

**PROCESOS DE CONTROL DE MARKOV
CON
COSTO PROMEDIO**

Tesis que presenta:

José Raúl Montes de Oca Machorro

Para la obtención del grado de Doctor en Ciencias.

Junio de 1994

Director: Dr. Onésimo Hernández-Lerma

**Universidad Autónoma Metropolitana-Iztapalapa
División de Ciencias Básicas e Ingeniería
Departamento de Matemáticas**

Agradecimientos

Recibí apoyo y motivación de mucha gente durante el período previo a la elaboración de este trabajo. Tengo la seguridad de que el tiempo que cada una de ellas me hizo el favor de dedicar ha contribuido de manera importante para la realización de esta Tesis. Por esto quiero dejar constancia de mi agradecimiento a:

*mis padres y hermano,
Guadalupe Benítez Rodríguez,
Dr. Onésimo Hernández Lerma,
Dr. Fernando Galaz Fontes, y a*

*Egidio Barrera, Alejandro Islas, Fernando Garibay, Lorenzo Gutiérrez,
Enrique Lemus y Francisco Salem.*

A todos ellos les dedico este trabajo.

Finalmente, agradezco al Consejo Nacional de Ciencia y Tecnología por el apoyo que me dió durante el tiempo en que realicé mis estudios de maestría, así como el apoyo que actualmente me brinda mediante el proyecto 1332-E9206.

Indice

PRIMERA PARTE: PRELIMINARES

1. Procesos de control de Markov.
 - 1.1 Introducción 3
 - 1.2 El modelo de control de Markov 3
 - 1.3 Políticas de control 4
 - 1.4 Problemas de control óptimo 5
 - 1.5 Resultados del problema con costo descontado 6

SEGUNDA PARTE: PLANTEAMIENTO DEL PROBLEMA

2. El problema de control óptimo con costo promedio: caso no acotado.
 - 2.1 Introducción 8
 - 2.2 El problema 8
 - 2.3 Métodos de solución del problema 11

TERCERA PARTE: SOLUCION AL PROBLEMA

3. Condiciones de optimalidad.
 - 3.1 Introducción 13
 - 3.2 La desigualdad de optimalidad 13
 - 3.3 Condiciones suficientes para la desigualdad de optimalidad 18
 - 3.4 Conclusiones 21
4. La ecuación de optimalidad.
 - 4.1 Introducción 22
 - 4.2 La propiedad de equicontinuidad 22
 - 4.3 Condición para la ecuación de optimalidad 23
 - 4.4 Conclusiones 26
5. Convergencia del algoritmo de iteración de valores.
 - 5.1 Introducción 27
 - 5.2 Ternas canónicas 27
 - 5.3 Iteración de valores 30
 - 5.4 Construcción de una política CP-óptima 37
 - 5.5 Comentarios y conclusiones 39

CUARTA PARTE: EJEMPLOS Y CONCLUSIONES.

Resumen de hipótesis	41
6. Un ejemplo: el problema lineal con costo cuadrático.	
6.1 Introducción	43
6.2 El problema lineal con costo cuadrático	43
6.3 Verificación de hipótesis	44
7. Conclusiones y problemas abiertos.	
7.1 Introducción	47
7.2 Conclusiones	47
7.3 Problemas abiertos	48
APENDICE	50
BIBLIOGRAFIA	52

PROCESOS
DE
CONTROL DE MARKOV
CON
COSTO PROMEDIO

Resumen

En esta tesis se estudian procesos de control de Markov en espacios de Borel con costo promedio, permitiendo que la función costo sea no acotada, y que los conjuntos de controles admisibles en cada estado no sean compactos.

El objetivo principal consiste en proponer condiciones que garanticen la existencia de políticas óptimas y que permitan caracterizar dichas políticas, así como el valor promedio óptimo, mediante técnicas de Programación Dinámica.

Las condiciones propuestas en este trabajo están dadas principalmente en base a la función de valor óptimo para problemas de control con costo total descontado, por lo que el trabajo cae en el contexto del método del "*factor de descuento desvaneciente*". Tales condiciones permiten obtener soluciones de la desigualdad y de la ecuación de optimalidad, que a su vez permiten obtener y caracterizar políticas óptimas y el valor promedio óptimo.

Bajo condiciones adicionales, también se obtiene la convergencia del algoritmo de iteración de valores.

La generalidad de nuestras condiciones permite que sean aplicables a la mayoría de los modelos de control de Markov que surgen en la teoría de control estocástico en tiempo discreto. Esto permite, en particular, ilustrar nuestros resultados con el problema del regulador lineal con costo cuadrático, para el cual no son aplicables resultados anteriores.

Introducción

En esta tesis se estudian procesos de control de Markov (PCMs) en los que el espacio de estados y el de controles son ambos espacios de Borel y el índice de funcionamiento es el costo promedio esperado. Es decir, se estudia el problema de control óptimo con costo promedio (PCOCP).

Este problema ha sido ampliamente estudiado para PCMs en los que: (i) el espacio de estados es un conjunto numerable, y/ó (ii) los conjuntos de controles admisibles son compactos, y/ó (iii) la función de costo. (ver e.g. [2]). Estas restricciones excluyen importantes problemas de control como, el problema lineal con costo cuadrático (ver ejemplo 2.1, Cap. II) en ninguna de las condiciones (i), (ii) e (iii) se cumplen. Por otro lado, la solución del PCOCP en muchas de las situaciones (i)-(iii) antes descritas, se ha obtenido encontrando soluciones de la desigualdad ó de la igualdad de optimalidad con costo promedio, así como por medio del algoritmo de iteración de valores.

Por lo tanto, aquí proponemos extender a nuestro contexto las técnicas de la desigualdad y de la igualdad de optimalidad y del algoritmo de iteración de valores haciendo uso del enfoque del factor de descuento desvaneciente. Es decir, la idea es formular condiciones que aseguren que el problema de control óptimo con costo descontado tiene solución y, basándose en un Teorema Tauberiano (ver Ap. B), estudiaremos el caso promedio como límite del caso descontado.

Nuestro trabajo es una extensión al caso de Borel de resultados de Sennott [37, 38] para PCMs con espacio de estados numerable y conjunto de controles finitos, así como del artículo de Fernández-Gaucherand, Marcus y Arapostathis [9] sobre PCMs convexos, y tiene como antecedentes los trabajos [15, 25, 26, 27, 28].

El trabajo está dividido en cuatro partes. La primera y la segunda parte contienen cada una de ellas, un capítulo y presentan respectivamente, las definiciones y resultados preliminares para PCMs y el planteamiento del problema que nos interesa estudiar. La tercera parte consiste de tres capítulos en los que se analizan la desigualdad y la igualdad de optimalidad y el algoritmo de iteración de valores. La cuarta parte contiene: un resumen de las hipótesis propuestas en los capítulos anteriores, seguido de dos capítulos donde se consideran un ejemplo y las conclusiones del trabajo. También se incluye una sección de problemas abiertos para investigaciones futuras. Finalmente, en los apéndices A, B y C se ha concentrado la notación, la terminología y teoremas de referencia que se citan en la tesis.

Primera Parte

PRELIMINARES

1 Procesos de Control de Markov

- 1.1 Introducción.
- 1.2 El Modelo de control de Markov.
- 1.3 Políticas de control
- 1.4 Problemas de control óptimo.
- 1.5 Resultados del problema con costo descontado.

1.1 Introducción

Este capítulo tiene por objetivos: (a) plantear los problemas de control óptimo que son de interés para este trabajo y (b) dar una lista de hipótesis y resultados que apoyan el planteamiento y la solución al problema que se proponen en esta tesis.

En las secciones 2 y 3 se definen los elementos necesarios para que en la sección 4 se formulen los problemas de control óptimo con costo descontado y con costo promedio. En la sección 5 se dan hipótesis que aseguran la existencia de "minimizadores medibles", así como algunos resultados relacionados con el problema con costo descontado.

1.2 El modelo de control de Markov

El modelo de control de Markov (a tiempo discreto, homogéneo en el tiempo) (X, A, Q, c) consiste de: el espacio de estados X , el conjunto de controles (ó acciones) A , la ley de transición Q y la función de costo c . Se supone que X y A son espacios de Borel [def. 2, Ap. A].

