00384 3 27

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

DIVISION DE ESTUDIOS DE POSGRADO

Modelos de Control Semi-Markoviano en Espacios de Borel

T E S I S

QUE PARA OBTENER EL GRADO ACADEMICO DE

Doctor en Ciencias (Matemáticas)

P R E S E N T A

Fernando Luque Vásquez

Director de Tesis: Dr. Onésimo Hernández Lerma

A Eva, Eva Luz y Fernando.

A la memoria de mi madre.

A mi padre.

A mis hermanos y sobrinos...

TESIS: Modelos de Control Semi-Markoviano en Espacios de Borel
Alumno: M. en C. Fernando Luque Vásquez

In this thesis we deal with semi-Markov control models with Borel state and control spaces, possibly unbounded cost and not necessarily compact constraint sets. We consider optimal control problems with performance indexes in discounted cost and average cost. The objective is to extend to the semi-Markovian context, classical results on Markov control models. In the problem with discounted cost index, we study the existance of optimal stationary policies, a characterization of the optimal policies and some results of approximation. Under the regularity and continuity-compacity conditions we show the existance of optimal stationary policies and we characterize the optimal cost function as the minimal solution of the optimality equation. Under suita ble conditions, the optimal policies are those whose cost function is solution of the optimality equation. In the corresponding part of the approximations we show the convergence of the value iteration and the policy iteration algorithms; other approximations are also discussed. The average caseis studied in two parts: In the first one, we considerthe problem using the approximation with discounted problems; under the regularity and continuity compacity conditions and impossing conditions on the optimal cost functions of the discounted case, we show the existance of solutions to the optimality equation and the existence of optimal stationary policies. In the second part, the problem is studied applying a transformation on the semi-Markov model and a Markov model is obtained from which is possible deduce results of the original model. In addition of the regularity and continuity-compacity conditionswe consider ergodicity conditions; We show the existence of solutions to the inequality of optimality and of the optimality equation and the existence of stationary optimal policies.

Agradecimientos

Agradezco al Dr. Onésimo Hernández Lerma su asesoría y apoyo durante mis estudios de doctorado. Aprovecho para dejar constancia de mi reconocimiento al trabajo que realiza por el desarrollo de la Matemática y la formación de matemáticos en México.

Doy las gracias al Dr. Miguel Angel García Alvarez por su ayuda en la realización de este trabajo y por el apoyo que me ha brindado en distintas etapas de mi carrera.

Muy especialmente, agradezco al Dr. Juan Ruiz de Chávez y a Blanca Edith Reyes, por el valioso apoyo que de ellos recibí durante mi programa de doctorado.

Agradezco a mis compañeros Oscar Vega Amaya, Ma. Teresa Robles Alcaraz y Adolfo Minjárez Sosa por su colaboración y su interés en este trabajo.

No hay palabras para agradecer a mi padre, a mi esposa y a mis hijos por su comprensión e incondicional apoyo que siempre me han brindado. Consideren este trabajo como de ustedes...

Contenido

1 Introducción **3**

2 El modelo de control semi-Markoviano **6**

2.1 Introducción. 6
2.2 Descripción del Modelo 7
2.3 Indices de funcionamiento 10
2.4 Condiciones de regularidad, continuidad y compacidad 12
2.5 Comentarios y conclusiones 15

3 Indice de costo descontado **16**

3.1 Introducción. 16
3.2 La ecuación de optimalidad en costo $\alpha-$descontado 17
3.3 Condiciones suficientes para (3.4) 22
3.4 Aproximaciones 24
3.5 Comentarios y conclusiones 27

4 Indice de costo promedio I **29**

4.1 Introducción 29
4.2 La desigualdad de optimalidad 30
4.3 La ecuación de optimalidad 33
4.4 Comentarios y conclusiones 35

5 Indice de costo promedio II **36**

5.1 Introducción 36
5.2 Condiciones de optimalidad 37
5.3 El modelo Markoviano asociado 40
5.4 Resultados de optimalidad 44
5.5 Comentarios y conclusiones 46

6 Conclusiones y problemas abiertos **47**

6.1 Introducción 47
6.2 Conclusiones 47

6.3 Problemas abiertos 48

A Cadenas de Markov **49**

Referencias **52**

Capítulo 1

Introducción

En esta tesis se estudian los modelos de control semi-Markovianos en los que el espacio de estados es un espacio de Borel, con costos posiblemente no acotados y conjuntos de control no necesariamente compactos. Se analizan problemas de control óptimo con índices de funcionamiento en costo descontado y en costo promedio.

Esta clase de modelos, cuyas aplicaciones se encuentran principalmente en teoría de inventarios, teoría de colas y modelos de mantenimiento-reemplazo (ver [5], [21], [22], [30], [32]), ha sido ampliamente estudiada en el caso en que el espacio de estados es un conjunto numerable (ver [4], [6], [5], [11], [21], [30], [32], [35], [37], [38], [41]). En el caso en que el espacio de estados es un espacio de Borel, podemos mencionar [2], [3], [22], [33] y, hasta donde conocemos, modelos de control semi-Markoviano con la generalidad que se considera aquí, se han estudiado únicamente en [24], [43] y [44].

El objetivo de este trabajo es extender al contexto semi-Markoviano mencionado, algunos resultados clásicos para modelos de control Markoviano.

La parte correspondiente al caso descontado (Capítulo 3), se basa en [24], [44] y en ella se estudia la existencia de políticas óptimas, iteración de valores, iteración de políticas y otras aproximaciones. La existencia de políticas óptimas para este problema cuando el espacio de estados es de Borel, se estudia en [2] donde se consideran conjuntos de control compactos, una condición de continuidad débil en la ley de transición y semi-continuidad en la función de costo. En este trabajo se consideran conjuntos de control no necesariamente compactos, se impone la condición de inf-compacidad a la función de costo y una condición de continuidad fuerte en la ley de transición. El análisis de aproximaciones a la función de costo descontado cuando el espacio de estados es espacio de Borel, se hace por primera vez en [24]. Los resultados que se obtienen, extienden algunos de los resultados en [15] para modelos de control Markoviano. El análisis del problema con costo promedio, se hace en dos partes. En la primera (Capítulo 4), se estudia la existencia de políticas óptimas mediante el método del factor desvaneciente y, escencialmente se extienden algunos resultados en [27] y [28]. Para modelos de control semi-Markoviano, este método se utiliza en [3] con las mismas condiciones de continuidad-compacidad que en [2]. Las condiciones en la familia de funciones de costo relativo que se consideran en esta tesis, son mas generales que las condiciones de equicontinuidad y acotamiento uniforme que se introducen en [3]. En la segunda parte (Capítulo 5), se estudia el mismo problema usando condiciones de ergodicidad en el modelo del tipo de condiciones utilizadas en [8] y [9] para modelos de control Markoviano. Aplicando una transformación, introducida por P.J. Schweitzer en [37], se obtiene un modelo de control Markoviano del cual es posible deducir resultados del modelo original. Para modelos de control semi-Markoviano con espacio de estados numerable, este método se usa en [5], [6], [30], [37], [40] y [41] pero, hasta donde conocemos, no se había usado en modelos de control semi-Markoviano con espacio de estados de Borel.

El contenido de la tesis está organizado de la siguiente manera: Después de esta introducción, en el capítulo 2 se definen los elementos principales del modelo de control, los índices de funcionamiento, se enuncian algunas condiciones de regularidad, continuidad y compacidad y se demuestran algunos resultados que se utilizan en los siguientes capítulos. En el capítulo 3 se estudia el problema con costo descontado y en los capítulos 4 y 5 se presentan los resultados correspondientes al problema en costo promedio. En el capítulo 6, se presentan las conclusiones del trabajo y algunos problemas abiertos. En el apéndice se incluyen las definiciones de algunos conceptos y resultados que son utilizados en la tesis.

Terminología y notación. Dado un *espacio de Borel* X, es decir, un subconjunto de Borel de un espacio métrico separable completo, $\mathcal{B}(X)$ denota la $\sigma-$álgebra de Borel de X. Si X y Y son espacios de Borel, un *kérnel estocástico* o *probabilidad de transición* en X dado Y es una función $P(\cdot \mid \cdot)$ tal que $P(\cdot \mid y)$ es una medida de probabilidad en X para cada $y \in Y$ fijo y $P(B \mid \cdot)$ es una función medible en Y para cada $B \in \mathcal{B}(X)$ fijo. $M(X)$ denota la familia de funciones reales medibles definidas en X y $M_+(X)$ es la subfamilia de funciones medibles no negativas. $\mathbf{R}$ $(\mathbf{R}_+)$ denota el conjunto de los números reales (no negativos).

Capítulo 2

El modelo de control semi-Markoviano

2.1 Introducción.

2.2 Descripción del modelo de control.

2.3 Indices de funcionamiento.

2.4 Condiciones de regularidad, continuidad y compacidad.

2.5 Comentarios y conclusiones.

2.1 Introducción.

En este capítulo se definen los elementos principales del modelo de control que estudiaremos en este trabajo. También se definen los criterios de optimalidad o índices de funcionamiento que se considerarán en esta tesis y se plantea el problema de control óptimo. Bajo condiciones de regularidad, continuidad y compacidad, se demuestran algunos resultados preliminares que se utilizarán en los siguientes capítulos.

2.2 Descripción del Modelo

Definición 2.1. El *modelo de control semi-Markoviano* (MCSM). que denotaremos por $(X, A, \{A(x) \mid x \in X\}, Q, F, D, d)$ consiste en los siguientes elementos:

1. Un espacio de Borel X, llamado el *espacio de estados.*

2. Un espacio de Borel A. llamado el *conjunto de controles* (o *acciones*).

3. Una colección $\{A(x) \mid x \in X\}$ de subconjuntos medibles no vacíos de A. donde $A(x)$ representa el *conjunto de controles admisibles* cuando el sistema está en el estado $x \in X$; el conjunto

$$\mathbf{K} := \{(x, a) \mid x \in X,\ a \in A(x)\}$$

es un subconjunto de Borel de $X \times A$ y supondremos que contiene la gráfica de una función medible $f : X \longrightarrow A$, de modo que $f(x) \in A(x)\ \forall x \in X$. Denotamos por $\mathbf{F}$ al conjunto de funciones con esta propiedad.

4. Una probabilidad de transición $Q(\cdot \mid \cdot)$ en X dado $\mathbf{K}$, llamada la *ley de transición.*

5. Una función medible $(t, (x, a)) \longmapsto F(t \mid x, a)$ sobre $\mathbf{R} \times \mathbf{K}$ tal que $F(\cdot \mid x, a)$ es una función de distribución para cada $(x, a) \in \mathbf{K}$, llamada la *distribución de los tiempos de transición.*

6. Las funciones de costo D y d en $M_+(\mathbf{K})$.

Dos casos particulares importantes de MCSMs son los siguientes:

a) *Modelo de control Markoviano en tiempo discreto* (MCM). Es el caso en que la distribución de los tiempos de transición está dada por

$$\begin{aligned} F(t \mid x, a) \ &:= 0 \quad \text{si } t < 1 \\ &:= 1 \quad \text{si } t \geq 1. \end{aligned}$$

b) *Modelo de control Markoviano en tiempo continuo:* En este caso, la función de distribución de los tiempos de transición está dada por:

$$F(t \mid x, a) \quad := 1 - \exp(-\lambda(x, a)t) \;\; \text{si } t \geq 0$$

$$:= 0 \;\; \text{si} \;\; t < 0.$$

con $\lambda : \mathbf{K} \longrightarrow (0, \infty)$.

Interpretación. Un modelo de control semi-Markoviano. representa un sistema estocástico controlado que evoluciona de la siguiente manera: En el tiempo $t = 0$. el sistema se encuentra en un estado inicial $x_0 \in X$ y se elige un control $a_0 \in A(x_0)$; entonces ocurre lo siguiente: Se produce un costo inmediato $D(x_0, a_0)$; el sistema permanece en el estado x_0 un tiempo δ_1 el cual es una variable aleatoria con función de distribución $F(\cdot \mid x_0, a_0)$; transcurrido δ_1, el sistema evoluciona al estado $x_1 \in X$ de acuerdo a la ley de probabilidad $Q(\cdot \mid x_0, a_0)$ y por el tiempo transcurrido hasta que la transición ocurre, se produce un costo cuya razón es $d(x_0, a_0)$. Cuando el sistema se encuentra en el estado x_1, se elige un control a_1 y el proceso continúa en esta forma indefinidamente.

Sea $H_0 := X$ y para $n = 1, 2, \ldots$, definimos el conjunto de historias admisibles hasta la n-ésima transición por

$$H_n := (\mathbf{K} \times \mathbf{R})^n \times X.$$

Un elemento en H_n, que llamaremos una $n-$*historia*, es un vector de la forma

$$h_n = (x_0, a_0, \delta_1, \ldots, x_{n-1}, a_{n-1}, \delta_n, x_n) \tag{2.1}$$

con $(x_k, a_k, \delta_{k+1}) \in \mathbf{K} \times \mathbf{R}_+$ para $k = 0, 1, \ldots, n-1$, y $x_n \in X$.

Definición 2.2. Una *política o estrategia de control* es una sucesión $\pi = \{\pi_n\}_{n=0}^{\infty}$ de probabilidades de transición en A dado H_n que satisfacen la restricción $\pi_n(A(x_n) \mid h_n) = 1 \; \forall h_n \in H_n$, $n = 0, 1, \ldots$. Denotamos por Π al conjunto de todas las políticas.

Una política $\pi = \{\pi_n\} \in \Pi$ se dice que es *estacionaria*, si existe $f \in \mathbf{F}$ tal que $\pi_n(B \mid h_n) = I_B(f(x_n))\ \forall B \in \mathcal{B}(A)$. Identificamos a la política π con f y a la subfamilia de las políticas estacionarias con el conjunto $\mathbf{F}$.

