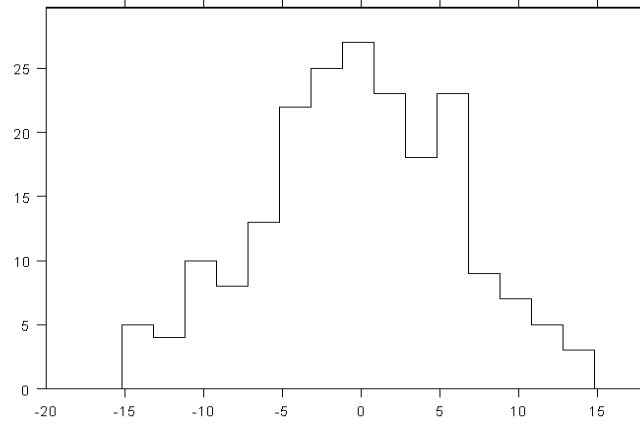


Figura 7. Método directo. Normalidad de $\tilde{v}_d^{(n)}$.

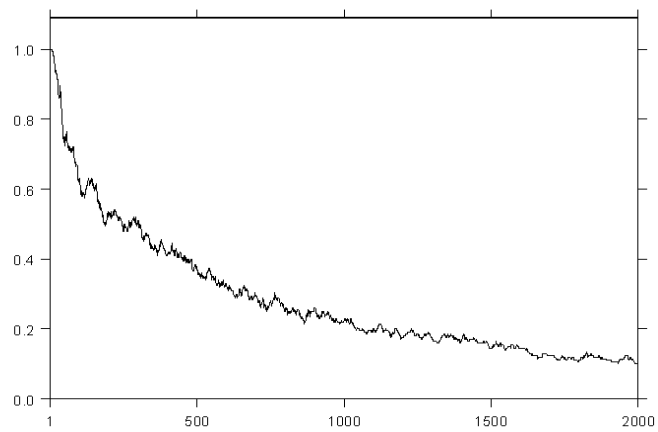
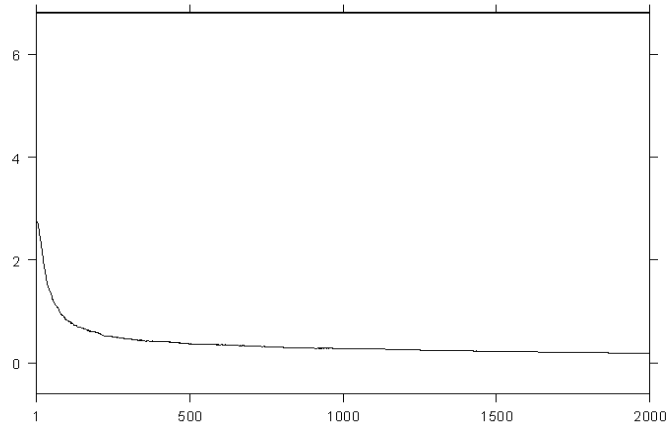
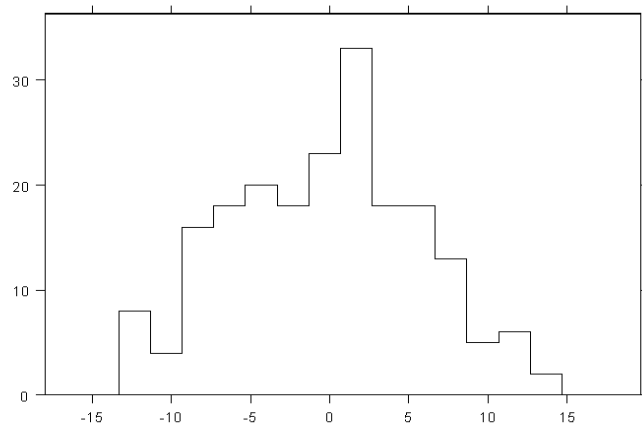
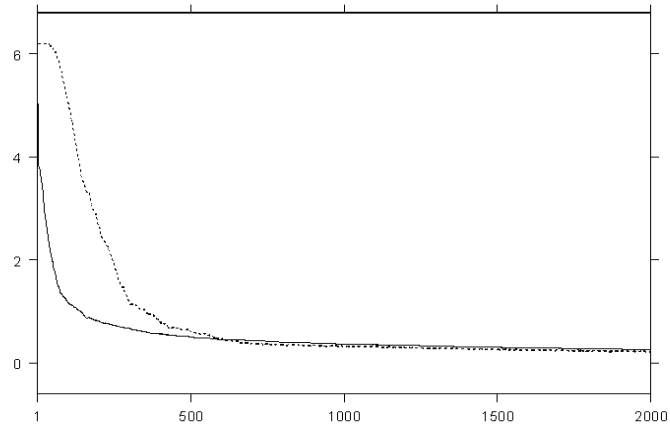
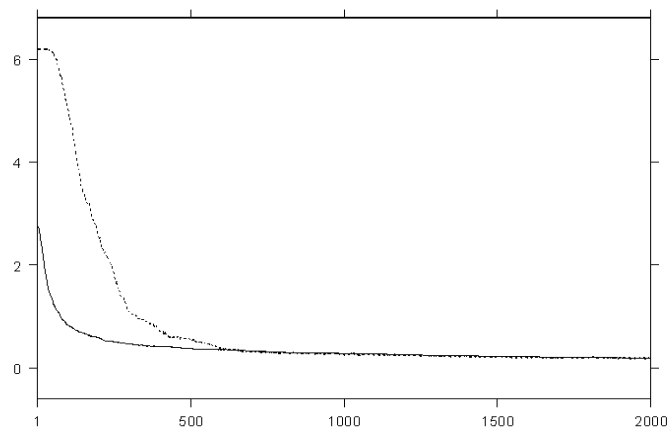