A cada estado $x \in X$ se asocia un subconjunto $A(x)$ de A el cual se supone que es medible y no vacío. $A(x)$ representa el conjunto de controles (o acciones) admisibles en el estado x . El conjunto

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

de las parejas admisibles estado-acción, suponemos que es medible en el producto cartesiano de X y A y que, además, contiene la gráfica de una función $f : X \rightarrow A$, medible. (La existencia de tal función f garantiza que el conjunto Δ , de políticas de control definido posteriormente, no es vacío). La ley de transición $Q(B|x, a)$, donde B está en $\mathcal{B}(X)$, la σ -álgebra de Borel del espacio X , y $(x, a) \in \mathbf{K}$, es un kernel estocástico [Def. 3, Ap. A] sobre X dado \mathbf{K} . Finalmente la función c , definida en \mathbf{K} , se supone medible y no negativa.

El modelo de control de Markov (X, A, Q, c) representa un sistema estocástico controlado el cual es observado en tiempos $t = 0, 1, \dots$. Si denotamos por x_t y a_t el estado del sistema y el control aplicado en el tiempo t , el desarrollo del sistema es el siguiente. Si el sistema se encuentra en el estado $x_t = x \in X$ y el control $a_t = a \in A(x)$ es aplicado, entonces se tiene: (1) se paga un costo $c(x, a)$ y (2) el sistema se mueve a un nuevo estado x_{t+1} de acuerdo a la distribución $Q(\cdot|x, a)$, i.e.,

$$Q(B|x, a) = P(x_{t+1} \in B | x_t = x, a_t = a), \quad B \in \mathcal{B}(X).$$

Una vez que la transición a un estado, digamos $x_{t+1} = x'$ ha ocurrido, se escoge un nuevo control $a_{t+1} = a' \in A(x')$ y se repite el proceso antes descrito. En este trabajo el estado inicial x_0 es un punto dado del espacio X , i.e., $x_0 = x \in X$.

1.3 Políticas de Control

Sea F el conjunto formado por las funciones medibles $f : X \rightarrow A$ tales que $f(x) \in A(x), \forall x \in X$. Para cada $t = 0, 1, \dots$ se define el espacio de *historias admisibles* hasta el tiempo t como $H_0 := X$, y $H_t := \mathbf{K}^t \times X = \mathbf{K} \times H_{t-1}$, $t = 1, 2, \dots$, i.e., H_t consiste de elementos de la forma:

$$h_t := (x_0, a_0, \dots, x_{t-1}, a_{t-1}, x_t),$$

donde $(x_n, a_n) \in \mathbf{K}$ para $n = 0, 1, \dots, t-1$ y $x_t \in X$.

Définición 3.1. (a) Una *política de control* es una sucesión $\delta = \{\delta_t\}$ tal que, para cada $t = 0, 1, \dots, \delta_t$ es un kernel estocástico sobre A dado H_t que satisface $\delta_t(A(x_t)|h_t) = 1, \forall h_t \in H_t$. El conjunto de todas las políticas se denotará Δ . (b) Diremos que una política de control δ es una *política*

estacionaria si existe una función $f \in F$ tal que $\delta_t(\cdot|h_t)$ está concentrado en $f(x)$, $\forall t \geq 0$. En este caso, identificaremos δ con $f \in F$, y llamaremos a F el conjunto de las políticas estacionarias.

1.4 Problemas de Control Optimo

Denotaremos por P_x^δ a la medida de probabilidad inducida en $(X \times A)^\infty$, cuando se fija la política δ y el estado inicial es x (ver [22] pág. 80 para la construcción de P_x^δ). El operador esperanza correspondiente a P_x^δ se denotará E_x^δ .

Para cada política $\delta \in A$ y estado inicial $x \in X$, definimos el *costo promedio esperado* como:

$$J(\delta, x) := \limsup_{n \rightarrow \infty} (n+1)^{-1} \sum_{t=0}^{\infty} E_x^\delta [c(x_t, a_t)], \quad (4.1)$$

y el *costo total descontado* como:

$$V_\alpha(\delta, x) := E_x^\delta \left[\sum_{t=0}^{\infty} \alpha^t c(x_t, a_t) \right], \quad (4.2)$$

donde $\alpha \in (0, 1)$ es un *factor de descuento*. Las funciones

$$J(x) := \inf_{\delta} J(\delta, x), \text{ y } V_\alpha(x) := \inf_{\delta} V_\alpha(\delta, x), x \in X, \quad (4.3)$$

serán llamadas el *costo promedio óptimo* y el *costo α -descontado óptimo*, respectivamente.

Para finalizar la sección y basándonos en (4.1)-(4.3) presentamos los problemas siguientes.

EL PROBLEMA DE CONTROL OPTIMO CON COSTO PROMEDIO (PCOCP)

Determinar una política $\delta \in \Delta$ tal que $J(x) = J(\delta, x)$, $\forall x \in X$. En este caso δ será llamado *costo-promedio óptima (CP-óptima)*.

Análogamente tenemos:

EL PROBLEMA DE CONTROL OPTIMO CON COSTO DESCONTADO (PCOCD).

Determinar una política $\delta^* \in \Delta$ tal que $V_\alpha(x) = V_\alpha(\delta^*, x)$, $\forall x \in X$. En este caso decimos que δ^* es *α -óptima*.

1.5 Resultados del Problema con Costo Descontado

Para garantizar la existencia de minimizadores medibles (ver Teorema 1, Ap. C), se necesita imponer condiciones de semicontinuidad y compacidad en las componentes del modelo de control. Para formular estas condiciones requerimos de la siguiente definición.

Definición 5.1. Una función V definida en \mathbf{K} se llamará *inf-compacta* en \mathbf{K} , si el conjunto $\{a \in A(x) | V(x, a) \leq r\}$ es compacto para cada $r \in \mathbf{R}$ y $x \in X$.

Hipótesis. (H) (a) La función de costo $c = c(x, a)$ se supone semicontinua inferiormente e inf-compacta en \mathbf{K} ;

(b) La ley de transición Q es fuertemente continua, i.e., la función $V'(x, a) = \int V(y)Q(dy|x, a)$ es continua y acotada en $(x, a) \in \mathbf{K}$ para cada función $V : X \rightarrow \mathbf{R}$ medible y acotada.

La hipótesis **H** se cumple, por ejemplo, en la situación siguiente.

Ejemplo 5.2. Sistemas con ruido aditivo y costo cuadrático. Considérese el sistema estocástico:

$$x_{t+1} = F(x_t, a_t) + \epsilon_t, \quad t \geq 0, \quad (5.1)$$

con costo cuadrático:

$$c(x, a) = x'qx + a'ra$$

(x' y a' denotan la transpuesta de x y de a respectivamente). Aquí $X = \mathbf{R}^n$, $A(\cdot) = A = \mathbf{R}^m$, $F : X \times A \rightarrow X$ es una función continua, q y r son matrices simétricas y definidas positivas. Este ejemplo incluye el problema lineal con costo cuadrático en el que F es de la forma $F(x, a) = \gamma x + \beta a$, con γ y β matrices dadas. Suponemos que $\epsilon_0, \epsilon_1, \dots$ son variables aleatorias independientes e idénticamente distribuidas con densidad continua $g(s)$.

En este ejemplo, la hipótesis **H** (a) se cumple trivialmente. Para verificar **H** (b) sea u una función medible y acotada definida en $X = \mathbf{R}^n$. Entonces

$$\begin{aligned} \int u(y)Q(dy|x, a) &= \int u(F(x, a) + s)g(s)ds \\ &= \int u(s)g(s - F(x, a))ds \end{aligned}$$

Así, usando el Teorema de Scheffé (ver [10] pág. 125), tenemos:

$$\begin{aligned} \left| \int u(y)Q(dy|\bar{x}, \bar{a}) - \int u(y)Q(dy|x, a) \right| \\ \leq \|u\| \int |g(s - F(\bar{x}, \bar{a})) - g(s - F(x, a))| ds \\ \rightarrow 0 \quad \text{si } (\bar{x}, \bar{a}) \rightarrow (x, a), \end{aligned}$$

i.e., H (b) se cumple.

El lema siguiente constituye el Teorema 4.23 en [21].