Sea $(\Omega, \mathcal{A})$ el espacio medible canónico en el que Ω es el espacio producto $(X \times A \times \mathbf{R}_+)^\infty$ y $\mathcal{A}$ la $\sigma-$álgebra producto correspondiente. Los elementos en Ω son de la forma $\varpi = (x_0, a_0, \delta_1, x_1, a_1, \delta_2, ...)$ y las proyecciones o variables coordenadas x_n, a_n y δ_{n+1} con valores en X, A, y $\mathbf{R}_+$ respectivamente, son las variables de estado, control y tiempo de transición. Por un teorema de C. Ionescu Tulcea (ver [1] Theorem 2.7.2, p. 109), para cada estado inicial $x \in X$ y cada política $\pi \in \Pi$, existe una medida de probabilidad P_x^π tal que $\forall B \in \mathcal{B}(A)$, $C \in \mathcal{B}(X)$ y $h_n \in H_n$ como en (2.1), $n = 0, 1, ...$, se tiene:

$$P_x^\pi(x_0 = x) = 1, \tag{2.2}$$

$$P_x^\pi(a_n \in B \mid h_n) = \pi_n(B \mid h_n), \tag{2.3}$$

$$P_x^\pi(x_{n+1} \in C \mid h_n, a_n, \delta_{n+1}) = Q(C \mid x_n, a_n), \tag{2.4}$$

y

$$P_x^\pi(\delta_{n+1} \leq t \mid h_n, a_n) = F(t \mid x_n, a_n). \tag{2.5}$$

Las variables aleatorias $\delta_1, \delta_2, ...$ son condicionalmente independientes dado el proceso $x_0, a_0, x_1, a_1, ...$

Para una política $\pi \in \Pi$, la variable x_n describe el estado del sistema en el tiempo de la $n-$ésima transición, cuando se eligen los controles de acuerdo a la política π. Es claro que, en general, dicho estado depende de la evolución del sistema en las primeras $n-1$ transiciones, pero en el caso de una política estacionaria f, $\{x_n\}_{n=0}^\infty$ resulta ser una cadena de Markov con probabilidad de transición

$$Q(\cdot \mid x, f(x)).$$

Lo anterior es una consecuencia de propiedades de la esperanza condicional y la propiedad (2.4) a la cual llamaremos la *propiedad de Markov*.

Sea $T_0 \equiv 0$ y para $n = 1, 2, ...$ sea $T_n := \sum_{k=1}^n \delta_k$; es decir, T_n es el tiempo de la $n-$ ésima transición.

Para $t \geq 0$, sea $y_t = x_n$ si $T_n \leq t < T_{n+1}$. Si la distribución de los tiempos de transición $F(\cdot \mid x, a)$ es de la forma definida en b), entonces $\{y_t; t \geq 0\}$ es un proceso de Markov, mientras que en cualquier otro caso (no exponencial), $\{y_t; t \geq 0\}$ es un proceso no-Markoviano.

Denotaremos por E_x^π a la esperanza correspondiente a P_x^π.

Nota 2.3. Para una política estacionaria $f \in \mathbf{F}$, usaremos la notación abreviada (x, f) por $(x, f(x))$. En particular, $D(x, f) := D(x, f(x))$, $d(x, f) := d(x, f(x))$ y $Q(\cdot \mid x, f) := Q(\cdot \mid x, f(x))$.

2.3 Indices de funcionamiento

Consideraremos el índice de funcionamiento en *costo descontado* y el índice de funcionamiento en *costo promedio*. En el primer caso se considera un descuento en costo continuo (con descuento $\alpha > 0$), es decir, una cantidad de c unidades en el tiempo t es equivalente a $ce^{-\alpha t}$ unidades en el tiempo presente.

Definición 2.4. Para $\alpha > 0$, $x \in X$ y $\pi \in \Pi$, se define el *costo total esperado* $\alpha-$*descontado* (es decir, con descuento α) por

$$V_\alpha(\pi, x) := E_x^\pi[\sum_{n=0}^{\infty} e^{-\alpha T_n}\{D(x_n, a_n) + d(x_n, a_n)\int_0^{\delta_{n+1}} e^{-\alpha t}dt\}].$$

La función de costo óptimo α *− descontado* se define entonces como

$$V_\alpha(x) := \inf_{\pi \in \Pi} V_\alpha(\pi, x)$$

y se dice que una política π^* es α *− óptima* si $V_\alpha(\pi^*, x) = V_\alpha(x)\ \forall x \in X$.

Definición 2.5. Para $x \in X$ y $\pi \in \Pi$, se define el *costo promedio esperado* por

$$J(\pi, x) := \limsup_{n \longrightarrow \infty} \frac{E_x^\pi \sum_{k=0}^{n-1}\{D(x_k, a_k) + \delta_{k+1} d(x_k, a_k)\}}{E_x^\pi T_n}.$$

La función

$$J(x) := \inf_{\pi \in \Pi} J(\pi, x)$$

es el *costo promedio óptimo* y una política $\pi^* \in \Pi$ es *óptima en costo promedio* (CP-óptima) si

$$J(x) = J(\pi^*, x) \quad \forall x \in X.$$

A continuación definimos algunas funciones que intervienen en el desarrollo del trabajo.

Definición 2.6.

$$\Delta_\alpha(x,a) := \int_0^\infty e^{-\alpha t} F(dt \mid x,a),$$

$$\tau_\alpha(x,a) := \frac{1 - \Delta_\alpha(x,a)}{\alpha}, \tag{2.6}$$

$$C_\alpha(x,a) := D(x,a) + \tau_\alpha(x,a) d(x,a).$$

$$\tau(x,a) := \int_0^\infty t F(dt \mid x,a),$$

$$C(x,a) := D(x,a) + \tau(x,a) d(x,a).$$

Usando propiedades de la esperanza condicional, podemos escribir

$$V_\alpha(\pi,x) = E_x^\pi \{ C_\alpha(x_0,a_0) + \sum_{n=1}^{\infty} \prod_{k=0}^{n-1} \Delta_\alpha(x_k,a_k) C_\alpha(x_n,a_n) \} \tag{2.7}$$

y

$$J(\pi,x) := \limsup_{n \to \infty} \frac{E_x^\pi \sum_{k=0}^{n-1} C(x_k,a_k)}{E_x^\pi \sum_{k=0}^{n-1} \tau(x_k,a_k)}.$$

Un *problema de control semi-Markoviano* (PCSM) se compone de un MCSM y un índice de funcionamiento. Dado un PCSM, el *problema de control óptimo* consiste en encontrar una política π^* bajo la cual, el índice de funcionamiento es óptimo; es decir, si $v(\pi, x)$ es el índice de funcionamiento, entonces el objetivo es encontrar una política $\pi^* \in \Pi$ tal que

$$v(\pi^*, x) = \inf_{\pi \in \Pi} v(\pi, x) \quad \forall x \in X.$$

2.4 Condiciones de regularidad, continuidad y compacidad

Como es usual en esta clase de modelos ([24], [30], [32], [33], [38], [44]), supondremos que se satisface la siguiente *condición de regularidad,* la cual garantiza que en un intervalo de tiempo finito, no puede ocurrir un número infinito de transiciones (ver Proposición 2.11 c)).

Hipótesis 2.7 (H2.7). Existen $\varepsilon > 0$ y $\theta > 0$ tal que

$$F(\theta \mid x, a) \leq 1 - \varepsilon \quad \forall (x, a) \in \mathbf{K}.$$

También introduciremos una clase de funciones y una hipótesis que asegurará, en particular, la exitencia de "mínimos medibles" (ver Lema 2.13 a)).

Definición 2.8. Una función $v : \mathbf{K} \longrightarrow \mathbf{R}$ se llama *inf-compacta en* $\mathbf{K}$ si para cada $x \in X$ y $r \in \mathbf{R}$, el conjunto $\{a \in A(x) : v(x, a) \leq r\}$ es compacto.

Observación 2.9. Si $v : \mathbf{K} \rightarrow \mathbf{R}$ es inf-compacta en $\mathbf{K}$ entonces para cada $x \in X$, $v(x, \cdot)$ es semicontinua inferiormente (s.c.i.) y acotada inferiormente en $A(x)$.

Hipótesis 2.10 (H2.10). a) D es inf-compacta en $\mathbf{K}$ y $d(x, \cdot)$ es s.c.i. en $A(x)$ para cada $x \in X$.

b) La ley de transición Q es fuertemente continua en A, es decir, para cada función $v : X \longrightarrow \mathbf{R}$ medible y acotada y $x \in X$, la función v' definida en $A(x)$ por

$$v'(a) := \int v(y) Q(dy \mid x, a)$$

es continua.

c) $F(t \mid x, \cdot)$ es continua en $A(x)$ para cada $x \in X$ y $t \in \mathbf{R}$.

En las siguientes dos proposiciones se demuestran propiedades importantes de las funciones introducidas en la Definición 2.6.

Proposición 2.11. Si H2.7 se cumple, entonces:
a) $\rho := \inf_{\mathbf{K}} \tau(x, a) \geq \varepsilon\theta$,

b) $\rho_\alpha := \sup_{\mathbf{K}} \Delta_\alpha(x,a) < 1$,

c) $P_x^\pi[\sum_{n=1}^\infty \delta_n = \infty] = 1 \ \ \forall x \in X,\ \pi \in \Pi$.

Demostración. a) De la definición de $\tau(x,a)$ y H2.7 se tiene,

$$\begin{aligned}\tau(x,a) \ &:= \textstyle\int_0^\infty tF(dt \mid x,a) = \int_0^\infty (1 - F(t \mid x,a))dt \\ &\geq \textstyle\int_0^\theta (1 - F(t \mid x,a))dt \geq \varepsilon\theta \quad \forall (x,a) \in \mathbf{K},\end{aligned}$$

por lo que $\rho \geq \varepsilon\theta$.

b) Sea $(x,a) \in \mathbf{K}$. De la definición de $\Delta_\alpha(x,a)$, la fórmula de integración por partes y H2.7,

$$\begin{aligned}\Delta_\alpha(x,a) \ &:= \textstyle\int_0^\infty e^{-\alpha t} F(dt \mid x,a) = \alpha \int_0^\infty e^{-\alpha t} F(t \mid x,a)dt \\ &= \alpha[\textstyle\int_0^\theta e^{-\alpha t} F(t \mid x,a)dt + \int_\theta^\infty e^{-\alpha t} F(t \mid x,a)dt] \\ &\leq (1-\varepsilon)(1 - e^{-\alpha\theta}) + e^{-\alpha\theta} < 1\end{aligned}$$

por lo que $\rho_\alpha < 1$.

c) Sean $x \in X$ y $\pi \in \Pi$. Entonces de propiedades de la esperanza condicional y b) se obtiene

$$E_x^\pi[e^{-\alpha \sum_{n=1}^\infty \delta_n} \mid x_0, a_0, x_1, a_1, \ldots] = E_x^\pi[\Pi_{n=0}^\infty \Delta_\alpha(x_n, a_n)] = 0.$$

de donde $\sum_{n=1}^\infty \delta_n = \infty \ \ P_x^\pi - c.d.$ ■

Proposición 2.12. Si se satisface H2.10, entonces para cada $x \in X$ y $\alpha > 0$,

a) $\Delta_\alpha(x,\cdot)$, $\tau_\alpha(x,\cdot)$ y $\tau(x,\cdot)$ son funciones continuas en $A(x)$.

b) C_α y C son funciones inf-compactas en $\mathbf{K}$.

Demostración. La parte a) se sigue de H2.10 c). Para demostrar b), notamos primero que las funciones C_α y C son funciones s.c.i. y, para $x \in X$ y $r \in \mathbf{R}$,

$$D_r := \{a \in A(x) : C_\alpha(x,a) \leq r\} \subset \{a \in A(x) : D(x,a) \leq r\} =: F_r.$$

Puesto que D_r es cerrado y F_r es compacto se tiene que D_r es compacto lo que demuestra la inf-compacidad de C_a. De la misma forma se demuestra la inf-compacidad de C. ■

Finalizamos esta sección con un lema en el que se demuestran dos resultados que serán de gran utilidad en los siguientes capítulos. El primero es un resultado de los llamados "teoremas de selección medible", mientras que el segundo es un resultado de intercambio de límite y el ínfimo.

Lema 2.13. a) Si $v : \mathbf{K} \to \mathbf{R}$ es medible e inf-compacta en $\mathbf{K}$, entonces la función

$$v^*(x) := \inf_{a\in A(x)} v(x,a)$$

es medible y existe $f^* \in \mathbf{F}$ tal que (usando la notación introducida en la Nota 2.3).

$$v^*(x) = v(x, f^*) = \min_{a\in A(x)} v(x,a) \quad \forall x \in X.$$

b) Sean v y v_n $(n = 1, 2, \ldots)$ funciones medibles e inf-compactas en $\mathbf{K}$ tales que $v_n \uparrow v$. Entonces para cada $x \in X$,

$$\lim_n \min_{a\in A(x)} v_n(x,a) = \min_{a\in A(x)} v(x,a) \quad .$$

Demostración. a) Puesto que v es medible e inf-compacta en $\mathbf{K}$, entonces para cada $r \in \mathbf{R}$, el conjunto $U_r := \{(x,a) \in \mathbf{K} : v(x,a) \le r\}$ es medible y el conjunto $\{a \in A(x) : v(x,a) \le r\}$ es compacto. Luego la demostración de a) se sigue del Corolario 4.3 en [31].

b) Puesto que $v_n \uparrow v$, para $x \in X$,

$$v_n(x,a) \le v(x,a) \quad \forall a \in A(x),\ n \in \mathbf{N}.$$

es decir,

$$\min_{a\in A(x)} v_n(x,a) \le \min_{a\in A(x)} v(x,a) =: v'(x)$$

y por lo tanto

$$l(x) := \lim_{n\to\infty} \min_{a\in A(x)} v_n(x,a) \le v'(x).$$

Fijemos $x \in X$ y definamos los conjuntos $A_0 := \{a \in A(x) : v(x,a) = v'(x)\}$ y $A_n := \{a \in A(x) : v_n(x,a) \leq v'(x)\}$. Estos conjuntos son compactos no vacíos y $A_n \downarrow A_0$. Para cada $n \in \mathbf{N}$, sea $a_n \in A_n$ tal que

$$v_n(x,a_n) = \min_{A(x)} v_n(x,a).$$

Por la compacidad y la convergencia $A_n \downarrow A_0$, existe $a_0 \in A_0$ y una subsucesión $\{a_{n_i}\}$ tal que $a_{n_i} \longrightarrow a_0$. Por la monotonía, para cada $n = 1, 2, ...$

$$v_{n_i}(x,a_{n_i}) \geq v_n(x,a_{n_i}) \quad \forall n_i \geq n$$

por lo que haciendo $i \longrightarrow \infty$, la semicontinuidad inferior implica

$$l(x) \geq v_n(x,a_0) \quad \forall n = 1, 2, ...$$

y, por lo tanto,

$$l(x) \geq v'(x). \blacksquare$$

2.5 Comentarios y conclusiones

En este capítulo se definieron los elementos principales del MCSM, los índices de funcionamiento y algunos conceptos relacionados. Se enunciaron las condiciones de regularidad, continuidad y compacidad bajo las cuales se estudiarán los problemas de control óptimo en los siguientes capítulos y se demostraron algunos resultados preliminares.