Figura 8. Método directo. Estimación de d^* .

Figura 9. Método directo. Divergencia de v^* .Figura 10. Método directo. Normalidad de $\hat{v}^{(n)}$.

Figura 11. Comparación de las divergencias de v_d .Figura 12. Comparación de las divergencias de v^* .

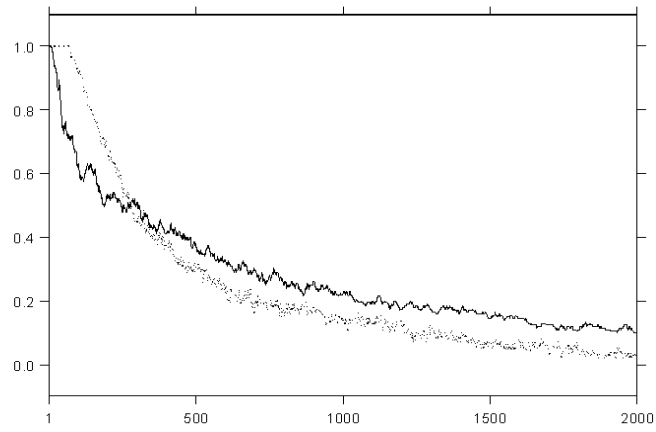


Figura 13. Comparación de las estimaciones de d^* .

Referencias

- ASH R.B., *Real Analysis and Probability*. Academic Press, 1972.
- BAZARAA M.S., *Programación Lineal y Flujo en Redes*. Limusa, 1995.
- BILLINGSLEY P., *Statistical Inference for Markov Processes*. University of Chicago Press, 1961.
- HALL P., HEYDE C.C., *Martingale Theory and Its Application*. Academic Press, 1980.
- IBARROLA P., PARDO L., QUESADA V., *Teoría de la Probabilidad*. Síntesis, 1997.
- LAHA R.G., ROHATGI V.K., *Probability Theory*. Wiley, 1979.
- PARDO L., FELIPE A., PARDO J.A., *Programación Lineal Entera*. Díaz de Santos, 1990.
- PUTERMAN M.L., *Markov Decision Processes*. Wiley, 1994.
- ROMANOVSKY V.I., *Discrete Markov Chains*. Wolters Noordhoff Publishing, 1970.
- WOLSEY L.A., *Integer Programming*. Wiley Interscience, 1998.