Lema 5.3. *Supóngase que las hipótesis H se cumplen. Sea $\alpha \in (0, 1)$ fijo. Si para cada $x \in X$,*

$$V_\alpha(x) < \infty,$$

entonces V_α es la mínima solución (puntual) de la Ecuación de Programación Dinámica para el caso descontado, i.e.,

$$V_\alpha(x) = \min_{a \in A(x)} \left[c(x, a) + \alpha \int V_\alpha(y)Q(dy|x, a) \right], \quad x \in X. \quad (5.2)$$

Además, existe una política estacionaria $f_\alpha \in F$ tal que, para todo $x \in X$, $f_\alpha(x) \in A(x)$ minimiza el lado derecho de (5.2), i.e.,

$$V_\alpha(x) = c(x, f_\alpha(x)) + \alpha \int V_\alpha(y)Q(dy|x, f_\alpha(x)) \quad \forall x \in X, \quad (5.3)$$

y f_α es α -óptima.

Segunda Parte

PLANTEAMIENTO
DEL
PROBLEMA

2 El Problema de Control Optimo con Costo Promedio: Caso no Acotado

2.1 Introducción.

2.2 El Problema.

2.3 Métodos de solución del problema.

2.1 Introducción

En este capítulo formularemos el problema que se presenta en esta tesis. Esta formulación del problema se da en la sección 2. En la sección 3 se exponen las técnicas con que el problema ha sido estudiado y se revisan algunos antecedentes de su solución.

2.2 El Problema

El análisis de los procesos de control de Markov (PCMs) a tiempo discreto y con costo promedio, el cual se remonta a los años 50 (en [2] se encuentra una recopilación bastante completa acerca de la historia de los PCMs con costo promedio), ha estado concentrado principalmente en PCMs para los cuales:

- (i) el espacio de estados es un conjunto numerable; y/ó
- (ii) los conjuntos de controles admisibles en cada estado son conjuntos compactos; y/ó
- (iii) la función de costo es acotada.

No obstante, hay casos importantes donde no se cumplen algunas de las condiciones (i)-(iii). El hecho que la función de costo no sea acotada ha sido considerado en algunas aplicaciones de control estocástico a colas [31, 32, 36, 37, 41]. También se tiene el ejemplo siguiente (el cual es un caso particular del ejemplo 5.2, Cap. I), en el que ninguna de las condiciones (i)-(iii) se cumple.

Ejemplo 2.1. El problema lineal con costo cuadrático (ó problema del regulador lineal). Consideremos el sistema estocástico:

$$x_{t+1} = \gamma x_t + \beta a_t + \epsilon_t, \quad t = 0, 1, \dots,$$

con función costo

$$c(x, a) = x'qx + a'ra,$$

donde x' y a' denotan la transpuesta de x y de a , respectivamente. Aquí, $X = \mathbb{R}^n$ y $A = A(\cdot) = \mathbb{R}^m$, $n, m \geq 1$, γ y β son matrices de dimensiones apropiada, y q y r son matrices dadas, simétricas y definidas positivas. $\epsilon_0, \epsilon_1, \dots$, son variables aleatorias independientes e idénticamente distribuidas, e independientes del estado inicial x_0 .

Para dar solución al PCOCP hay técnicas como el uso de una desigualdad ó igualdad de optimalidad, y el algoritmo de iteración de valores, las cuales han sido aplicadas en las situaciones (i)-(iii) antes descritas. Para precisar esto, supongamos que (X, A, Q, c) es un modelo de control fijo, y considérese el PCOCP para este modelo. Tenemos:

Definición 2.1. Un par (j^*, h) que consiste de un número real j^* y una función $h : X \rightarrow \mathbb{R}$ medible se dice que es *solución de la desigualdad de optimalidad con costo promedio* si

$$j^* + h(x) \geq \min_{a \in A(x)} \left[c(x, a) + \int h(y)Q(dy|x, a) \right] \quad \forall x \in X, \quad (2.1)$$

y diremos que (j^*, h) es *solución de la igualdad de optimalidad con costo promedio* si en (2.1) se tiene igualdad, i.e.,

$$j^* + h(x) = \min_{a \in A(x)} \left[c(x, a) + \int h(y)Q(dy|x, a) \right] \quad \forall x \in X. \quad (2.2)$$

Bajo ciertas circunstancias (e.g. suponiendo las hipótesis H, y que la función h en (2.1) es acotada por abajo), es conocido (ver e.g. [14]) que si (2.1) se cumple, entonces existe una política estacionaria $f^* \in F$ la cual minimiza el lado derecho de (2.1), i.e.,

$$\begin{aligned} \min_{a \in A(x)} \left[c(x, a) + \int h(y)Q(dy|x, a) \right] \\ = c(x, f^*(x)) + \int h(y)Q(dy|x, f^*(x)), \quad \forall x \in X; \end{aligned}$$

además, f^* es CP-óptima, y j^* es el valor promedio óptimo. Si también se tiene que (2.2) se cumple, entonces un procedimiento para aproximar j^* y h es el siguiente.

Definición 2.2. Definimos las funciones u_0, u_1, \dots , conocidas como funciones de iteración de valores como sigue.

Sea $u_0(\cdot) \equiv 0$, y para $n \geq 1$, sea

$$u_n(x) = \min_{a \in A(x)} \left[c(x, a) + \int u_{n-1}(y) Q(dy|x, a) \right] \quad x \in X. \quad (2.3)$$

Es conocido [3] que:

$$u_n = \inf_{\delta} E_x^{\delta} \left[\sum_{t=0}^{n-1} c(x_t, a_t) \right], \quad x \in X.$$

Sean (j^*, h) solución de la igualdad de optimalidad con costo promedio, y $\bar{x} \in X$ un estado fijo. Entonces el objetivo del algoritmo de iteración de valores consiste en estudiar la convergencia de $\{u_n(x) - u_n(\bar{x})\}$ a $h(x)$, $x \in X$, y de $\{u_n(x) - u_{n-1}(x)\}$ a j^* .

Un problema relacionado con la convergencia del algoritmo de iteración de valores ocurre cuando tenemos que para cada $n \geq 1$, existe una política estacionaria $f_n \in F$ que minimiza el lado derecho de (2.3); entonces es importante tener condiciones que permitan construir una política estacionaria $\hat{f} \in F$ tal que para cada $x \in X$, $\hat{f}(x)$ resulte ser un punto de acumulación de $\{f_n(x)\}$, y que además \hat{f} sea CP-óptima.

A continuación damos el enunciado de nuestro problema.

Para esto, supongamos que:

- (a) X y A son espacios de Borel; y además se permite que:
- (b) los conjuntos $A(x)$, $x \in X$ no sean compactos; y que
- (c) la función costo no sea acotada.

En este contexto se plantea el problema:

PROBLEMA 2.4. Establecer condiciones que aseguren la existencia de políticas CP-óptimas y que permitan, caracterizar dichas políticas y el valor promedio óptimo por medio de: la desigualdad y/ó la igualdad de optimalidad con costo promedio (ver definición 2.2), y del algoritmo de iteración de valores (ver definición 2.3).

2.3 Métodos de solución del problema

El problema 2.4 ha sido considerado en trabajos recientes. Para el caso numerable, i.e., cuando X es un conjunto numerable, se puede encontrar un estudio completo acerca de esta situación en [2], y también se dan como referencias [4] y [5]. En el caso general, i.e., considerando X y A espacios de Borel, del cual nos ocupamos en esta tesis, se tienen antecedentes utilizando los enfoques siguientes (nota: se supone que las hipótesis H se cumplen).

2.3.1 Enfoque del factor de descuento desvaneciente

En este método se estudia el PCOCP por medio del PCOCD cuando el factor de descuento α tiende a 1. Se basa fundamentalmente en la desigualdad dada en el Teorema Tauberiano (ver Ap. B), i.e.,

$$\limsup_{\alpha \uparrow 1} (1 - \alpha)V_\alpha(x) \leq J(x) \quad \forall \delta \in \Delta, \quad x \in X,$$

donde $V_\alpha(\cdot)$ es la función de valor óptimo para el problema descontado (ver (4.3), Cap. I), y $J(x)$ es la función de valor óptimo para el caso promedio cuando el estado inicial es x (ver (4.3), Cap. I). Hernández-Lerma ha propuesto condiciones (hipótesis 2.1 y 3.1 en [11]) sobre V_α , los cuales, por medio del enfoque del factor de descuento desvaneciente, le permiten obtener políticas estacionarias CP-óptimas por medio de la desigualdad (2.1). De igual forma, Hernández-Lerma y Lasserre en [14], y Schäl en [35] han utilizado el enfoque del factor de descuento desvaneciente para obtener resultados similares a los obtenidos en [11], pero en ambos casos se considera que los conjuntos $A(x)$, $x \in X$ son compactos.