Nótese que el "teorema de selección medible" en el Lema 2.13 a), generaliza la Proposición D.6 en [13], donde la función v es s.c.i., acotada inferiormente e inf-compacta en $\mathbf{K}$. Asimismo, el Lema 2.13 b) es una generalización del Lema 4.2.4 en [13] donde las funciones v y v_n $(n = 1, 2, ...)$ son s.c.i., acotadas inferiormente e inf-compactas en $\mathbf{K}$.

Capítulo 3

Indice de costo descontado

2.1 Introducción.

2.2 La ecuación de optimalidad en costo α−descontado.

2.3 Condiciones suficientes para (3.4).

2.4 Aproximaciones.

2.5 Comentarios y conclusiones.

3.1 Introducción.

En este capítulo se considera el PCSM con índice de costo descontado que consiste en el MCSM $(X, A, \{A(x) : x \in X\}, Q, F, D, d)$ (ver Definición 2.1) y el criterio de optimalidad de la Definición 2.4. Los resultados que se presentan son una extensión al contexto semi-Markoviano de los resultados en [15] para modelos de control Markoviano. Se demuestra que bajo las condiciones de regularidad, continuidad y compacidad (ver H2.7 y H2.10) e imponiendo una nueva hipótesis (H3.4), la función de valor óptimo es la mínima solución en $M_+(X)$ a la ecuación de optimalidad. También se demuestra la existencia de políticas estacionarias óptimas y algunos resultados de aproximación.

Problemas de este tipo con espacio de estados numerable y conjunto de controles finito o compacto, se consideran en [21], [22], [30] y [32]. En [2] se estudia el problema cuando el espacio de estados es de Borel y el conjunto de controles es un conjunto compacto.

3.2 La ecuación de optimalidad en costo $\alpha-$descontado

Sea $\alpha > 0$ y sean $C_\alpha(\cdot,\cdot)$ y $\Delta_\alpha(\cdot,\cdot)$ como en la Definición 2.6. Decimos que una función medible $u : X \longrightarrow \mathbf{R}$ es una *solución de la ecuación de optimalidad en costo $\alpha-$descontado* ($\alpha-$EOCD) si satisface

$$u(x) = \min_{a\in A(x)}\{C_\alpha(x,a) + \Delta_\alpha(x,a)\int_X u(y)Q(dy \mid x,a)\} \quad \forall x \in X.$$

El resultado principal en esta sección (Teorema 3.5) establece, bajo condiciones adecuadas, la existencia de políticas estacionarias óptimas y una caracterización de la función de valor óptimo $V_\alpha(x)$ como solución de la $\alpha-$EOCD, es decir,

$$V_\alpha(x) = \min_{a\in A(x)}\{C_\alpha(x,a) + \Delta_\alpha(x,a)\int_X V_\alpha(y)Q(dy \mid x,a)\} \quad \forall x \in X. \tag{3.1}$$

Para $x \in X$ y $\pi \in \Pi$, denotamos por $V_\alpha^n(\pi,x)$ $(n = 1,2,...)$ el costo esperado $\alpha-$descontado hasta la $n-$ésima transición, es decir,

$$V_\alpha^n(\pi,x) = E_x^\pi\{\sum_{k=0}^{n-1} e^{-\alpha T_k}\left[D(x_k,a_k) + d(x_k,a_k)\int_0^{\delta_{k+1}} e^{-\alpha s}ds\right]\}.$$

Nótese que

$$V_\alpha^n(\pi,x) \uparrow V_\alpha(\pi,x)$$

y [cf. (2.7)]

$$V_\alpha^n(\pi,x) = E_x^\pi\{C_\alpha(x_0,a_0) + \sum_{k=1}^{n-1}\prod_{j=0}^{k-1}\Delta_\alpha(x_j,a_j)C_\alpha(x_k,a_k)\}.$$

En lo que sigue, usaremos la siguiente notación:

$$\Lambda_\alpha^0 := 1 \quad \text{y} \quad \Lambda_\alpha^k := \prod_{j=0}^{k-1}\Delta_\alpha(x_j,a_j) \quad \text{para } k = 1,2,....$$

por lo que podemos escribir

$$V_\alpha^n(\pi, x) = E_x^\pi[\sum_{k=0}^{n-1} \Lambda_\alpha^k C_\alpha(x_k, a_k)], \quad n = 1, 2, ...$$

y [ver (2.7)]

$$V_\alpha(\pi, x) = E_x^\pi[\sum_{k=0}^{\infty} \Lambda_\alpha^k C_\alpha(x_k, a_k)].$$

Definición 3.1. Para $u \in M_+(X)$, definimos la función Tu en X por

$$Tu(x) = \min_{a \in A(x)} \{C_\alpha(x, a) + \Delta_\alpha(x, a) \int u(y) Q(dy \mid x, a)\}$$

Nótese que la ecuación (3.1) puede escribirse en la forma

$$V_\alpha(x) = TV_\alpha(x)$$

y por el Lema 2.13 a), si se satisface H2.10 entonces T es una función de $M_+(X)$ en $M_+(X)$. Nótese que la ecuación anterior, $V_\alpha = TV_\alpha$, significa que V_α es un *punto fijo* del operador T. Este es un hecho importante en el que algunos autores se han basado para demostrar la ecuación (3.1) usando técnicas de punto fijo (por ejemplo, el Teorema de Punto Fijo de Banach). Aquí seguiremos una variante de estas técnicas, usando la funciones que definimos a continuación.

Definición 3.2. La sucesión $\{v_n\}_{n=0}^\infty$ de *funciones de iteración de valores* se define por $v_0(x) := 0 \; \forall x \in X$ y para $n = 1, 2, ...$

$$\begin{aligned} v_n(x) \quad &:= \min\nolimits_{a \in A(x)} \{C_\alpha(x, a) + \Delta_\alpha(x, a) \textstyle\int v_{n-1}(y) Q(dy \mid x, a)\} \\ &= Tv_{n-1}(x) = T^n v_0(x). \end{aligned}$$

Observación 3.3. Para $x \in X$ y $\pi = \{\pi_n\} \in \Pi$.

$$v_n(x) \leq \int_A C_\alpha(x, a_0) \pi_0(da_0 \mid x) + \int_A \Delta_\alpha(x, a_0) \int_X v_{n-1}(x_1) Q(dx_1 \mid x, a_0) \pi_0(da_0 \mid x).$$

Iterando esta desigualdad se obtiene, para $x \in X$, $n = 1, 2, ...$

$$v_n(x) \leq V_\alpha^n(\pi, x). \tag{3.2}$$

De hecho, para cada n *fijo*, se puede demostrar que la *igualdad* se cumple en (3.2) si π es una política (que depende de n) tal que $V_\alpha^n(\pi,\cdot) = \inf_{\pi'\in\Pi} V_\alpha^n(\pi',\cdot)$.

Denotamos por Π^0, el conjunto de políticas $\pi \in \Pi$ tales que $V_\alpha(\pi,x) < \infty\ \forall x \in X$.

Hipótesis 3.4 (H3.4) Π^0 es no vacío.

Por ejemplo, si H2.10 se cumple y si $0 \le C_\alpha(x,a) \le R\ \forall(x,a) \in \mathbf{K}$ y una constante R, entonces H3.4 se cumple ya que en este caso (por la Proposición 2.11 b)),

$$V_\alpha(\pi,x) \le \sum_{n=0}^{\infty} \rho_\alpha^n R = \frac{R}{1-\rho_\alpha} \quad \forall x \in X,\ \pi \in \Pi.$$

En la siguiente ecuación (3.3) usamos la notación de la Nota 2.3.

Teorema 3.5. Si el MCSM satisface H2.7, H2.10 y H3.4, entonces:

a) $v_n \uparrow V_\alpha$ y V_α es la función mínima (puntualmente) en $M_+(X)$ que es solución de la $\alpha-$EOCD (3.1);

b) Existe $f_\alpha \in \mathbf{F}$ tal que $f_\alpha(x)$ minimiza el lado derecho de (3.1), es decir,

$$V_\alpha(x) = C_\alpha(x,f_\alpha) + \Delta_\alpha(x,f_\alpha)\int_X V_\alpha(y)Q(dy \mid x,f_\alpha) \quad \forall x \in X \tag{3.3}$$

y f_α es $\alpha-$óptima. Recíprocamente, si $f_\alpha \in \mathbf{F}$ es una política $\alpha-$óptima entonces satisface (3.3).

c) Si $\pi^* \in \Pi$ es tal que $V_\alpha(\pi^*,\cdot)$ es solución de la $\alpha-$EOCD y satisface

$$\lim_{n\to\infty} E_x^\pi\left[\Lambda_\alpha^n V_\alpha(\pi^*,x_n)\right] = 0 \quad \forall x \in X,\ \pi \in \Pi^0. \tag{3.4}$$

entonces π^* es $\alpha-$óptima, es decir, $V_\alpha(\pi^*,x) = V_\alpha(x)\ \forall x \in X$.

d) Si C_α es acotada en $\mathbf{K}$, entonces V_α es la *única* solución acotada de la $\alpha-$EOCD.

En la demostración del Teorema 3.5 se usará el siguiente lema.

Lema 3.6. Supóngase que se satisfacen H2.7, H2.10 y H3.4.

a) Si $u \in M_+(X)$ es tal que $u \ge Tu$, entonces $u \ge V_\alpha$.

b) Si u es una función medible tal que $u \le Tu$ y

$$\lim_{n\to\infty} E_x^\pi\left[\Lambda_\alpha^n u(x_n)\right] = 0 \quad \forall x \in X,\ \pi \in \Pi^0, \tag{3.5}$$

entonces $u \leq V_\alpha$.

Demostración. Si $u \geq Tu$, se sigue del Lema 2.13 a) que existe $f \in \mathbf{F}$ tal que $\forall x \in X$,

$$u(x) \geq C_\alpha(x, f) + \Delta_\alpha(x, f) \int u(y) Q(dy \mid x, f).$$

Iterando esta desigualdad se obtiene

$$u(x) \geq E_x^f[C_\alpha(x, f) + \sum_{k=1}^{n-1} \prod_{j=0}^{k-1} \Delta_\alpha(x_j, f) C_\alpha(x_k, a_k)] + E_x^f[\prod_{j=0}^{n-1} \Delta_\alpha(x_j, f) u(x_n)]$$

y, puesto que $u \geq 0$,

$$u(x) \geq V_\alpha^n(f, x).$$

Haciendo $n \to \infty$ se tiene

$$u(x) \geq V_\alpha(f, x) \geq V_\alpha(x) \quad \forall x \in X.$$

b) Supóngase que $u \leq Tu$ y sean $x \in X$ y $\pi \in \Pi^0$. Entonces, para $n = 0, 1, \ldots,$

$$E_x^\pi[\Lambda_\alpha^{n+1} u(x_{n+1}) \mid h_n, a_n] = \Lambda_\alpha^{n+1} \int u(y) Q(dy \mid x_n, a_n)$$

$$= \Lambda_\alpha^n [C_\alpha(x_n, a_n) + \Delta_\alpha(x_n, a_n) \int u(y) Q(dy \mid x_n, a_n) - C_\alpha(x_n, a_n)]$$

$$\geq \Lambda_\alpha^n [u(x_n) - C_\alpha(x_n, a_n)].$$

Luego,

$$\Lambda_\alpha^n C_\alpha(x_n, a_n) \quad \geq -E_x^\pi[\Lambda_\alpha^{n+1} u(x_{n+1}) - \Lambda_\alpha^n u(x_n) \mid h_n, a_n]. \tag{3.6}$$

Tomando esperanza E_x^π en cada lado de (3.6) y sumando sobre $n = 0, 1, \ldots t-1$ se obtiene

$$V_\alpha^t(\pi, x) \geq u(x) - E_x^\pi[\Lambda_\alpha^t u(x_t)].$$

Haciendo $t \to \infty$ en la desigualdad anterior, de (3.5) se sigue que

$$V_\alpha(\pi, x) \geq u(x)$$

y, puesto que $\pi \in \Pi^0$ es arbitraria,

$$V_\alpha(x) \geq u(x). \blacksquare$$

Demostración del Teorema 3.5. Puesto que el operador $T : M_+(X) \longrightarrow M_+(X)$ es monótono (es decir, $u \geq v \Longrightarrow Tu \geq Tv$), se sigue por inducción que la sucesión $\{v_n\} \subset M_+(X)$ es no decreciente y, por lo tanto, existe $u \in M_+(X)$ tal que $v_n \uparrow u$.