2.3.2 Enfoque cuasimonótono

En este procedimiento se impone la condición sobre la función costo c de ser *estrictamente no acotada*, i.e., por ejemplo si $X = A = \mathbf{R}$, se pide que:

$$\inf_{|x| > n} \inf_a c(x, a) \rightarrow \infty \quad \text{si } n \rightarrow \infty. \quad (3.1)$$

Hernández-Lerma ha propuesto una condición (hipótesis 2.1 (c) en [12]) semejante a la mencionada en (3.1); bajo ésta, y algunas condiciones de recurrencia (ver corolario 5.4 (b) en [12]) obtiene resultados que garantizan la

existencia de políticas estacionarias CP-óptimas por medio de la desigualdad (2.1).

Finalmente tenemos:

2.3.3 Enfoque utilizando programación lineal

En este enfoque se introduce un programa lineal EP y su dual EP* cuyos valores $\sup(\text{EP}^*)$ e $\inf(\text{EP})$ están relacionados con $J^* := \inf_x \inf_{\delta} J(\delta, x)$ por las desigualdades:

$$\sup \text{EP}^* \leq J^* \leq \inf \text{EP}. \quad (3.2)$$

Si se tiene que

$$\sup \text{EP}^* = \inf \text{EP}, \quad (3.2)$$

y además que EP y EP* tienen soluciones óptimas, i.e.,

$$\inf \text{EP} = \min \text{EP}; y$$

$$\sup \text{EP} = \max \text{EP},$$

se obtiene J^* , y también políticas CP-óptimas.

En [19] se estudia el problema 2.4 bajo el enfoque de programación lineal. En dicho trabajo se consideran condiciones (hipótesis 4.2 y 4.3 en [19]) que permiten obtener políticas estacionarias CP-óptimas y soluciones a la igualdad (2.2), pero los resultados obtenidos se cumplen *casi seguramente* con respecto a una medida de probabilidad relacionada a la solución óptima de EP. Cabe mencionar también el trabajo [20] en donde se combinan los enfoques de programación lineal y del factor de descuento desvaneciente para analizar el caso promedio. Las hipótesis impuestas en [20] (hipótesis 2.3, 3.1, 3.2 (a) y 5.1) permiten obtener soluciones de la igualdad (2.2) y políticas CP-óptimas casi seguramente; pero considerando una hipótesis adicional (ver inciso (c) del Teorema 6.2 en [20]) se obtienen resultados que se cumplen en todos los puntos de X .

Tercera Parte

SOLUCION
AL
PROBLEMA

3 Condiciones de Optimalidad

3.1 Introducción.

3.2 La desigualdad de Optimalidad.

3.3 Condiciones suficientes para la desigualdad de optimalidad.

3.4 Conclusiones.

3.1 Introducción

En este capítulo se propondrán condiciones (condiciones C1, C2, y C3) dadas principalmente en términos de la función de valor óptimo V_α para el PCOCD, y se utilizará el enfoque del factor de descuento desvaneciente (ver Secc.3 del Cap. II) como método de solución. Las condiciones C1, C2 y C3 dan respuesta al problema de existencia de políticas CP-óptimas, por medio de la obtención de soluciones de la igualdad de optimalidad (ver def. 2.2, Cap II). Los resultados de este capítulo están contenidos en los artículos [25] y [26].

En la sección 2 se formula la condición C1, la cual garantiza la existencia de políticas CP-óptimas. En la sección 3 se enuncian las condiciones C2 y C3, y se dan relaciones (implicaciones y/o equivalencias) entre C1, C2 y C3. Finalmente, en la sección 4 se dan las conclusiones del capítulo.

3.2 La desigualdad de optimalidad

Comenzaremos por describir la condición principal que consideraremos en este capítulo.

Sea V_α la función del valor óptimo para PCOCD, y sea $\bar{x} \in X$ un estado fijo pero arbitrario. Definimos:

$$h_\alpha(x) := V_\alpha(x) - V_\alpha(\bar{x}), \quad x \in X, \quad \alpha \in (0, 1). \quad (2.1)$$

CONDICION 1. (C1). Existen constantes no negativas N y M , una función no negativa b definida en X , no necesariamente medible, y $\alpha_0 \in (0, 1)$ tal que:

- (a) $V_\alpha(x) < \infty$ para todo $x \in X$ y $\alpha \in (0, 1)$;

(b) $(1 - \alpha)V_\alpha(\bar{x}) \leq M$ para todo $\alpha \in [\alpha_0, 1)$;

(c) $-N \leq h_\alpha(x) \leq b(x)$ para todo $x \in X$ y $\alpha \in [\alpha_0, 1)$.

La condición C1 fue introducida por L.I. Sennott en [38], para PCMs con espacio de estados numerable y conjunto de controles finito. En [38] se utiliza la condición C1 para garantizar la existencia de políticas estacionarias CP-óptimas.

Lema 3.1. *Bajo la condición C1, existe una constante j^* y una sucesión de factores de descuento $\alpha_n \uparrow 1$ tal que:*

$$\lim_{n \rightarrow \infty} (1 - \alpha_n)V_{\alpha_n}(x) = j^* \quad \forall x \in X. \quad (2.2)$$

DEMOSTRACIÓN. Sea \bar{x} el estado fijo considerado en (2.1), y sea j^* un punto límite de $(1 - \alpha)V_\alpha(\bar{x})$ cuando $\alpha \uparrow 1$. Sea $\alpha_n \uparrow 1$ tal que

$$\lim_{n \rightarrow \infty} (1 - \alpha_n)V_{\alpha_n}(\bar{x}) = j^*. \quad (2.3)$$

Ahora, para cada $x \in X$, (2.1), C1 y (2.3) permiten obtener:

$$\begin{aligned} |(1 - \alpha_n)V_{\alpha_n}(x) - j^*| &\leq (1 - \alpha_n)|h_{\alpha_n}(x)| + |(1 - \alpha_n)V_{\alpha_n}(\bar{x}) - j^*| \\ &\leq (1 - \alpha_n)\max\{N, b(x)\} + |(1 - \alpha_n)V_{\alpha_n}(\bar{x}) - j^*| \\ &\rightarrow 0 \quad \text{si } n \rightarrow \infty. \end{aligned}$$

Esto demuestra el lema. ■

A continuación daremos el Teorema principal del capítulo.

Teorema 3.2. *Supongamos que las hipótesis H y la condición C1 se cumplen. Sea j^* la constante obtenida en el lema 3.1. Entonces:*

(a) Existe una función medible h definida en X tal que: $-N \leq h(\cdot) \leq b(\cdot)$, y se cumple la desigualdad de optimalidad con costo promedio, i.e.,

$$j^* + h(x) \geq \min_{a \in A(x)} \left[c(x, a) + \int h(y)Q(dy|x, a) \right] \quad \forall x \in X; \quad (2.4)$$

(b) Existe una política estacionaria $f^* \in F$ tal que:

$$j^* + h(x) \geq c(x, f^*(x)) + \int h(y)Q(dy|x, f^*(x)) \quad \forall x \in X; \quad (2.5)$$

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

DEMOSTRACIÓN. (a) Sea $\alpha_n \uparrow 1$ la sucesión cuya existencia es garantizada en el lema 3.1, y sea h_α la función definida en (2.1).

Definimos $h : X \rightarrow \mathbf{R}$ como

$$\begin{aligned} h(x) &= \liminf_{n \rightarrow \infty} h_{\alpha_n}(x) \\ &= \lim_{n \rightarrow \infty} H_n(x), \quad x \in X, \end{aligned} \quad (2.6)$$

donde $H_n(x) := \inf_{k \geq n} h_{\alpha_k}(x) \uparrow h(x)$ si $n \rightarrow \infty$. Sea x un estado fijo, pero arbitrario. Entonces, por (2.1) y el lema 5.3, Cap. I, para cada n existe $a_n \in A(x)$ tal que:

$$\begin{aligned} (1 - \alpha_n)V_{\alpha_n}(\bar{x}) + h_{\alpha_n}(x) \\ = c(x, a_n) + \alpha_n \int h_{\alpha_n}(y)Q(dy|x, a_n) \quad \forall n \geq 1. \end{aligned} \quad (2.7)$$

Por otro lado, de (2.2) y (2.6), para cada $\epsilon > 0$ existe un entero $n(\epsilon)$ y una subsucesión $\{\alpha_{n_i}\}$ de $\{\alpha_n\}$ tal que:

$$j^* + h(x) + \epsilon \geq (1 - \alpha_{n_i})V_{\alpha_{n_i}}(\bar{x}) + h_{\alpha_{n_i}}(x) \quad \forall n_i \geq n(\epsilon). \quad (2.8)$$

Entonces combinando (2.7) y (2.8) tenemos:

$$\begin{aligned} j^* + h(x) + \epsilon \\ \geq c(x, a_{n_i}) + \alpha_{n_i} \int h_{\alpha_{n_i}}(y)Q(dy|x, a_{n_i}) \\ \geq c(x, a_{n_i}) + \alpha_{n_i} \int H_{n_i}(y)Q(dy|x, a_{n_i}) \quad \forall n_i \geq n(\epsilon). \end{aligned} \quad (2.9)$$

Sumando $\alpha_{n_i} \cdot N$ a ambos lados de (2.9) y usando que $\alpha_{n_i} < 1$, tenemos

$$\begin{aligned} j^* + (h(x) + N) + \epsilon \\ \geq c(x, a_{n_i}) + \alpha_{n_i} \int (H_{n_i}(y) + N)Q(dy|x, a_{n_i}) \quad \forall n_i \geq n(\epsilon) \end{aligned} \quad (2.10)$$

Definimos $S_i^T(\cdot) := \min\{T, H_{n_i}(\cdot) + N\}$, $T \geq 1$. Nótese que para cada $T \geq 1$, $S_i^T : X \rightarrow \mathbf{R}$ es medible, acotada, no negativa y $S_i^T(\cdot) \uparrow H_{n_i}(\cdot) + N$ si $T \uparrow \infty$. En (2.10) tenemos:

$$\begin{aligned} j^* + (H(x) + N) + \epsilon \\ \geq c(x, a_{n_i}) + \alpha_{n_i} \int D_i^T(y) Q(dy|x, a_{n_i}) \quad \forall n_i \geq n(\epsilon), T \geq 1. \end{aligned} \quad (2.11)$$

Para $i = 1, 2, \dots$, y $T \geq 1$ fijo sean

$$\begin{aligned} D_i^T(x) &:= \{a \in A(x) | c(x, a) + \alpha_{n_i} \int S_i^T(y) Q(dy|x, a) \\ &\leq j^* + (h(x) + N) + \epsilon\}. \end{aligned}$$

Sean $r := j^* + (h(x) + N) + \epsilon$ y $A_r(x) := \{a \in A(x) | c(x, a) \leq r\}$. Obsérvese que de las hipótesis H se sigue que $D_i^T(x)$ es un subconjunto cerrado de $A_r(x)$ y, por tanto, $D_i^T(x)$ es compacto. Además, como

$$a_{n_i} \in D_i^T(x) \text{ y } S_i^T(\cdot) \uparrow S^T(\cdot) := \min\{T, h(\cdot) + N\} \text{ si } i \rightarrow \infty,$$

tenemos que la sucesión de compactos $D_i^T(x) \downarrow D^T(x)$, donde

$$\begin{aligned} D^T(x) &:= \{a \in A(X) | c(x, a) + \int S^T(y) Q(dy|x, a) \\ &\leq j^* + (h(x) + N) + \epsilon\}, \end{aligned}$$

y $D^T(x)$ es compacto y no vacío. Consecuentemente, existe una subsucesión de $\{n_i\}$ (la cual denotamos nuevamente por $\{n_i\}$ para no complicar más la notación) y un control $a_x^T \in D^T(x)$ tal que $a_{n_i} \rightarrow a_x^T$. Escógase $L > n(\epsilon)$; entonces en (2.11) tenemos:

$$\begin{aligned} j^* + (h(x) + N) + \epsilon \\ \geq c(x, a_{n_i}) + \int S_L^T(y) Q(dy|x, a_{n_i}) \quad \forall n_i > L \end{aligned} \quad (2.12)$$

Luego, haciendo $i \rightarrow \infty$, las hipótesis H nos llevan a:

$$\begin{aligned} j^* + (h(x) + N) + \epsilon \\ \geq c(x, a_x^T) + \int S_L^T(y) Q(dy|x, a_x^T). \end{aligned}$$

Ahora, si hacemos $L \rightarrow \infty$,

$$\begin{aligned} j^* + (h(x) + N) + \epsilon \\ \geq c(x, a_x^T) + \int S^T(y)Q(dy|x, a_x^T). \end{aligned} \quad (2.13)$$

Como T era arbitrario, un argumento similar, pero ahora sobre la sucesión $\{D^T(x)\}_{t \geq 1}$ de conjuntos compactos y no vacíos, nos lleva a garantizar la existencia de una subsucesión de $\{a_x^T\}$ (denotada igual para evitar confusión) y de un control a_x en el conjunto

$$\begin{aligned} D(x) := \{a \in A(x) | c(x, a) + \int (h(y) + N)Q(dy|x, a) \\ \leq j^* + (h(x) + N) + \epsilon\}, \end{aligned}$$

tal que $a_x^T \rightarrow a_x$, $T \rightarrow \infty$. Considérese $T' \geq 1$; luego en (2.13) tenemos:

$$\begin{aligned} j^* + (h(x) + N) + \epsilon \\ \geq c(x, a_x^T) + \int S^{T'}(y)Q(dy|x, a_x^T) \quad \forall T \geq T'. \end{aligned}$$

Haciendo $T \uparrow \infty$ y luego $T' \uparrow \infty$ (nótese que $S^{T'}(\cdot) \uparrow h(\cdot) + N$ si $T' \rightarrow \infty$), obtenemos:

$$\begin{aligned} j^* + (h(x) + N) + \epsilon \\ \geq c(x, a_x) + \int (h(y) + N)Q(dy|x, a_x). \end{aligned}$$

Por tanto,

$$\begin{aligned} j^* + h(x) + \epsilon \\ \geq c(x, a_x) + \int h(y)Q(dy|x, a_x) \\ \geq \min_{a \in A(x)} [c(x, a) + \int h(y)Q(dy|x, a)], \end{aligned}$$

y como $\epsilon > 0$ y $x \in X$ eran arbitrarios, se concluye (2.4).

(b) Tomando $u(\cdot) := h(\cdot) + N \geq 0$ en el Teorema 1, Apéndice C, garantizamos la existencia de $f^* \in F$ tal que:

$$\begin{aligned} \min_{a \in A(x)} [c(x, a) + \int (h(y) + N)Q(dy|x, a)] \\ = c(x, f^*(x)) + \int (h(y) + N)Q(dy|x, f^*(x)) \quad \forall x \in X. \end{aligned}$$

Finalmente, restando la constante N y usando (2.4) obtenemos (2.5).

(c) Si iteramos (2.5) n veces obtenemos:

$$h(x) \geq \sum_{t=0}^{n-1} E_x^{f^*} [c(x_t, f^*(x_t))] - nj^* + E_x^{f^*} [h(x_n)],$$

$n \geq 1$. Como $h(\cdot) \geq -N$, tenemos

$$\sum_{t=0}^{n-1} E_x^{f^*} [c(x_t, f^*(x_t))] \leq h(x) + nj^* + N,$$

dividiendo entre n y haciendo $n \uparrow \infty$ da como resultado:

$$J(f^*, x) \leq j^* \quad \forall x \in X. \quad (2.14)$$

Por otro lado, del Teorema Tauberiano en el Ap. B y de (2.2) se tiene que

$$j^* \leq J(\delta, x) \quad \forall \delta \in \Delta, \quad x \in X. \quad (2.15)$$

De (2.14) y (2.15) se tiene que f^* es CP-óptima y además

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

Esto concluye la demostración del Teorema 2.2. ■

3.3 Condiciones suficientes para la desigualdad de optimalidad

Veremos en el teorema 3.3 que la condición C1, está relacionada con otras dos condiciones, C2 y C3 que en seguida detallaremos, anexando en cada una de ellas referencias de los trabajos donde previamente han sido usadas para estudiar el PCOCP.

CONDICION 2 (C2). [14]. Existe una constante N no negativa, una función b definida en X y no negativa, $\alpha_0 \in (0, 1)$ y una política estacionaria $\hat{f} \in F$ tal que:

(a) $V_\alpha(x) < \infty$ para todo $x \in X$ y $\alpha \in [\alpha_0, 1)$;

(b) $h_\alpha(x) \geq -N$ para todo $x \in X$ y $\alpha \in [\alpha_0, 1)$;

(c) $h_\alpha(x) \leq b(x)$ y $\int b(y)Q(dy|x, \hat{f}(x)) < \infty$ para todo $x \in X$ y $\alpha \in [\alpha_0, 1)$.