Por el Teorema de la Convergencia Monótona se tiene que para $(x, a) \in \mathbf{K}$,

$$C_\alpha(x,a) + \Delta_\alpha(x,a) \int v_n(y) Q(dy \mid x, a) \uparrow C_\alpha(x,a) + \Delta_\alpha(x,a) \int u(y) Q(dy \mid x, a)$$

y por el Lema 2.13 b) se tiene que $\lim_{n\to\infty} Tv_n(x) = Tu(x)$, es decir, u es una solución a la $\alpha-$EOCD

$$u(x) = Tu(x).$$

Por el Lema 3.6 a), se sigue que

$$u(x) \geq V_\alpha(x) \quad \forall x \in X. \tag{3.7}$$

Por otra parte, haciendo $n \longrightarrow \infty$ en (3.2) se obtiene

$$u(x) \leq V_\alpha(\pi, x) \quad \forall x \in X, \ \pi \in \Pi,$$

de donde se sigue que

$$u(x) \leq V_\alpha(x) \quad \forall x \in X. \tag{3.8}$$

Combinando (3.7) y (3.8) se obtiene que $u = V_\alpha$ y por lo tanto, V_α es una solución a la $\alpha-$EOCD. Finalmente, si $u' \in M_+(X)$ satisface $Tu' = u'$ entonces por el Lema 3.6 a), $u' \geq V_\alpha$ lo que completa la demostración de a).

b) Por la Proposición 2.12 y el Lema 2.13 a), existe $f_\alpha \in \mathbf{F}$ que satisface la ecuación (3.3). Iterando se obtiene

$$V_\alpha(x) = E_x^{f_\alpha}[\textstyle\sum_{n=0}^{N-1} \Lambda_\alpha^n C_\alpha(x_n, a_n)] + E_x^{f_\alpha}[\Lambda_\alpha^N V_\alpha(x_N)]$$

$$\geq V_\alpha^N(f_\alpha, x) \quad \forall N \geq 1, \ x \in X.$$

Haciendo $N \longrightarrow \infty$ se tiene $V_\alpha(x) \geq V_\alpha(f_\alpha, x)$. Puesto que la desigualdad inversa se cumple trivialmente, se sigue que $V_\alpha(x) = V_\alpha(f_\alpha, x) \ \forall x \in X$.

La segunda parte en b) se sigue del hecho de que para cualquier política estacionaria $f \in \mathbf{F}$,

$$V_\alpha(f,x) = C_\alpha(x,f) + \Delta_\alpha(x,f) \int V_\alpha(f,y) Q(dy \mid x, f).$$

c) Aplicando el Lema 3.6 b) se obtiene

$$V_\alpha(\pi^*,x) \leq V_\alpha(x) \ \ \forall x \in X,$$

por lo que π^* es $\alpha-$óptima.

d) Sea $u \in M_+(X)$ acotada tal que

$$u = Tu.$$

De a), se tiene que $u \geq V_\alpha$ y, puesto que se satisface (3.5), se sigue del Lema 3.6 b) que $u \leq V_\alpha$. Por lo tanto, $u = V_\alpha$. ■

Nótese que en la demostración de d), únicamente se usó el hecho de que cualquier función acotada en $M_+(X)$ satisface (3.5). Por lo anterior, si V_α satisface (3.5) entonces es la única solución de la $\alpha-$EOCD con esta propiedad.

3.3 Condiciones suficientes para (3.4)

Como puede verse en el Teorema 3.5, la condición (3.4) permite caracterizar a las políticas $\alpha-$óptimas. En esta sección suponemos que se satisfacen H2.7, H2.10 y H3.4 y deseamos dar condiciones suficientes para (3.4). Denotamos por A_0, A_1, A_2 y A_3 las siguientes condiciones:

A_0 : $C_\alpha(x,a)$ es acotada en $\mathbf{K}$.

A_1 : Existen dos números positivos m y β con $\beta\rho_\alpha < 1$, y una función medible no negativa $\omega : X \longrightarrow \mathbf{R}$ tal que $\forall (x,a) \in \mathbf{K}$,

1) $C_\alpha(x,a) \leq m\omega(x)$,

2) $\int \omega(y) Q(dy \mid x,a) \leq \beta\omega(x)$.

A_2 : $C(x) := \sum_{t=0}^{\infty} C_t(x) < \infty \;\; \forall x \in X$, donde $C_0(x) := \sup_{a\in A(x)} C_\alpha(x,a)$ y
$$C_t(x) := \sup_{a\in A(x)} \Delta_\alpha(x,a) \int C_{t-1}(y)Q(dy \mid x,a)$$
para $t = 1, 2, \ldots$.

A_3 : $\lim_{n\to\infty} E_x^\pi[\Lambda_\alpha^n V_\alpha(\pi', x_n)] = 0 \;\; \forall \pi,\; \pi' \in \Pi^0$ y $x \in X$.

Teorema 3.7. a) A_i implica A_{i+1} $(i = 0, 1, 2)$ y A_3 implica (3.4).

b) Si A_i se cumple para algún $i = 0, 1, 2, 3$, entonces una política π^* es óptima si y solo si $V_\alpha(\pi^*, \cdot)$ es una solución de la $\alpha-$EOCD.

Demostración. a) $A_0 \Longrightarrow A_1$. Es suficiente con tomar m cualquier cota superior de $C_\alpha(x,a)$ y $\omega(\cdot) \equiv 1$.

$A_1 \Longrightarrow A_2$. Por un argumento de inducción, se prueba que
$$C_t(x) \le (\beta\rho_\alpha)^t m\omega(x) \quad \forall x \in X,\; t = 0, 1, \ldots.$$
Por lo tanto,
$$C(x) \le \frac{m\omega(x)}{1-\beta\rho_\alpha} < \infty.$$
$A_2 \Longrightarrow A_3$. Sean $x \in X$ y $\pi \in \Pi^0$ arbitrarios. Probaremos primero que
$$V_\alpha(\pi, x) \le C(x). \tag{3.9}$$
Obsérvese que
$$\begin{aligned} E_x^\pi[\Lambda_\alpha^n C_0(x_n) \mid h_{n-1}, a_{n-1}] &= \Lambda_\alpha^{n-1}\Delta_\alpha(x_{n-1}, a_{n-1}) \textstyle\int C_0(y)Q(dy \mid x_{n-1}, a_{n-1}) \\ &\le \Lambda_\alpha^{n-1} C_1(x_{n-1}) \end{aligned}$$
y tomando esperanza se obtiene
$$E_x^\pi[\Lambda_\alpha^n C_0(x_n)] \le E_x^\pi[\Lambda_\alpha^{n-1} C_1(x_{n-1})].$$
De la misma forma, para $n \ge 1$,
$$\begin{aligned} E_x^\pi[\Lambda_\alpha^n C_0(x_n)] &\le E_x^\pi[\Lambda_\alpha^{n-1} C_1(x_{n-1})] \le \ldots \\ &\le E_x^\pi[\Delta_\alpha(x_0, a_0) C_{n-1}(x_1)] \\ &\le E_x^\pi[C_n(x_0)] = C_n(x). \end{aligned} \tag{3.10}$$

De (3.10) y el hecho que $C_\alpha(x_n, a_n) \leq C_0(x_n)$ se tiene que

$$E_x^\pi[\Lambda_\alpha^n C_\alpha(x_n, a_n)] \leq E_x^\pi[\Lambda_\alpha^n C_0(x_n)] \leq C_n(x),$$

lo cual implica (3.9). Ahora, si π y π' son dos políticas arbitrarias en Π^0 entonces, de (3.9),

$$E_x^\pi[\Lambda_\alpha^n V_\alpha(\pi', x_n)] \leq E_x^\pi[\Lambda_\alpha^n C(x_n)].$$

Además,

$$\begin{aligned} E_x^\pi[\Lambda_\alpha^n C(x_n) \mid h_{n-1}, a_{n-1}] &= \Lambda_\alpha^n \textstyle\int \sum_{t=0}^\infty C_t(y) Q(dy \mid x_{n-1}, a_{n-1}) \\ &= \Lambda_\alpha^{n-1} \textstyle\sum_{t=0}^\infty \Delta_\alpha(x_{n-1}, a_{n-1}) \int C_t(y) Q(dy \mid x_{n-1}, a_{n-1}) \\ &\leq \Lambda_\alpha^{n-1} \textstyle\sum_{t=0}^\infty C_{t+1}(x_{n-1}). \end{aligned}$$

Tomando esperanza E_x^π y procediendo inductivamente,

$$\begin{aligned} E_x^\pi[\Lambda_\alpha^n C(x_n)] &\leq E_x^\pi[\Lambda_\alpha^{n-1} \textstyle\sum_{t=0}^\infty C_{t+1}(x_{n-1})] \leq \cdots \\ &\leq E_x^\pi[\textstyle\sum_{t=0}^\infty C_{t+n}(x_0)] = \sum_{t=0}^\infty C_{t+n}(x), \end{aligned}$$

por lo que

$$E_x^\pi[\Lambda_\alpha^n V_\alpha(\pi', x_n)] \leq \sum_{t=n}^\infty C_t(x) \longrightarrow 0$$

cuando $n \to \infty$ puesto que $C(x)$ es finito.

b) Se sigue de a) y del Teorema 3.5. ■

3.4 Aproximaciones

En esta sección, consideramos otros tipos de aproximaciones a la función de costo V_α, diferentes a las "aproximaciones sucesivas" o iteración de valores que se utilizó en el Teorema 3.5. Suponemos que se cumplen H2.7, H2.10 y H3.4.

a) Aproximaciones con costos acotados.

Sea $C_\alpha^n(x,a)$ una sucesión de funciones no negativas, acotadas e inf-compactas en $\mathbf{K}$ tal que $C_\alpha^n(x,a) \uparrow C_\alpha(x,a)$ [por ejemplo, las funciones truncadas $C_\alpha^n(x,a) = \min(C_\alpha(x,a), n)$]. Para $x \in X$ y $\pi \in \Pi$, defínanse las correspondientes funciones de costo, por:

$$U_n(\pi, x) := E_x^\pi \sum_{k=0}^{\infty} e^{-\alpha T_k} C_\alpha^n(x_k, a_k) \tag{3.11}$$

y

$$U_n(x) := \inf_{\pi \in \Pi} U_n(\pi, x). \tag{3.12}$$

Para cada $v \in M_+(X)$, sea $T_n v$ la función en $M_+(X)$ definida por

$$T_n v(x) := \min_{a \in A(x)} \{C_\alpha^n(x,a) + \Delta_\alpha(x,a) \int v(y) Q(dy \mid x, a)\}. \tag{3.13}$$

Observación 3.8. Si C_α se reemplaza por C_α^n, entonces por el Teorema 3.5 d), la función $U_n(\cdot)$ es la única función acotada en $M_+(X)$ que satisface

$$U_n(x) = T_n U_n(x). \tag{3.14}$$

Proposición 3.9. La sucesión $\{U_n\}$ es monótona creciente y converge a V_α.

Demostración. Puesto que $C_\alpha^n \uparrow C_\alpha$, la sucesión $\{U_n\}$ es creciente y por lo tanto existe una función u en $M_+(X)$ tal que $U_n \uparrow u$. Haciendo $n \to \infty$ en (3.14), se sigue del Lema 2.13 b) que $u = Tu$ y por el Lema 3.6 a), $u \geq V_\alpha$. Por otra parte, $U_n \leq V_\alpha \; \forall n$, de donde se sigue que $u \leq V_\alpha$ y, por lo tanto, $u = V_\alpha$, es decir, $U_n \uparrow V_\alpha$. ■

b) Aproximaciones recursivas con costos acotados.

Sea T_n como en (3.13) y sea $\{u_n\}$ la sucesión definida por

$$u_0 \equiv 0$$

y

$$u_n := T_n u_{n-1} \;\text{ para } n \geq 1.$$

es decir,

$$u_n(x) = \min_{a \in A(x)} \{C_\alpha^n(x,a) + \Delta_\alpha(x,a) \int u_{n-1}(y) Q(dy \mid x, a)\}.$$

Proposición 3.10. La sucesión $\{u_n\}$ es monótona creciente y converge a V_α.

Demostración. Con argumentos análogos a los que se usaron en la demostración de la Proposición 3.9, se obtiene que existe una función v tal que $u_n \uparrow v$ y $v \geq V_\alpha$. Por otra parte, para cada $n = 0, 1, \ldots,$ $u_n(x) \leq V_n(\pi, x)$ $\forall \pi \in \Pi$, $x \in X$ (ver (3.2)) por lo que $u_n(x) \leq V_\alpha(x)$. Haciendo $n \to \infty$ se tiene que $v(x) \leq V_\alpha(x)$ y por lo tanto $u_n \uparrow V_\alpha$. ■

c) Iteración de políticas.

Sea $f_0 \in \mathbf{F}$ una política estacionaria tal que $\omega_0(x) := V_\alpha(f_0, x) < \infty$ $\forall x \in X$. Entonces

$$\omega_0(x) = C_\alpha(x, f_0) + \Delta_\alpha(x, f_0) \int \omega_0(y) Q(dy \mid x, f_0) \quad \forall x \in X \tag{3.15}$$

y, por el Lema 2.13 a), existe $f_1 \in \mathbf{F}$ tal que

$$T\omega_0(x) = C_\alpha(x, f_1) + \Delta_\alpha(x, f_1) \int \omega_0(y) Q(dy \mid x, f_1). \tag{3.16}$$

donde T es el operador en la Definición 3.1.

Sea $\omega_1(x) := V_\alpha(f_1, x)$. En general, dada $f_n \in \mathbf{F}$, sean $\omega_n(x) := V_\alpha(f_n, x)$ y $f_{n+1} \in \mathbf{F}$ tales que

$$T\omega_n(x) = C_\alpha(x, f_{n+1}) + \Delta_\alpha(x, f_{n+1}) \int \omega_n(y) Q(dy \mid x, f_{n+1}). \tag{3.17}$$

Proposición 3.11. Existe $\omega \in M_+(X)$ tal que $\omega_n \downarrow \omega$ y $T\omega = \omega$. Si además,

$$\lim_{n\to\infty} E_x^\pi \left[\Lambda_\alpha^n \omega(x_n)\right] = 0 \quad \forall x \in X,\ \pi \in \Pi, \tag{3.18}$$

entonces $\omega = V_\alpha$.