Definimos:

$$m_\alpha := \inf_x V_\alpha(x) \text{ y } g_\alpha(x) := V_\alpha(x) - m_\alpha, \quad x \in X, \alpha \in (0, 1). \quad (3.1)$$

CONDICION 3 (C3). [11,35]. (a) Existe una política $\hat{\delta} \in A$ y un estado inicial $\hat{x} \in X$ tal que $J(\hat{\delta}, \hat{x}) < \infty$;

(b) Existe $\beta \in [0, 1)$ tal que $\sup_{\beta < \alpha < 1} g_\alpha(x) < \infty$, para cada $x \in X$.

Tenemos el resultado siguiente

Teorema 3.3. *Supóngase que las hipótesis H se cumplen. Entonces:*

(a) *C1 es equivalente a C3;*

(b) *C2 implica C1.*

DEMOSTRACIÓN. (a) *C3 implica C1:* supóngase que C3 se cumple, y sea $\bar{x} := \hat{x}$. Usando C3 (a) y el Teorema Tauberiano (ver Ap. B) se puede demostrar (ver [11] y [35]) que:

$$0 \leq g^L \leq g^U \leq j < \infty \quad (3.2)$$

donde

$$j := \inf_{x \in X} \inf_{\delta \in A} J(\delta, x),$$

$$g^L := \liminf_{\alpha \uparrow 1} (1 - \alpha)m_\alpha,$$

$$g^U = \lim_{\alpha \uparrow 1} (1 - \alpha)m_\alpha.$$

Ahora demostraremos que $m_\alpha < \infty \forall \alpha \in (0, 1)$, procediendo por contradicción. Si existiera $\alpha_1 \in (0, 1)$ tal que $m_{\alpha_1} = \infty$, entonces $m_\alpha = \infty \forall \alpha \geq \alpha_1$, ya que la función $\alpha \rightarrow m_\alpha$, $\alpha \in (0, 1)$ es no decreciente (recuérdese que $c \geq 0$). Por lo tanto

$$\limsup_{\alpha \uparrow 1} (1 - \alpha)m_\alpha = \infty,$$

lo que contradice (3.2). De aquí se sigue que $m_\alpha < \infty \forall \alpha \in (0, 1)$. Luego,

$$\begin{aligned} V_\alpha(x) &= (V_\alpha(x) - m_\alpha) + m_\alpha \\ &= g_\alpha(x) + m_\alpha \\ &\leq \sup_{\beta < \alpha < 1} g_\alpha(x) + m_\alpha \\ &< \infty \quad [\text{por C3 (b)}], \end{aligned}$$

i.e., $V_\alpha(x) < \infty$ para cada $x \in X$ y $\alpha \in (\beta, 1)$. Como $V_\alpha(x)$ es creciente en α , se tiene que $V_\alpha(x) < \infty \forall \alpha \in (0, 1)$, de suerte que C1 (a) se cumple.

Por otro lado, de C3 (a) y del Teorema Tauberiano (Ap. B) también se tiene que:

$$\limsup_{\alpha \uparrow 1} (1 - \alpha)V_\alpha(\bar{x}) \leq J(\bar{x}) < \infty,$$

donde $J(\bar{x}) = \inf_\delta J(\delta, \bar{x})$. Por tanto, existen constantes $M = M(\bar{x}) > 0$ y $\mu = \mu(\bar{x})$ tales que:

$$0 \leq (1 - \alpha)V_\alpha(\bar{x}) \leq M \quad \forall \alpha \in [\mu, 1).$$

Esto permite obtener C1 (b) con $\alpha_0 := \mu$. Finalmente, definiendo

$$N := \sup_{\beta < \alpha < 1} g_\alpha(\bar{x}) \quad \text{y} \quad b(x) := \sup_{\beta < \alpha < 1} g_\alpha(x) \quad x \in X$$

y usando que $h_\alpha(\cdot) \leq g_\alpha(\cdot)$, obtenemos:

$$-N \leq h_\alpha(x) \leq b(x) \quad \forall x \in X \quad \text{y} \quad \alpha \in (\beta, 1).$$

Por tanto, C1 (c) se cumple para cualquier $\alpha_0 \in (\beta, 1)$ fijo; y C1 (b) y C1 (c) se cumplen con $\alpha_0 := \max(\mu, \beta)$. Esto completa la demostración de que C3 implica C1.

C1 implica C3: Bajo las hipótesis H y la condición C1, se ha demostrado (ver Teorema 2.2) que existe una política f^* que es CP-óptima, y tiene costo óptimo constante. Entonces definiendo $\hat{\delta} = f^*$ se obtiene C3 (a), para todo estado inicial. Por otro lado C1 da como resultado que para cada $x \in X$ y $\alpha \in [\alpha_0, 1]$:

$$\begin{aligned} g_\alpha(x) &= V_\alpha(x) - m_\alpha \\ &= h_\alpha(x) - m_\alpha \\ &\leq b(x) + N \quad [\text{por C1 (c)}]. \end{aligned}$$

Entonces, definiendo $\beta := \alpha_0$ se obtiene C3 (b). Por tanto, C1 implica C3. Esto finaliza la demostración de (a).

(b) Supóngase que C2 se cumple. Evidentemente C1(a) y C1 (b) se cumplen. De (5.2) Cap. I y de (2.1) se tiene que para cada $\alpha \in (0, 1)$:

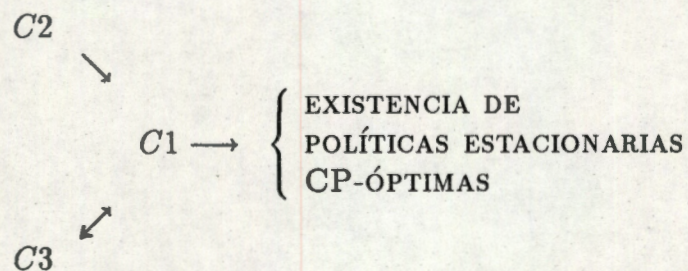
$$\begin{aligned} (1 - \alpha)V_\alpha(\bar{x}) &= V_\alpha(\bar{x}) - \alpha V_\alpha(\bar{x}) \\ &\leq c(\bar{x}, \hat{f}(\bar{x})) + \alpha \int V_\alpha(y)Q(dy|\bar{x}, \hat{f}(\bar{x})) - \alpha V_\alpha(\bar{x}) \\ &= c(\bar{x}, \hat{f}(\bar{x})) + \alpha \int h_\alpha(y)Q(dy|\bar{x}, \hat{f}(\bar{x})) \\ &\leq c(\bar{x}, \hat{f}(\bar{x})) + \int b(y)Q(dy|\bar{x}, \hat{f}(\bar{x})) \quad [\text{por C2 (c)}]. \end{aligned}$$

Esto da como resultado C1 (b) con $m := c(\bar{x}, \hat{f}(\bar{x})) + \int b(y)Q(dy|\bar{x}, \hat{f}(\bar{x}))$. Por tanto C1 se cumple con lo que se concluye la demostración de (b) y del Teorema 3.3. ■

3.4 Conclusiones

En este capítulo se han formulado tres condiciones: C1, C2 y C3, cada una dada en términos de la función de valor óptimo para el PCOCD. Los resultados obtenidos quedan resumidos como sigue:

Bajo las hipótesis H se tiene que



(“ \rightarrow ” denota implicación, y “ \leftrightarrow ” equivalencia). Además la existencia de políticas estacionarias CP-óptimas queda determinada mediante la obtención de una solución de la desigualdad de optimalidad con costo promedio.

4 La Ecuación de Optimalidad

4.1 Introducción.

4.2 La propiedad de equicontinuidad.

4.3 Condición para la ecuación de optimalidad.

4.4 Conclusiones.

4.1 Introducción

En el capítulo anterior se consideraron hipótesis (condición C1) sobre la función de valor óptimo para el PCOCD, las que permitieron asegurar la existencia de políticas estacionarias CP-óptimas mediante la obtención de una solución de la desigualdad de optimalidad con costo promedio (ver Teorema 3.2, Cap. III). Cabe observar que bajo C1 se puede tener la desigualdad *estricta*; esto ha sido mostrado con un contraejemplo en el caso numerable por R. Cavazos [6]. En este capítulo se considerará una versión más fuerte (dada en la Sección 3) de la condición C1 que garantiza la *igualdad*. Los resultados de este capítulo se encuentran en [25]. Dicho artículo extiende al caso de Borel el trabajo de Fernández-Gaucherand, Marcus y Arapostathis [9] sobre PCMs convexos.