Demostración. Probaremos primero que $\{\omega_n\}$ es una sucesión decreciente. De (3.15) y (3.16) se sigue que

$$\omega_0(x) \geq C_\alpha(x, f_1) + \Delta_\alpha(x, f_1) \int \omega_0(y) Q(dy \mid x, f_1)$$

e iterando esta desigualdad se obtiene

$$\omega_0(x) \geq V_\alpha(f_1, x) = \omega_1(x).$$

Por un argumento similar se obtiene que para cada $n = 1, 2, ...$

$$\omega_n \geq T\omega_n \geq \omega_{n+1} \tag{3.19}$$

y, por lo tanto, existe una función $\omega \in M_+(X)$ tal que $\omega_n \downarrow \omega$. Puesto que $\omega_n \geq V_\alpha \; \forall n$,

$$\omega \geq V_\alpha. \tag{3.20}$$

Haciendo $n \to \infty$ en (3.19) y aplicando el Lema 3.3 de [16] se concluye que

$$\omega = T\omega.$$

Si ω satisface (3.18), entonces por el Lema 3.6 b), se sigue que $\omega \leq V_\alpha$ y por (3.20) se obtiene $\omega = V_\alpha$. ■

3.5 Comentarios y conclusiones

En este capítulo se obtuvieron los siguientes resultados relativos al problema de control óptimo en costo descontado:

1) Existencia de políticas estacionarias óptimas y caracterización de la función de costo V_α como la mínima solución (puntual) de la $\alpha-$EOCD. En el caso en que C_α es acotada en $\mathbf{K}$, V_α resulta ser la única solución acotada de la $\alpha-$EOCD.

2) Caracterización de las políticas óptimas como las políticas $\pi \in \Pi$ tales que $V_\alpha(\pi, \cdot)$ es solución de la $\alpha-$EOCD. Aquí es importante resaltar la importancia de la condición (3.4) y las condiciones A_0, A_1, y A_2 suficientes para (3.4). Una forma mas general de la condición A_1 aparece en Lippman [22], mientras que en [2], se considera una condición del tipo A_2.

3) Aproximaciones a la función de costo V_α por los métodos de iteración de valores, iteración de políticas y aproximaciones con costos acotados. Es importante notar el carácter monótono de la convergencia en cada caso lo que hace posible, mediante la elección de aproximaciones adecuadas, deducir propiedades de V_α. Específicamente, suponer que $\{w_n\}$ es una sucesión de funciones en X tal que $w_n \uparrow V_\alpha$, como es el caso de las funciones de iteración de valores y las funciones en las aproximaciones a) y b)

de la sección 3.4. Entonces, por ejemplo, si las funciones w_n $(n = 1, 2, ...)$ son s.c.i. o convexas o monótonas, se sigue, V_α es del mismo tipo, respectivamente.

En relación con 3), observamos que en [2] se consideran conjuntos de control $A(x)$ compactos, C_α es s.c.i. y la siguiente condición de continuidad débil en la ley de transición: Para cada $v : X \longrightarrow \mathbf{R}$ continua y acotada, la función

$$v'(x, a) := \int v(y)Q(dy \mid x, a)$$

es continua y acotada en $\mathbf{K}$. Bajo estas condiciones, las funciones de iteración de valores v_n resultan ser s.c.i. y, por lo tanto, V_α es s.c.i. Si los conjuntos $A(x)$ no son compactos, las funciones v_n no necesariamente resultan ser s.c.i.; véase [23].

Capítulo 4

Indice de costo promedio I

4.1 Introducción.

4.2 La desigualdad de optimalidad.

4.3 La ecuación de optimalidad.

4.4 Comentarios y conclusiones.

4.1 Introducción

Una forma de estudiar los PCSMs con índice de costo promedio (ver Definición 2.5), es mediante la aproximación con problemas con índice de costo descontado, como puede verse en [3], [4], [21], [22], [30], [32], [33] y [38]. Para modelos de control Markoviano, este método está relacionado con Teoremas Tauberianos. Siguiendo este enfoque, en este capítulo se dan condiciones para la existencia de soluciones a la ecuación de optimalidad así como para la existencia de políticas estacionarias óptimas. En el Capítulo 5 veremos otra forma de estudiar esta clase de problema.

4.2 La desigualdad de optimalidad

En todo este capítulo, $z \in X$ es un punto arbitrario pero fijo. Para $\alpha > 0$, sea $V_\alpha(\cdot)$ la función de costo óptimo $\alpha-$descontado (Definición 2.4) y definimos

$$h_\alpha(x) := V_\alpha(x) - V_\alpha(z), \quad x \in X. \tag{4.1}$$

Uno de nuestros principales objetivos es llegar a la *ecuación de optimalidad en costo promedio* (EOCP) (4.14), para lo cual empezamos por garantizar la existencia de soluciones de la *desigualdad de optimalidad* (4.3). Con tal fin se introducen las siguientes condiciones.

Hipótesis 4.1 (H4.1). Existen $N \in M_+(X)$ y constantes positivas K, γ y α_0 tales que $\forall\alpha \in (0, \alpha_0)$

a) $\alpha V_\alpha(z) \leq K$;

b) $|h_\alpha(x)| \leq \gamma N(x) \;\; \forall x \in X$;

c) Para cada $x \in X$, $a \longmapsto \int N(y)Q(dy \mid x, a)$ es continua y

$$\sup_{a\in A(x)} \int N(y)Q(dy \mid x, a) =: L_x < \infty;$$

d) $\sup_{a\in A(x)} \tau(x, a) =: M_x < \infty$ para cada $x \in X$. (recuérdese que $\tau(x, a)$ es una función estrictamente positiva -ver Proposición 2.11 d).)

Proposición 4.2. Si H2.7, H2.10 y H4.1 se cumplen, entonces existen una función medible h definida en X, una constante j^* y una política estacionaria $f^* \in \mathbf{F}$ tales que $\forall x \in X$,

$$|h(x)| \leq \gamma N(x) \tag{4.2}$$

y

$$\begin{aligned} h(x) &\geq \min_{a\in A(x)}\{C(x, a) - j^*\tau(x, a) + \textstyle\int h(y)Q(dy \mid x, a)\} \\ &= C(x, f^*) - j^*\tau(x, f^*) + \textstyle\int h(y)Q(dy \mid x, f^*). \end{aligned} \tag{4.3}$$

Para modelos Markovianos en tiempo discreto, (4.3) implica que f^* es CP-óptima como puede verse en [28]. Sin embargo, en el caso semi-Markoviano, no se tiene un resultado que permita afirmar que f^* es CP-óptima y, por lo tanto, (4.3) es un paso intermedio para obtener la EOCP (4.14).

Para la demostración de la Proposición 4.2 se necesita el siguiente lema en el cual se supone que se cumplen las hipótesis de la proposición.

Lema 4.3. a) Existen una constante j^* y una sucesión $\alpha(n) \downarrow 0$ tales que

$$\lim_{n\to\infty} \alpha(n) V_{\alpha(n)}(z) = j^*. \tag{4.4}$$

b) Para cada $x \in X$ existen a_x y a_n $(n = 1, 2, ...)$ en $A(x)$ tales que

$$V_{\alpha(n)}(x) = C_{\alpha(n)}(x, a_n) + \Delta_{\alpha(n)}(x, a_n) \int V_{\alpha(n)}(y) Q(dy \mid x, a_n). \tag{4.5}$$

y a_x es un punto de acumulación de $\{a_n\}$.

Demostración. a) Sea $j^* := \limsup_{\alpha\downarrow 0} \alpha V_\alpha(z)$. Por H4.1 a), j^* es finito y existe una sucesión $\alpha(n) \downarrow 0$ que satisface (4.4).

b) La existencia de a_n $(n = 1, 2, ...)$ que satisface (4.5) se sigue del Teorema 3.5 b). Para la demostración de la segunda parte en b), definimos las funciones $h, G_n : X \to \mathbf{R}$ $(n = 1, 2, ...)$ por

$$h(y) := \liminf_{n\to\infty} h_{\alpha(n)}(y) = \lim_{n\to\infty} G_n(y). \tag{4.6}$$

y

$$G_n(y) := \inf_{k\geq n} h_{\alpha(k)}(y). \tag{4.7}$$

Sea $x \in X$ fijo. Si definimos

$$W_\alpha(x, a, g) := C_\alpha(x, a) + \Delta_\alpha(x, a) \int g(y) Q(dy \mid x, a).$$

entonces de (2.6), (4.1) y (4.5), se sigue que

$$W_{\alpha(n)}(x, a_n, h_{\alpha(n)}) \;= h_{\alpha(n)}(x) + \alpha(n) V_{\alpha(n)}(z) \tau_{\alpha(n)}(x, a_n). \tag{4.8}$$

Para cualquier $\epsilon > 0$, por H3.1 a), d) y (4.6), existe $n(\epsilon)$ y una subsucesión $\{\alpha(n_i)\}$ de $\{\alpha(n)\}$ tal que $\forall n_i \geq n(\epsilon)$,

$$W_{\alpha(n_i)}(x, a_{n_i}, h_{\alpha(n_i)}) \leq h(x) + KM_x + \epsilon \tag{4.9}$$

y sumando $\Delta_{\alpha(n_i)}(x, a_{n_i}) \int \gamma N(y)Q(dy \mid x, a_{n_i})$ en ambos lados de (4.9), de **H3.1** c), (4.7) y el hecho que $0 < \Delta_\alpha(x, a) < 1$, se obtiene

$$W_{\alpha(n_i)}(x, a_{n_i}, G_{n_i} + \gamma N) \leq h(x) + \gamma L_x + K M_x + \epsilon.$$

Puesto que $C_{\alpha(n_i)}(x, a) \uparrow C(x, a)$, $\Delta_{\alpha(n_i)}(x, a) \uparrow 1$ y $G_{n_i}(y) \uparrow h(y)$ cuando $n_i \longrightarrow \infty$, los conjuntos

$$D_i := \{a \in A(x) \mid W_{\alpha(n_i)}(x, a, G_{n_i} + \gamma N) \leq r\},$$

con $r := h(x) + \gamma L_x + K M_x + \epsilon$, forman una sucesión decreciente de conjuntos compactos no vacíos que convergen al conjunto compacto no vacío

$$D(x) := \{a \in A(x) \mid C(x, a) + \int (h + \gamma N)(y)Q(dy \mid x, a) \leq r\}.$$

Por lo tanto, existe $a_x \in A(x)$ y una subsucesión de $\{n_i\}$ que denotamos de nuevo por $\{n_i\}$ tal que $a_{n_i} \longrightarrow a_x$. ■

Demostración de la Proposición 4.2. Sean $x \in X$ un estado arbitrario pero fijo, $\epsilon > 0$ y $\{\alpha(n_i)\} =: \{\beta(i)\}$ y $\{a_{n_i}\} =: \{b_i\}$ las sucesiones en la demostración del Lema 4.3. De (4.8) y el hecho que $\tau_\alpha(x, a) \leq \tau(x, a)$ se tiene que

$$W_{\beta(i)}(x, b_i, h_{\beta(i)}) \leq h_{\beta(i)}(x) + \beta(i)V_{\beta(i)}(z)\tau(x, b_i).$$

Por (4.4), (4.6), el Lema 4.3 y la continuidad de $a \longmapsto \tau(x, a)$ (ver Proposición 2.12), existe $i(\epsilon)$ tal que $\forall i \geq i(\epsilon)$,

$$W_{\beta(i)}(x, b_i, G_i) \leq h(x) + j^* \tau(x, a_x) + \epsilon. \tag{4.10}$$

Sea $l > i(\epsilon)$ un número entero arbitrario y sea $i \geq l$. Entonces, de (4.10),

$$W_{\beta(i)}(x, b_i, G_l + \gamma N) \quad \leq h(x) + j^* \tau(x, a_x) + \gamma \int N(y)Q(dy \mid x, b_i) + \epsilon. \tag{4.11}$$

Tomando $\liminf_{i \to \infty}$ en (4.11), de **H4.1** c) se obtiene:

$$\begin{aligned} C(x, a_x) + \int (G_l + \gamma N)(y)Q(dy \mid x, a_x) \quad &\leq h(x) + j^* \tau(x, a_x) \\ &\quad + \int \gamma N(y)Q(dy \mid x, a_x) + \epsilon. \end{aligned} \tag{4.12}$$

y por el Teorema de la Convergencia Monótona (puesto que $G_n \uparrow h$),

$$C(x, a_x) + \int h(y)Q(dy \mid x, a_x) \leq h(x) + j^*\tau(x, a_x) + \epsilon. \tag{4.13}$$

Puesto que ϵ es arbitrario, la desigualdad en (4.3) se sigue de (4.13). Por último, la existencia de $f^* \in \mathbf{F}$ tal que se satisface la igualdad en (4.3), se sigue del Lema 2.13 a). ∎

4.3 La ecuación de optimalidad

En esta sección, se establecen condiciones que garantizan la existencia de soluciones de la EOCP (4.14) y la existencia de políticas estacionarias CP-óptimas.

Sea $\{\alpha(n)\}$, la sucesión en el Lema 4.3.

Hipótesis 4.4 (H4.4). La sucesión $\{h_{\alpha(n)}\}$ es equicontinua.

Teorema 4.5. Supóngase que las hipótesis del Teorema 4.2, así como H4.4 se cumplen. Entonces existen una función continua h, una constante $j^* \geq 0$ y $f \in \mathbf{F}$ tal que:

a) $\forall x \in X$:

$$\begin{aligned} h(x) &= \min_{a \in A(x)} \{C(x,a) - j^*\tau(x,a) + \textstyle\int h(y)Q(dy \mid x,a)\} \\ &= C(x, f^*) - j^*\tau(x, f^*) + \textstyle\int h(y)Q(dy \mid x, f^*). \end{aligned} \tag{4.14}$$

b) Si

$$\lim_{n \to \infty} \frac{E_x^\pi N(x_n)}{n} = 0 \quad \forall x \in X, \ \pi \in \Pi,$$

entonces f^* es CP-óptima y $J(f^*, x) = J(x) = j^* \quad \forall x \in X$.