4.2 La Propiedad de equicontinuidad

Si la condición C1 se cumple, utilizaremos la notación: $h_n(x) = h_{\alpha_n}(x)$, $n \geq 1$, $x \in X$, donde $h_\alpha(\cdot)$ está definida en (2.1), Cap. III, y $\{\alpha_n\}$, $\alpha \uparrow 1$ es la sucesión obtenida en el lema 3.1, Cap. III.

Lema 4.1. *Supóngase que la condición C1 se cumple y que la sucesión $\{h_n\}$ es equicontinua. Entonces existe una subsucesión $\{h_{n_k}\}$ que converge puntualmente a una función h continua, y dicha convergencia es uniforme en cada subconjunto compacto de X .*

DEMOSTRACIÓN. Fijemos $x \in X$ y sea $D_x := \{h_n(x) : n \geq 0\}$.

Note que $D_x \subset [-N, b(x)]$, por lo que $\bar{D}_x \subset [-N, b(x)]$, donde \bar{D}_x denota la cerradura del conjunto D_x . Como \bar{D}_x es cerrado, y $[-N, b(x)]$ es compacto

se sigue que \bar{D}_x es compacto. Por lo tanto, como x era arbitrario y $\{h_n\}$ es equicontinua, el lema se sigue del Teorema de Ascoli-Arzelá (ver Ap. B). ■

Finalizaremos esta sección con algunos ejemplos en los cuales $\{h_n\}$ resulta ser equicontinua. En cada uno de los ejemplos se supone que C1 se cumple.

Ejemplo 2.2. Si X es numerable, evidentemente $\{h_n\}$ es equicontinua.

Ejemplo 2.3. En [9] se supone que X es un subconjunto convexo y abierto de \mathbb{R}^n , y que $h_n(\cdot)$ es una función convexa. Bajo esas hipótesis se demuestra en [9] que $\{h_n(\cdot)\}$ es localmente equilipschitziana, en particular, (localmente) equicontinua.

Ahora daremos un ejemplo en el cual $\{h_n\}$ es equicontinua, pero no es localmente equilipschitziana. En particular, $h_n(\cdot)$ no es convexa. Por lo tanto, no satisface las hipótesis de [9].

Ejemplo 2.4. Sean $X = \mathbb{R}$ y $A = A(x) = \{a\}$ para cada $x \in X$. La función de costo está dada por $c(x, a) = \sqrt{x}$ si $x \geq 0$, y $c(x, a) = 0$ si $x < 0$. La ley de transición está dada por:

$$Q(\{0\} | x, a) = 1 \quad \forall x \in X.$$

Evidentemente, este ejemplo satisface las hipótesis H. Luego, por (5.2), Cap. I, tenemos:

$$V_\alpha(x) = c(x, a) + \alpha V_\alpha(0) = c(x, a), \quad x \in X.$$

Sea $\bar{x} = 0$. Entonces $h_\alpha(x) = V_\alpha(x) = c(x, a) \forall x \in X$. Por tanto, para cualquier sucesión $\alpha_n \uparrow 1$, se tiene que $h_n(x) = c(x, a) \forall x \in X$, y $n \geq 1$. De aquí se sigue que $\{h_n\}$ es equicontinua. Por otro lado, es claro que $\{h_n\}$ no es localmente equilipschitziana (considérese $x = 0$).

4.3 Condición para la ecuación de optimalidad

Consideraremos la condición siguiente:

CONDICION PARA LA ECUACION DE OPTIMALIDAD.
(CEO).

- (a) C1 se cumple;

- (b) Sea $\alpha_n \uparrow 1$ como en el lema 3.1, Cap. III, y defínase $h_n(x) := h_{\alpha_n}(x)$, $x \in X$. Se supone que $\{h_n\}$ es equicontinua;
- (c) La función b en C1 es medible y, además $\int b(y)Q(dy|x, a) < \infty$, para cada $a \in A(x)$ y $x \in X$;
- (d) X es localmente compacto.

Teorema 4.2. *Supóngase que las hipótesis H y la condición CEO se cumplen. Sea j^* como en el lema 3.1, Cap. III. Entonces existe una función h definida en X tal que:*

- (i) *La pareja (j^*, h) es solución de la ecuación de optimalidad con costo promedio, i.e.,*

$$j^* + h(x) = \min_{a \in A(x)} \left[c(x, a) + \int h(y)Q(dy|x, a) \right] \quad \forall x \in X. \quad (3.1)$$

Además

- (ii) *Existe una política estacionaria $f^* \in F$ tal que $f^*(x) \in A(x)$ minimiza el lado derecho de (3.1), f^* es CP-óptima y $J(f^*, x) = j^* \forall x \in X$.*

DEMOSTRACIÓN. (i) Por CEO (b) y el lema 4.1, existe una subsucesión $\{h_{n_k}\}$ de $\{h_n\}$ que converge puntualmente a una función continua h , i.e.,

$$h(x) := \lim_{k \rightarrow \infty} h_{n_k}(x), \quad x \in X. \quad (3.2)$$

Además, la convergencia es uniforme en subconjuntos compactos de X . Sin pérdida de generalidad (ver C1 (c)) podemos suponer $h \geq 0$. (En caso contrario, se puede considerar la función $h'(\cdot) := h(\cdot) + N \geq 0$). Denotaremos $h_k(\cdot) := h_{n_k}(\cdot)$, β_{n_k} , y $V_{\beta_k}(\cdot) : V_{\alpha_{n_k}}$. Para demostrar (3.1) verificaremos que se cumplen las desigualdades:

$$j^* + h(x) \geq \min_{a \in A(x)} \left[c(x, a) + \int h(y)Q(dy|x, a) \right], \quad x \in X, \quad (3.3.a)$$

y

$$j^* + h(x) \leq \min_{a \in A(x)} \left[c(x, a) + \int h(y)Q(dy|x, a) \right], \quad x \in X, \quad (3.3.b)$$

DEMOSTRACIÓN DE (3.3a). Usaremos argumentos similares a los de la demostración del Teorema 3.2, Cap. III. Como X es separable y localmente compacto, existe una sucesión de conjuntos abiertos G_ℓ , $\ell = 1, 2, \dots$, los cuales tienen cerradura compacta, y $\bar{G}_\ell \uparrow X$. Sea \bar{x} el estado fijo considerado en (2.1) Cap. III, y sea x un estado arbitrario, pero fijo. Escójase G_ℓ tal que $x \in G_\ell$ y nótese que $x \in G_L \forall L \geq \ell$. Del Lema 5.3, Cap. I para cada $k \geq 1$, existe $a_k \in A(x)$ tal que:

$$\begin{aligned} (1 - \beta_k)V_{\beta_k}(\bar{x}) + h_k(x) \\ = c(x, a_k) + \beta_k \int h_k(y)Q(dy|x, a_k) \quad \forall k \geq 1. \end{aligned} \quad (3.4)$$

Por otro lado, dado $\epsilon > 0$, selecciónese $T > 0$ tal que

$$(1 - \beta_k)V_{\beta_k}(\bar{x}) \leq j^* + \epsilon/3, \quad y \quad (3.5a)$$

$$h(y) - \epsilon/3 \leq h_k(y) \leq h(y) + \epsilon/3 \quad (3.5b)$$

$\forall k \geq T$ y $\epsilon \bar{G}_\ell$. Por tanto, de (3.4), (3.5) y usando que $x \in G_\ell$, $\beta_k > 0$ y $h(\cdot) \geq 0$, tenemos:

$$j^* + h(x) + \epsilon \geq c(x, a_k) + \beta_k \int I_{G_\ell}(y)h(y)Q(dy|x, a_k) \quad (3.6)$$

$\forall k \geq T$, donde I_{G_ℓ} denota la función indicadora de G_ℓ .