Demostración. a) Por el Teorema de Ascoli (ver [34] p. 179), junto con H4.1 b) y H4.4, existe una subsucesión de $\{\alpha(n)\}$ que denotaremos de nuevo por $\{\alpha(n)\}$ y una función continua h definida en X tal que

$$\lim_{n \to \infty} h_{\alpha(n)}(x) = h(x).$$

De la $\alpha-$EOCD y (4.1), $\forall x \in X$ y $a \in A(x)$

$$\begin{aligned} h_{\alpha(n)}(x) \quad & \leq C_{\alpha(n)}(x,a) + \Delta_{\alpha(n)}(x,a) \textstyle\int h_{\alpha(n)}(y)Q(dy \mid x,a) \\ & -\alpha(n)V_{\alpha(n)}(z)\tau_{\alpha(n)}(x,a). \end{aligned} \tag{4.15}$$

Tomando $\lim_{n\to\infty}$ en ambos lados de (4.15) y usando el Teorema de la Convergencia Dominada se obtiene

$$h(x) \leq C(x,a) - j^*\tau(x,a) + \int h(y)Q(dy \mid x,a) \quad \forall x \in X, \quad a \in A(x)$$

y, por lo tanto,

$$h(x) \leq \min_{a\in A(x)}\{C(x,a) - j^*\tau(x,a) + \int h(y)Q(dy \mid x,a)\}.$$

La desigualdad $\geq$ se obtiene de la Proposición 4.2. Por último, la segunda igualdad en (4.14) se sigue del Lema 2.13 a).

b) Sean $x \in X$ y $\pi \in \Pi$ arbitrarios. Entonces, para $t = 0, 1, \ldots$, se sigue de (2.4) que

$$\begin{aligned} E_x^\pi[h(x_{t+1}) \mid h_t, a_t)] \quad & = \textstyle\int h(y)Q(dy \mid x_t, a_t) \\ & = C(x_t,a_t) - j^*\tau(x_t,a_t) + \textstyle\int h(y)Q(dy \mid x_t,a_t) \\ & \quad -C(x_t,a_t) + j^*\tau(x_t,a_t) \\ & \geq h(x_t) - C(x_t,a_t) + j^*\tau(x_t,a_t) \quad \text{[por (4.14)]}. \end{aligned}$$

Tomando esperanza E_x^π y sumando de $t = 0$ a $t = n-1$ se obtiene

$$j^* E_x^\pi \sum_{t=0}^{n-1} \tau(x_t,a_t) \leq E_x^\pi h(x_n) - h(x) + E_x^\pi \sum_{t=0}^{n-1} C(x_t,a_t).$$

que con H3.1 b) lleva a

$$j^* + \frac{h(x)}{E_x^\pi \sum_{t=0}^{n-1} \tau(x_t,a_t)} \leq \frac{E_x^\pi \sum_{t=0}^{n-1} C(x_t,a_t)}{E_x^\pi \sum_{t=0}^{n-1} \tau(x_t,a_t)} + \frac{\gamma E_x^\pi N(x_n)}{E_x^\pi \sum_{t=0}^{n-1} \tau(x_t,a_t)}$$

y tomando $\limsup_{n\to\infty}$ se obtiene

$$j^* \leq J(\pi, x). \tag{4.16}$$

De la segunda igualdad en (4.14), $j^* = J(f^*, x)$ lo cual combinado con (4.16) demuestra la CP-optimalidad de f^*. ■

4.4 Comentarios y conclusiones

En este capítulo, se establecieron condiciones para la existencia de políticas estacionarias CP-óptimas. Estas condiciones son más generales que las que se usan en [3], donde la familia de funciones $\{h_\alpha\}$ (ver (4.1)) es equicontinua y uniformemente acotada y, al igual que en [2], los conjuntos de control $A(x)$ son compactos y Q satisface una condición de continuidad débil (ver sección 3.5).

Nótese que H4.4 se usó únicamente para garantizar la existencia de una subsucesión convergente. Si X es numerable, H4.4 no es necesaria pues en este caso, la subsucesión convergente se construye por el método de diagonalización (ver [34], p. 167).

Sennott [38], introduce condiciones de este tipo para MCSM con espacio de estados numerable. Demuestra la existencia de soluciones de la desigualdad de optimalidad y, utilizando un Teorema Tauberiano, obtiene la existencia de políticas estacionarias óptimas para el índice de funcionamiento definido por

$$\Phi(\pi, x) := \limsup_{t\to\infty} \frac{E_x^\pi Z(t)}{t}$$

donde $Z(t)$ es el costo hasta el tiempo t que se define de la siguiente manera: Si $T_n \leq t < T_{n+1}$,

$$Z(t) := \sum_{k=0}^{n} D(x_k, a_k) + \sum_{k=0}^{n-1} d(x_k, a_k)\delta_{k+1} + d(x_n, a_n)(t - T_n).$$

Una condición adicional en [38] es la igualdad de los índices $J(f, x) = \Phi(f, x)$. Un trabajo ineresante sería el extender estos resultados al contexto de esta tesis.

Capítulo 5

Indice de costo promedio II

4.1 Introducción.

4.2 Condiciones de optimalidad.

4.3 El modelo Markoviano asociado.

4.4 Resultados de optimalidad.

4.5 Comentarios y conclusiones.

5.1 Introducción

Otra forma de estudiar el problema de control semi-Markoviano con índice de costo promedio es aplicando una transformación al MCSM mediante la cual, se obtiene un modelo de control Markoviano (MCM) del cual es posible deducir resultados del modelo original. En el caso de MCSM con espacio de estados numerable, este método se ha utilizado en [5], [6], [30], [37], [40], [41]. Cabe mencionar que dicha transformación también se ha aplicado a MCMs, en cuyo caso los procesos resultantes son aperiódicos (ver [41] y [42]). En este capítulo, se establecen condiciones del tipo de condiciones introducidas en [8], [9] y [10] para MCMs y se obtienen resultados relativos a la existencia de políticas óptimas para el MCSM y el MCM asociado.

5.2 Condiciones de optimalidad

En todo este capítulo suponemos que (el espacio de Borel) X es localmente compacto. También usaremos las siguientes hipótesis.

Hipótesis 5.1 (H5.1). a) $\sup_{\mathbf{K}} \tau(x,a) =: M < \infty$;

b) Existe una función medible $v : X \longrightarrow [1,\infty)$ y $\gamma \in \mathbf{R}$ tal que para cada $x \in X$,

$$\sup_{a \in A(x)} C(x,a) \leq \gamma v(x)$$

y

$$a \longmapsto \int v(y) Q(dy \mid x,a) \quad \text{es continua en } A(x).$$

Hipótesis 5.2 (H5.2). Para cada $f \in \mathbf{F}$, la cadena de Markov con probabilidad de transición $Q(\cdot \mid x, f)$ es Harris-recurrente positiva (ver Definición A.7), con medida de probabilidad invariante q_f.

Hipótesis 5.3 (H5.3). Existe una medida de probabilidad ν en X y $0 < \alpha < 1$ para los cuales se cumple: Para cada $f \in \mathbf{F}$, existe una función ϕ_f en $M_+(X)$ tal que para cada $x \in X$ y $B \in \mathcal{B}(X)$:

a) $Q(B \mid x, f) \geq \phi_f(x)\nu(B)$;

b) $\int v(y) Q(dy \mid x, f) \leq \phi_f(x) \|\nu\|_v + \alpha v(x)$, con $\|\nu\|_v := \int v(y)\nu(dy) < \infty$;

c) $\inf_{\mathbf{F}} \int \phi_f(y)\nu(dy) > 0$.

En [9] se considera un MCM el cual satisface H5.1-H5.3 y representa sistemas que aparecen en algunas areas de aplicación como teoría de inventarios y teoría de colas. Basándonos en ese trabajo, presentamos un MCSM para una clase de sistemas de producción-inventario.

Ejemplo. Considérese un sistema de producción-inventario en el que la demanda de un producto (la cual se atiende inmediatamente), ocurre en ciertas *épocas* $T_1, T_2, \ldots$ tales que $\delta_n := T_n - T_{n-1}$ ($n = 1, 2, \ldots, T_0 = 0$) es una variable aleatoria. Sea ξ_n ($n = 1, 2, \ldots$) la demanda en la época T_n y sean x_n y a_n respectivamente, el nivel de inventario y la cantidad ordenada (o producida) inmediatamente después que ocurre la demanda ξ_n. Si x_0 y a_0 representan el nivel de inventario y la cantidad ordenada en el tiempo T_0, entonces la dinámica del sistema está dada por

$$x_{n+1} = (x_n + a_n - \xi_{n+1})^+, \quad n = 0, 1, ..., \ x_0 \text{ dado},$$

donde $X = [0, \infty)$. En la época T_n, se produce un costo inmediato $D(x_n, a_n)$ y por el tiempo que transcurre hasta la siguiente época de demanda, se produce un costo cuya razón es $d(x_n, a_n)$.

A continuación se dan condiciones bajo las cuales, el sistema descrito puede representarse por un MCSM que satisface H2.7, H2.10, H5.1-H5.3.

Condición 1. a) $\{\xi_n\}_{n=1}^{\infty}$ es una sucesión de variables aleatorias no negativas i.i.d.

b) $A(x) = A$ para todo $x \in X$ donde A es un subconjunto compacto del intervalo $[0, \theta]$ para algún número $\theta > 0$.

c) La variable aleatoria $\zeta = \theta - \xi$, donde ξ denota la variable aleatoria genérica con la distribución de ξ_1, satisface $E(\zeta) < 0$.

d) ξ tiene densidad continua y acotada.

La condición en c) significa que en cada época, la cantidad que se produce es menor que el valor esperado de la demanda. Por otra parte, puesto que para $z \geq 0$,

$$g(z) := E(e^{z\zeta}) = e^{z\theta}E(e^{-z\xi}) \leq e^{z\theta},$$

$g(0) = 1$ y $g'(0) = E(\zeta) < 1$, se sigue que existe un número $q > 0$ tal que $E(e^{q\zeta}) < 1$.

Condición 2. Las funciones de costo $D(x, a)$ y $d(x, a)$ son funciones medibles no negativas, $D(x, \cdot)$ y $d(x, \cdot)$ son s.c.i. en $A(x)$ para todo $x \in X$ y

$$\sup_{A} C(x, a) \leq \gamma v(x) \quad \text{para algún } \gamma > 0.$$

con $v(x) := e^{qx}$.

Condición 3. Dados $x_n = x$ y $a_n = a$, la función de distribución de δ_{n+1} es $F(\cdot \mid x, a)$ la cual satisface H2.7, H2.10 c) y H5.1 a).

La ley de transición Q está dada por:

$$Q(\{0\} \mid x, a) := P(x + a - \xi \leq 0)$$

y

$$Q(B \mid x, a) := P(x + a - \xi \in B) \text{ para } B \in \mathcal{B}(0, \infty).$$

El sistema de producción-inventario puede entonces representarse mediante el MCSM $(X, A, \{A(x)\}, Q, F, D, d)$. De las propiedades de F y con argumentos análogos a los utilizados en [9], se demuestra que si se cumplen las Condiciones 1-3, el MCSM satisface H2.7, H2.10, H5.1-H5.3.

Denotamos por L_v^∞ el espacio lineal normado de las funciones medibles $u : X \longrightarrow \mathbf{R}$

tales que

$$\|u\|_v := \sup_{x \in X} \frac{|u(x)|}{v(x)} < \infty.$$

La siguiente proposición se sigue directamente del Lema 3.4 en [8].

Proposición 5.4. Si H5.2 y H5.3 se cumplen, entonces existen constantes $k \geq 0$ y $0 \leq \eta < 1$ tales que para cada $g \in L_v^\infty$ y $x \in X$,

$$\sup_{f \in \mathbf{F}} \left| \int g(y) Q^n(dy \mid x, f) - \int g(y) q_f(dy) \right| \leq \|g\|_v k v(x) \eta^n. \tag{5.1}$$

Una consecuencia inmediata de la proposición anterior es que para cada $f \in \mathbf{F}$, el costo promedio $J(f, \cdot)$ es constante:

$$J(f, x) = \lim_{n \to \infty} \frac{E_x^f \sum_{k=0}^n C(x_k, a_k)}{E_x^f \sum_{k=0}^n \tau(x_k, a_k)} = \frac{\int C(y, f) q_f(dy)}{\int \tau(y, f) q_f(dy)}. \tag{5.2}$$

Lema 5.5. Si H5.1-H5.3 se cumplen, entonces:

a) Para cada $f \in \mathbf{F}$, existen $J_f \in \mathbf{R}$ y $h_f \in L_v^\infty$ tales que

$$h_f(x) = C(x, f) - J_f \tau(x, f) + \int h_f(y) Q(dy \mid x, f) \quad \forall x \in X. \tag{5.3}$$

b) Si la pareja (J'_f, h'_f) con $J'_f \in \mathbf{R}$ y $h'_f \in L_v^\infty$ es también una solución de la ecuación (5.3), es decir,

$$h'_f(x) = C(x, f) - J'_f \tau(x, f) + \int h'_f(y) Q(dy \mid x, f) \quad \forall x \in X. \tag{5.4}$$

entonces $J'_f = J_f$ y $h'_f(x) = h_f(x) + k'$ $\forall x \in X$ con k' una constante; es decir, la función h_f en (5.3) es única excepto por una constante aditiva y J_f es única.

Demostración. a) Sean

$$J_f := \frac{\int C(y,f)q_f(dy)}{\int \tau(y,f)q_f(dy)}$$

y

$$h_f(x) := \sum_{n=0}^{\infty} E_x^f\left[C(x_n,f) - J_f\tau(x_n,f)\right].$$

Probaremos primero que h_f pertenece a L_v^∞ : Para cada $n = 0, 1, \ldots$

$$\begin{aligned}\left|E_x^f[C(x_n,f) - J_f\tau(x_n,f)]\right| &= \left| \int C(y,f)Q^n(dy \mid x,f) \right. \\ &\left. -\frac{\int C(y,f)q_f(dy)}{\int \tau(y,f)q_f(dy)} \int \tau(y,f)Q^n(dy \mid x,f) \right| \\ &\leq \left|\int C(y,f)Q^n(dy \mid x,f) - \int C(y,f)q_f(dy)\right| \\ &+J_f \left|\int \tau(y,f)Q^n(dy \mid x,f) - \int \tau(y,f)q_f(dy)\right| \\ &\leq M_f v(x)\eta^n\end{aligned}$$

donde la última desigualdad se sigue de (5.1), con M_f una constante, por lo que $h_f \in L_v^\infty$. (5.3) se sigue de la definición de h_f y de la propiedad de Markov.

b) Integrando con respecto a q_f en ambos miembros de la ecuación (5.4) y usando el hecho de que q_f es invariante se obtiene que $J'_f = J_f$. Por otra parte, restando (5.4) de (5.3) se tiene que la función $u \in L_v^\infty$ definida por $u(x) := h_f(x) - h'_f(x)$ es una función *f-armónica*, es decir,

$$u(x) = \int u(y)Q(dy \mid x,f)$$

y, como se demuestra en [14], u es una función constante. ■

5.3 El modelo Markoviano asociado

En esta sección suponemos que se cumple la condición de regularidad H2.7. Recordamos que $\inf_{\mathbf{K}} \tau(x,a) > 0$ (ver Proposición 2.11 a)).

Definición 5.6. Sea τ tal que $0 < \tau < \inf_{\mathbf{K}} \tau(x,a)$. Se definen la función $\hat{C} : \mathbf{K} \to \mathbf{R}$ y el kérnel estocástico $\hat{Q}$ en X dado $\mathbf{K}$ por

$$\hat{C}(x,a) := \frac{C(x,a)}{\tau(x,a)} \tag{5.5}$$

y

$$\hat{Q}(B \mid x,a) := \frac{\tau}{\tau(x,a)} Q(B \mid x,a) + (1 - \frac{\tau}{\tau(x,a)})\delta_x(B) \tag{5.6}$$

donde $\delta_x(\cdot)$ es la medida de Dirac concentrada en x.

El modelo $(X, A, \{A(x) : x \in X\}, \hat{Q}, \hat{C})$ es un *modelo de control Markoviano asociado* al MCSM.

En los siguientes lemas se demuestra que el modelo de control Markoviano asociado satisface condiciones análogas al MCSM.

Lema 5.7. Si H2.10 y H5.1 a) se cumplen, entonces, $\hat{C}$ es inf-compacta en $\mathbf{K}$ y $\hat{Q}$ es fuertemente continua en A.

Demostración. Sea $x \in X$ fijo. Puesto que $C(x,\cdot)$ es s.c.i. en $A(x)$ y $\tau(x,\cdot)$ es continua en $A(x)$, vemos que $\hat{C}(x,\cdot)$ es s.c.i. en $A(x)$ por lo que para cada $r \in \mathbf{R}$, el conjunto

$$E_r := \{a \in A(x) : \hat{C}(x,a) \leq r\}$$

es cerrado. Como el conjunto

$$D_r := \{a \in A(x) : C(x,a) \leq rM\}$$

es compacto y $E_r \subset D_r$ (por H5.1 a)), se tiene que E_r es compacto lo que demuestra la primera parte.

Considérese de nuevo $x \in X$ fijo. Sea u una función medible y acotada en X y sea u' la función definida en $A(x)$ por

$$u'(a) := \int u(y)\hat{Q}(dy \mid x,a).$$

De (5.6) se obtiene

$$u'(a) = \frac{\tau}{\tau(x,a)} \int u(y)Q(dy \mid x,a) + u(x)(1 - \frac{\tau}{\tau(x,a)})$$

y puesto que $\tau(x,\cdot)$ es continua en $A(x)$ y Q es fuertemente continua en A (por H2.10 b)), se tiene que u' es medible y continua en $A(x)$, lo que completa la demostración. ∎

Lema 5.8. Supóngase que H5.1 a) y H5.3 se cumplen y, para $f \in \mathbf{F}$, sea $\hat{\phi}_f(x) := \frac{\tau}{\tau(x,f)}\phi_f(x)$. Entonces $\forall B \in \mathcal{B}(X)$, $x \in X$:

a) $\hat{Q}(B \mid x, f) \geq \hat{\phi}_f(x)\nu(B)$;

b)

$$\int v(y)\hat{Q}(dy \mid x, f) \quad \leq \hat{\phi}_f(x) \, \|\nu\|_v + \hat{\alpha} v(x),$$

donde $0 < \hat{\alpha} < 1$ es la constante $\hat{\alpha} := (1 - \frac{\tau}{M}(1-\alpha))$.

c) $\inf_{\mathbf{F}} \int \hat{\phi}_f(y)\nu(dy) > 0$.

Demostración. Las partes a) y c) se siguen directamente de H5.3 a), c) y (5.6). Por otra parte, definiendo $T(f,x)) := \tau/\tau(x,f)$, se tiene:

$$\begin{aligned}
\int v(y)\hat{Q}(dy \mid x, f) &= T(f,x) \int v(y)Q(dy \mid x, f) + [1 - T(f,x)]v(x) \\
&\leq T(f,x)\phi_f(x) \, \|\nu\|_v + T(f,x)\alpha v(x) + [1 - T(f,x)]v(x) \\
&\leq \hat{\phi}_f(x) \, \|\nu\|_v + v(x)[1 - T(f,x)(1-\alpha)] \\
&\leq \hat{\phi}_f(x) \, \|\nu\|_v + \hat{\alpha} v(x),
\end{aligned}$$

lo que completa la demostración. ∎

A continuación se presentan algunas propiedades del kérnel $\hat{Q}$.

Observación 5.9. a) Para una política estacionaria $f \in \mathbf{F}$, sea ψ_f una medida de irreducibilidad (ver Definición A.1) para la cadena de Markov con probabilidad de transición $Q(\cdot \mid x, f)$. De (5.6) puede verse que $\forall x \in X$, $B \in \mathcal{B}(X)$

$$\hat{Q}(B \mid x, f) \geq \frac{\tau}{\tau(x,f)} Q(B \mid x, f) \geq \frac{\tau}{M} Q(B \mid x, f)$$

por lo que la cadena de Markov con probabilidad de transición $\hat{Q}(\cdot \mid x, f)$ es también ψ_f-irreducible. Además, puesto que para cada $x \in X$,

$$\hat{Q}(\{x\} \mid x, f) \geq (1 - \frac{\tau}{\rho})$$

(ρ es la constante definida en la Proposición 2.11 a)), la cadena resulta ser aperiódica (ver Nota A.5).

b) Si C_f es un conjunto petite (ver Definición A.8) con respecto a $Q(\cdot \mid x, f)$ entonces del Lema 2.1 en [7] y del hecho que $\forall n \in \mathbf{N}$

$$\hat{Q}^n(B \mid x, f) \geq (\frac{\tau}{M})^n Q^n(B \mid x, f)$$

se sigue que C_f es un conjunto petite con respecto al kernel $\hat{Q}(\cdot \mid x, f)$.

Lema 5.10. Si H5.1 a), H5.2 y H5.3 se cumplen, entonces para cada $f \in \mathbf{F}$, la cadena de Markov con probabilidad de transición $\hat{Q}(\cdot \mid x, f)$ es Harris-recurrente positiva.

Demostración. Primero observamos que por el Teorema E en [19], la cadena con probabilidad de transición $Q(\cdot \mid x, f)$ es aperiódica y puesto que es irreducible, por el Lema 3.4 en [8] y el Teorema 16.0.1 en [25], existen un conjunto petite C_f, una función medible $V_f : X \longrightarrow [1, \infty)$ y constantes positivas $\beta_f < 1$ y b_f tales que

$$\int V_f(y) Q(dy \mid x, f) \leq \beta_f V_f(x) + b_f 1_{C_f}(x). \tag{5.7}$$

De (5.7) y la definición de $\hat{Q}$ se obtiene

$$\begin{aligned}
\int V_f(y)\hat{Q}(dy \mid x, f) &\leq T(f, x) \int V_f(y) Q(dy \mid x, f) + (1 - T(f, x))V_f(x) \\
&\leq T(f, x)\beta_f V_f(x) + T(f, x) b_f 1_{C_f}(x) + (1 - T(f, x))V_f(x) \\
&= (1 - T(f, x)(1 - \beta_f))V_f(x) + T(f, x) b_f 1_{C_f}(x) \\
&\leq \beta'_f V_f(x) + b'_f 1_{C_f}(x),
\end{aligned}$$

con $\beta'_f := 1 - (\tau/M)(1 - \beta_f) < 1$ y $b'_f := (\tau/\rho) b_f$. Por lo tanto, de la Proposición A.10 se sigue que la cadena es Harris-recurrente positiva. ∎

5.4 Resultados de optimalidad

En esta sección se demuestran resultados relativos a la existencia de políticas estacionarias CP-óptimas para el MCSM y el MCM asociado.

Teorema 5.11. Si H2.10, H5.1-H5.3 se cumplen, entonces existen $j^* \in \mathbf{R}$, $\hat{h} \in L_v^\infty$ y $f \in \mathbf{F}$ tales que

a) $\forall x \in X$,

$$\begin{aligned} \hat{h}(x) &\geq \min_{a\in A(x)}\{\hat{C}(x,a) - j^* + \textstyle\int \hat{h}(y)\hat{Q}(dy \mid x, a)\} \\ &= \hat{C}(x,f) - j^* + \textstyle\int \hat{h}(y)\hat{Q}(dy \mid x, f)\} \end{aligned} \tag{5.8}$$

b)

$$\hat{h}(x) = \min_{a\in A(x)}\{\hat{C}(x,a) - j^* + \int \hat{h}(y)\hat{Q}(dy \mid x,a)\} \quad \hat{q}_f - c.d. \tag{5.9}$$

c) $j^* = \hat{J}(f,x) = \inf_{\pi\in\hat{\Pi}} \hat{J}(\pi,x)$, donde

$$\hat{J}(\pi,x) := \limsup_{n\to\infty} \frac{E_x^\pi[\sum_{k=0}^{n-1} \hat{C}(x_k,a_k)]}{n};$$

es decir, f es CP-óptima para el MCM asociado.

Demostración. De los resultados en la sección anterior, se tiene que el MCM asociado satisface las hipótesis del Teorema 2.6 en [8] con lo que se tiene la prueba de a) y c). La demostración de b), se hace usando los mismos argumentos en la demostración del Teorema 5.3 en [17]. ■

Ahora usaremos el Teorema 5.11 sobre el MCM asociado para deducir resultados sobre el MCSM.

Teorema 5.12. Supóngase que H2.10, H5.1-H5.3 se cumplen. Entonces existen $h \in L_v^\infty$, $j^* \geq 0$ y $f \in \mathbf{F}$ tal que:

a) $\forall x \in X$.

$$\begin{aligned} h(x) &\geq \min_{a\in A(x)}\{C(x,a) - j^*\tau(x,a) + \textstyle\int h(y)Q(dy \mid x,a)\} \\ &= C(x,f) - j^*\tau(x,f) + \textstyle\int h(y)Q(dy \mid x,f)\ . \end{aligned}$$

b)

$$h(x) = \min_{a \in A(x)} \{C(x,a) - j^* \tau(x,a) + \int h(y) Q(dy \mid x,a)\} \quad q_f - c.d.$$

c) $j^* = J(f,x) = \inf_{g \in \mathbf{F}} J(g,x)$, es decir, f es CP-óptima en la clase de las políticas estacionarias $\mathbf{F}$.

Demostración. Sean j^* y $\hat{h}$ y f como en el Teorema 5.11. Entonces, $\forall x \in X$, se sigue de (5.8) que

$$\hat{h}(x) \geq \frac{C(x,f)}{\tau(x,f)} - j^* + \frac{1}{\tau(x,f)} \int \tau \hat{h}(y) Q(dy \mid x,f) + (1 - \frac{\tau}{\tau(x,f)}) \hat{h}(x). \tag{5.10}$$

de donde se obtiene

$$\tau \hat{h}(x) \geq C(x,f) - j^* \tau(x,f) + \int \tau h(y) Q(dy \mid x,f). \tag{5.11}$$

Haciendo $h(\cdot) = \tau \hat{h}(\cdot)$, de (5.11) se sigue la parte a). De la misma forma se demuestra b), observando que q_f es absolutamente continua con respecto a $\hat{q}_f$.

Para probar c), supóngase que la política f no es CP-óptima entre las políticas estacionarias, es decir, existe $g \in \mathbf{F}$ tal que para alguna $x \in X$ (y entonces para toda x),

$$J(g,x) < J(f,x). \tag{5.12}$$

Por el Lema 5.5 a), existen $J_g \in \mathbf{R}$ y $h_g \in L_v^\infty$ tal que $\forall x \in X$,

$$h_g(x) = C(x,g) - J_g \tau(x,g) + \int h_g(y) Q(dy \mid x,g).$$

Entonces $J_g = J(g,x)$ y de (5.6) se tiene que

$$\frac{h_g(x)}{\tau} = \hat{C}(x,g) - J_g + \int \frac{h_g(y)}{\tau} \hat{Q}(dy \mid x,g).$$

Integrando con respecto a $\hat{q}_g$, resulta

$$J_g = \int \hat{C}(x,g) \hat{q}_f(dy) = \hat{J}(g,x)$$

y en la misma forma $J_f = \hat{J}(f,x) = J(f,x)$. Pero por el Teorema 5.11 c), $\hat{J}(f,x) \leq \hat{J}(g,x)$ $\forall x \in X$ lo que contradice (5.12) y por lo tanto

$$J(f,x) = \min_{g \in \mathbf{F}} J(g,x) \quad \forall x \in X. \blacksquare$$

5.5 Comentarios y conclusiones

En este capítulo se establecieron condiciones en el MCSM bajo las cuales se obtiene la existencia de una política estacionaria f CP-óptima *entre las políticas estacionarias*. Asimismo, se demostró que la política f es CP-óptima para el problema de control Markoviano asociado. Nótese que una solución de la ecuación de optimalidad para el MCM asociado, induce una solución de la ecuación de optimalidad para el MCSM lo que a su vez, garantiza la existencia de políticas estacionarias CP-óptimas.

Capítulo 6

Conclusiones y problemas abiertos

6.1 Introducción

Como se mencionó en el Capítulo 1, el objetivo de este trabajo consistía en extender al contexto semi-Markoviano algunos resultados sobre modelos de control Markoviano. En este capítulo presentamos las conclusiones del trabajo así como algunos problemas abiertos.

6.2 Conclusiones

Para el problema de control semi-Markoviano con índice de costo descontado, suponiendo que se satisfacen H2.7, H2.10 y H3.4 se tienen los siguientes resultados:

1) Existencia de políticas estacionarias $\alpha-$óptimas y caracterización de la función de costo $\alpha-$óptimo como la solución mínima (puntual) de la $\alpha-$EOCD.

2) Caracterización, bajo condiciones adecuadas de las políticas $\alpha-$óptimas como las políticas π tales que $V_\alpha(\pi,\cdot)$ es una solución de la $\alpha-$EOCD.

3) Aproximaciones a la función de costo V_α por los métodos de iteración de valores, iteración de políticas y aproximaciones con costos acotados.

Para el problema de control semi-Markoviano con índice de costo promedio, se concluye lo siguiente:

1) Las condiciones H2.7, H2.10, H4.1 y H4.4 implican la existencia de soluciones de la ecuación de optimalidad en costo promedio, así como la existencia de políticas estacionarias CP-óptimas.

2) Las condiciones H2.7. H2.10, H5.1-H5.3, implican la existencia de soluciones a la desigualdad de optimalidad y la existencia de políticas estacionarias óptimas en la familia de las políticas estacionarias. Para el modelo de control Markoviano asociado, las mismas condiciones implican la existencia de políticas estacionarias CP-óptimas.

6.3 Problemas abiertos

1) Como es conocido, en el análisis del problema de control Markoviano con costo promedio usando el método del factor desvaneciente, bajo diversos tipos de condiciones, es suficiente la existencia de soluciones a la desigualdad de optimalidad para garantizar la existencia de políticas estacionarias CP-óptimas; esto es una consecuencia de los llamados Teoremas Tauberianos. Un problema importante es el encontrar un resultado de este tipo que relacione los criterios de optimalidad que se estudiaron aquí.

2) Establecer condiciones que garanticen la igualdad de los índices de funcionamiento J y Φ (ver sección 4.4). Esto permitiría obtener resultados de optimalidad si se tiene la existencia de soluciones de la desigualdad de optimalidad.

3) En el Capítulo 5, se establecieron condiciones para la existencia de una política estacionaria f, CP-óptima en la subfamilia de las políticas estacionarias. Un problema interesante es el encontrar alguna(s) condición(es) adicional(es) para que la política f sea CP-óptima.

Apéndice A

Cadenas de Markov

Sea $\{x_n\}$ una cadena de Markov con espacio de estados el espacio de Borel X y probabilidad de transición $P(\cdot \mid \cdot)$. La probabilidad de transición en el n−ésimo paso se define inductivamente por

$$P^0(B \mid x) := \delta_x(B), \quad x \in X, \; B \in \mathcal{B}(X),$$

donde $\delta_x(\cdot)$ es la medida de Dirac definida por

$$\begin{aligned} \delta_x(B) \; &:= 1 \text{ si } x \in B \\ &:= 0 \text{ si } x \notin B \end{aligned}$$

y, para $n = 1, 2, ...$

$$P^n(B \mid x) := \int_X P(dy \mid x) P^{n-1}(B \mid y), \quad x \in X, \; B \in \mathcal{B}(X).$$

Definición A.1. Sea φ una medida σ−finita en $\mathcal{B}(X)$. La cadena se dice ser φ*−irreducible*, si para cada $x \in X$ y $B \in \mathcal{B}(X)$ con $\varphi(B) > 0$, existe $n \in \mathbf{N}$ tal que $P^n(B \mid x) > 0$. A φ se le llama una medida de irreducibilidad para la cadena.

Definición A.2. Una sucesión $(E_0, E_1, ..., E_{m-1})$ de subconjuntos ajenos no vacíos en $\mathcal{B}(X)$ es un $m-$ *ciclo* ($m \in \mathbf{N}$) para $\{x_n\}$ si para $i = 0, 1, ...$ y para todo $x \in E_i$,

$$P(E_j^c \mid x) = 0 \quad \text{para } j = i + 1 \text{ (mod } m)$$

Proposición A.3. ([29] p. 21). Si $\{x_n\}$ es $\varphi-$irreducible, entonces existe un número entero $d \geq 1$ tal que
i) Existe un $d-$ciclo $\{E_0, E_1, ..., E_{d-1}\}$.
ii) Si $\{E'_0, E'_1, ..., E'_{d'-1}\}$ es cualquier otro $d'-$ciclo, entonces d' divide a d.

Definición A.4. El número d en la Proposición A.3 se llama el *periodo* de $\{x_n\}$. Si $d = 1$, la cadena se dice que es *aperiódica*.

Nota A.5. De lo anterior se sigue que si $P(\{x\} \mid x) > 0 \quad \forall x \in X$, entonces la cadena es aperiódica.

Definición A.6. La cadena $\{x_n\}$ se dice ser *Harris-recurrente*, si existe una medida $\sigma-$finita λ en $\mathcal{B}(X)$ tal que para cada $x \in X$ y $B \in \mathcal{B}(X)$ con $\lambda(B) > 0$,

$$P_x(x_n \in B \ i.o.) = 1.$$

Definición A.7. $\{x_n\}$ es *Harris-recurrente positiva*, si es Harris-recurrente y tiene una medida de probabilidad invariante, es decir, una medida de probabilidad $p \in \mathcal{B}(X)$ tal que

$$p(B) = \int P(B \mid x)p(dx) \quad \forall B \in \mathcal{B}(X).$$

Para una distribución de probabilidad $a = \{a(n)\}$ en $\mathbf{N}_0$, se define el kérnel estocástico K_a en X por

$$K_a(B \mid x) = \sum_{n=0}^{\infty} P^n(B \mid x)a(n).$$

Definición A.8. Sea a una distribución de probabilidad en $\mathbf{N}_0$ y ν_a una medida no trivial en X. Un conjunto $C \subset X$, se dice que es ν_a-petite, si

$$K_a(B \mid x) \geq \nu_a(B) \quad \forall x \in C, \ B \in \mathcal{B}(X).$$

Si la medida ν_a no es importante, se dice simplemente que C es un conjunto *petite*.

Proposición A.9. ([7], Lemma 2.1). Si $\{x_n\}$ es $\varphi-$irreducible, entonces $C \in \mathcal{B}(X)$ es petite si y solo si existen $n \in \mathbf{N}$ y una medida no trivial ν tales que

$$\sum_{i=1}^{n} P^i(\cdot \mid x) \geq \nu(\cdot), \ x \in C.$$

Proposición A.10. ([7], Theorem 2.3, [25], Theorem 16.0.1). Sea $\{x_n\}$ $\varphi-$irreducible y aperiódica que satisface la siguiente condición: Existen una función medible V : $X \longrightarrow [1, \infty)$, un conjunto petite $C \in \mathcal{B}(X)$ y constantes $\beta < 1$ y $b > 0$ tal que

$$\int V(y)P(dy \mid x) \leq \beta V(x) + b1_C(x), \quad x \in X.$$

Entonces:

a) La cadena $\{x_n\}$ es Harris-recurrente positiva. Denótese por p la probabilidad invariante.

b) $\int V(y)p(dy) < \infty$.

c) La cadena es $V-$uniformemente ergódica, es decir, existen constantes positivas $\gamma < 1$ y $M < \infty$ tales que

$$\left|\int u(y)P^n(dy \mid x) - \int u(y)p(dy)\right| \leq M\gamma^n V(x).$$

para todo $x \in X$, $n = 0, 1, \ldots$, y $u \in L_V^\infty$.

Referencias

[1] R.B. Ash. *Real Analysis and Probability.* Academic Press, New York. 1972.

[2] R.N. Bhattacharya and M. Majumdar. *Controlled semi-Markov models–the discounted case.* J.Statist. Plann. Inference **21** (1989), 365-381.

[3] R.N. Bhattacharya and M. Majumdar. *Controlled semi-Markov models under long-run average rewards.* J.Statist. Plann. Inference **22** (1989). 223-242.

[4] A. Federgruen, A. Hordijk and H.C. Tijms. *Denumerable state semi-Markov decision processes with unbounded costs, average cost criterion.* Stoch. Proc. Appl. **9** (1979), 222-235.

[5] A. Federgruen, P.J. Schweitzer and H.C. Tijms. *Denumerable undiscounted semi-Markov decision processes with unbounded rewards.* Math. Oper. Res. **8-2** (1983). 298-313.

[6] A. Federgruen and H.C. Tijms. *The optimality equation in average cost denumerable state semi-Markov decision problems. Recurrence conditions and algorithms.* J. Appl. Prob. **15** (1978). 356-373.

[7] P.W. Glynn and S.P. Meyn. *A Lyapounov bound for solutions of Poisson's equation.* Ann. Probab. **24** (1996). 916-931.

[8] E. Gordienko and O. Hernández-Lerma. *Average cost Markov control processes with weighted norms: existence of canonical policies.* Appl. Math. **23** (1995). 199-218.

[9] E. Gordienko and O. Hernández-Lerma. *Average cost Markov control processes with weighted norms: value iteration*. Appl. Math. **23** (1995), 219-237.

[10] E. Gordienko, R. Montes-de-Oca and J.A. Minjárez-Sosa. *Approximation of average cost optimal policies for general Markov decision processes with unbounded costs*. Math. Methods Oper. Res. (por aparecer)

[11] U.G. Haussmann. *On the optimal long-run control of Markov renewal processes*. J. Math. Anal. Appl. **36** (1971), 123-140.

[12] O. Hernández-Lerma. *Adaptive Markov Control Processes*. Springer-Verlag, New York (1989).

[13] O. Hernández-Lerma and J.B. Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer-Verlag, New York (1996).

[14] O. Hernández-Lerma and J.B. Lasserre. *Policy iteration for average cost Markov control processes on Borel spaces*. Acta Appl. Math. (por aparecer).

[15] O. Hernández-Lerma and M. Muñoz de Ozak. *Discrete-time MCPs with discounted unbounded costs: Optimality criteria*. Kybernetika (Prague) **28** (1992), 191-212.

[16] O. Hernández-Lerma and W.J. Runggaldier. *Monotone approximations for convex stochastic control problems*. J. Math. Syst., Estimation, and Control **4** (1994), 99-140.

[17] O. Hernández-Lerma and O. Vega-Amaya. *Infinite-horizon Markov control processes with undiscounted cost criteria: from average to overtaking optimality*. (Sometido)

[18] Q.Y. Hu. *Discounted and average MDPs with unbounded rewards: new conditions*. J. Math. Anal Appl. **171** (1992), 111-124.

[19] N.V. Kartashov. *Inequalities in theorems of ergodicity and stability of Markov chains with common phase space*. I. Theory Probab. Appl. **30** (1985), 247-259.

[20] M.Y. Kitayev. *Semi-Markov and jump Markov control models: average cost criterion.* Theory Probab. Appl. **30** (1985), 272-288.

[21] S.A. Lippman. *Semi-Markov decision processes with unbounded rewards.* Manage. Sci. **19** (1973), 717-731.

[22] S.A. Lippman. *On dynamic programming with unbounded rewards.* Manage Sci. **21** (1975), 1225-1233.

[23] F. Luque-Vásquez and O. Hernández-Lerma. *A counterexample on the semicontinuity of minima.* Proc. Amer. Math. Soc. **123** (1995), 3175-3176.

[24] F. Luque-Vásquez and M.T. Robles-Alcaraz. *Controlled semi-Markov models with discounted unbounded costs.* Bol. Soc. Mat. Mexicana **39** (1994), 51-68.

[25] S.P. Meyn and R.L Tweedie. *Markov Chains and Stochastic Stability.* Springer-Verlag, London (1993).

[26] S.P. Meyn and R.L. Tweedie. *Computable bounds for geometric convergence rates of Markov chains.* Ann. Appl. Prob. **4** (1994), 981-1011.

[27] R. Montes-de-Oca. *The average cost optimality equation for Markov control processes on Borel spaces.* Syst. Control Lett. **22** (1994), 351-357.

[28] R. Montes-de-Oca and O. Hernández-Lerma. *Conditions for average optimality in Markov control processes with unbounded costs and controls.* J. Math. Syst., Estimation, and Control **4** (1994), 145-148.

[29] E. Nummelin. *General Irreducible Markov Chains and Non-Negative Operators.* Cambridge University Press, Cambridge, 1984.

[30] M.L. Puterman. *Markov Decision Processes. Discrete Stochastic Dynamic Programming.* Wiley, New York (1994).

[31] U. Rieder. *Measurable selection theorems for optimization problems.* Manuscripta Math. **24** (1978), 115-131.

[32] S.M. Ross. *Applied Probability Models with Optimization Applications*. Holden-Day, San Francisco (1970).

[33] S.M. Ross. *Average cost semi-Markov decision processes*. J. Appl. Prob. **7** (1970), 649-656.

[34] H.L. Royden. *Real Analysis*. Macmillan, New York (1988).

[35] M. Schäl. *Conditions for optimality and for the limit of n-stage optimal policies to be optimal*. Z. Wahrs. verw. Gerb. **32** (1975), 179-196.

[36] M. Schäl. *On the second optimatility equation for semi-Markov decision models*. Math. Op. Res. **17** (1992), 470-486.

[37] P.J. Schweitzer. *Iterative solutions of the functionals equations of undiscounted Markov renewal programming*. J. Math. Anal. Appl. **34** (1971), 495-501.

[38] L.I. Sennott. *Average cost semi-Markov decision processes and the control of queueing systems*. Probab. Eng. Inform. Sci. **3** (1989), 247-272.

[39] L.I. Sennott. *Another set of conditions for average optimality in Markov control processes*. Syst. Control Lett. **24** (1995), 147-151.

[40] R. Serfozo. *An equivalence between continuous and discrete time Markov decision processes*. Op. Res. **27** (1979), 616-620.

[41] H.C. Tijms. *Stochastic Modeling and Analysis: A Computational Approach*. Wiley, Chichester (1986).

[42] H.C. Tijms. *Average reward optimality equation in Markov decision processes with a general state space*. Probability, Statistics and Optimisation (F.P. Kelly, Ed.). Wiley, New York (1994), 485-495.

[43] O. Vega-Amaya. *Average optimality in semi-Markov control models on Borel spaces: Unbounded costs and controls*. Bol. Soc. Mat. Mex. **38** (1993), 47-60; Erratum. **39** (1994), 68.

[44] O. Vega-Amaya y F. Luque-Vásquez. *Modelos de control semi-Markoviano.* Aportaciones Matemáticas, Soc. Mat. Mex. **11** (1994), 171-183.