Definimos:

$$\begin{aligned} B_k(x) := \{a \in A(x) | c(x, a) + \beta_k \int I_{G_\ell}(y)h(y)Q(dy|x, a) \\ \leq j^* + h(x) + \epsilon\} \end{aligned}$$

y

$$\begin{aligned} B(x) := \{a \in A(x) | c(x, a) + \beta_k \int I_{G_\ell}(y)h(y)Q(dy|x, a) \\ \leq j^* + h(x) + \epsilon\}. \end{aligned}$$

Nótese que $B_k(x)$ es compacto y no vacío. Además, $B_k(x) \downarrow B(x)$, y $B(x)$ también es compacto y no vacío. Por tanto, existe una subsucesión de $\{a_k\}$ (la que también denotaremos por $\{a_k\}$ para no complicar la notación) y $a_x \in B(x)$ tal que $a_k \rightarrow a_x$ si $k \rightarrow \infty$. Usando que $I_{G_\ell}(\cdot)h(\cdot)$ es medible y acotada junto con las hipótesis H, hacemos $k \rightarrow \infty$ en (3.6) obteniendo:

$$j^* + h(x) + \epsilon \geq c(x, a_x) + \int I_{G_\ell}(y)h(y)Q(dy|x, a). \quad (3.7)$$

En (3.7) hacemos $\ell \rightarrow \infty$ y luego $\epsilon \rightarrow 0$ para obtener:

$$j^* + h(x) \geq c(x, a_x) + \int h(y)Q(dy|x, a_x);$$

por tanto,

$$j^* + h(x) \geq \min_{a \in A(x)} \left[c(x, a) + \int h(y)Q(dy|x, a) \right].$$

como x fue arbitrario, concluimos 3.3 (a).

DEMOSTRACIÓN DE (3.3b). Del lema 5.3, Cap. I, tenemos:

$$\begin{aligned} (1 - \beta_k)V_{\beta_k}(\bar{x}) + h_k(x) \\ \leq c(x, a) + \beta_k \int h_k(y)Q(dy|x, a) \quad \forall x \in X \text{ y } a \in A(x). \end{aligned} \quad (3.8)$$

Usando el lema 2.1, (3.2), CEO (a), CEO (b) y el Teorema de Convergencia Dominada, hacemos $k \rightarrow \infty$ en (3.8) obteniendo:

$$j^* + h(x) \leq c(x, a) + \int h(y)Q(dy|x, a)$$

$\forall x \in X$ y $a \in A(x)$. Esto implica (3.3b).

(ii) Como h es continua y no negativa, el Teorema 1, Ap. C, garantiza la existencia de una política estacionaria $f^* \in F$ tal que $f^*(x) \in A(x)$ minimiza el lado derecho de (3.1) i.e.,

$$j^* + h(x) = c(x, f^*(x)) + \int h(y)Q(dy|x, f^*(x)) \quad \forall x \in X. \quad (3.8)$$

Finalmente, de (3.8), y como $h \geq 0$, se sigue que f^* es CP-óptima, y que $J(f^*, x) = j^*$ (ver la demostración del Teorema 3.2 (c), Cap. III). Con esto se concluye la demostración del Teorema 4.2. ■

4.4 Conclusiones

En este capítulo se dieron hipótesis (condición CEO) que garantizan la existencia de una solución de la ecuación de optimalidad con costo promedio, que a su vez, permite obtener políticas estacionarias CP-óptimas.

5 Convergencia del Algoritmo de iteración de valores

5.1 Introducción.

5.2 Ternas canónicas.

5.3 Iteración de valores.

5.4 Construcción de una política CP-óptima.

5.5 Comentarios y conclusiones

5.1 Introducción

En este capítulo estamos interesados en estudiar la convergencia de las funciones de iteración de valores (ó aproximaciones sucesivas) hacia el costo óptimo. Este problema importante ha sido considerado por muchos autores [2,3,10,13,16,18,42] y, en particular, aquí se extiende al caso de Borel resultados de L. Sennott [37] para PCM con espacio de estados numerable y conjunto de controles finito. Las condiciones y resultados del capítulo se encuentran en [27].

En la sección 2 damos algunas preliminares y notación y en las secciones 3 y 4 se estudia la convergencia del algoritmo de iteración de valores, y la construcción de una política CP-óptima. Finalmente, en la sección 5 se dan algunos comentarios relacionados con las hipótesis consideradas en las secciones 3 y 4 así como, las conclusiones del capítulo.

5.2 Ternas canónicas

Sea $h : X \rightarrow \mathbf{R}$ una función medible dada. Definimos el *costo total esperado en n etapas con costo terminal h* como:

$$\begin{aligned} J_n(\delta, x, h) &= E_x^\delta \left[\sum_{t=0}^{n-1} c(x_t, a_t) + h(x_n) \right] \\ &= J_n(\delta, x) + E_x^\delta [h(x_n)], \end{aligned} \quad (2.1)$$

donde $J_n(\delta, x) := E_x^\delta \left[\sum_{t=0}^{n-1} c(x_t, a_t) \right]$. Se tiene también que la función de costo óptimo está dada por

$$J_n^* := \inf_{\delta} J_n(\delta, x, h). \quad (2.2)$$

Definición 5.1. Sean ρ y h funciones reales definidas en X y sea $f^* \in F$ una política estacionaria. Diremos que la terna (ρ, h, f^*) es una *terna canónica* si para cada $x \in X$ y $n \geq 0$:

$$J_n(f^*, x, h) = J_n(x, h) = n\rho(x) + h(x). \quad (2.3)$$

Una política estacionaria es *canónica* si forma parte de alguna terna canónica.

Se tiene el resultado siguiente el cual se incluye por completez del trabajo.

Lema 5.2. (ρ, h, f^*) es una terna canónica si y sólo si, para cada $x \in X$:

- (a) $\rho(x) = \inf_{a \in A(x)} \int \rho(y) Q(dy|x, a)$;
- (b) $\rho(x) + h(x) = \inf_{a \in A(x)} \left[c(x, a) + \int h(y) Q(dy|x, a) \right]$,
- (c) $f^*(x) \in A(x)$ alcanza el mínimo en (a) y (b), i.e.

$$\rho(x) = \int \rho(y) Q(dy|x, f^*), \quad (2.4)$$

$$\rho(x) + h(x) = c(x, f^*(x)) + \int h(y) Q(dy|x, f^*(x)). \quad (2.5)$$

DEMOSTRACIÓN. Ver, e.g., [21, Secc. 5.2]. ■

Nota 5.3 (a) Con referencia a (2.2) es bastante conocido (ver [3]) que:

$$J_{n+1}^*(x, h) = \inf_{a \in A(x)} \left[c(x, a) + \int J_n^*(y, h) Q(dy|x, a) \right], \quad x \in X. \quad (2.6)$$

(b) Si se itera (2.5) n veces y se usa (2.4) obtenemos:

$$n\rho(x) + h(x) = J_n(f^*, x) + E_x^{f^*} h(x_n) = J_n(f^*, x, h), \quad (2.7)$$

$\forall x \in X, n \geq 0.$

Definición 5.4 Definimos las funciones de iteración de valores como:

$$\begin{aligned} u_0(\cdot) &\equiv 0, \text{ y} \\ u_n(x) &:= J_n^*(x, 0) = \inf_{\delta} J_n(\delta, x), \quad n \geq 1, \quad x \in X \end{aligned} \quad (2.8)$$

Por (2.6) tenemos:

$$u_n(x) = \inf_{a \in A(x)} \left[c(x, a) + \int u_{n-1}(y) Q(dy | x, a) \right], \quad n \geq 1, \quad x \in X. \quad (2.9)$$

Nota 5.5. Sea $J^* := \inf_x \inf_{\delta} J(\delta, x)$. Nótese que si (J^*, h, f^*) es una terna canónica, entonces mediante (2.7), para cada $x \in X$ y $n \geq 0$:

$$nj^* + h(x) = J_n(f^*, x) + E_x^{f^*} h(x_n), \quad (2.10)$$

y de (2.8) obtenemos

$$nj^* + h(x) \geq u_n(x) + E_x^{f^*} h(x_n). \quad (2.11)$$

Las relaciones (2.8)-(2.11) se usarán repetidamente en las secciones siguientes.

Considérese la condición CEO dada en el capítulo anterior. El enunciado del lema siguiente consiste en reescribir el Teorema 4.2, Cap. IV, en la notación de esta sección.

Lema 5.4. *Supóngase que las hipótesis H y CEO se cumplen. Sea j^* la constante obtenida en el lema 3.1, Cap. III. Entonces existe una función continua h definida en X tal que:*

$$-N \leq h(x) \leq b(x) \quad \forall x \in X, \quad h(\bar{x}) = 0, \quad (2.12)$$

y una política estacionaria $f^ \in F$ tal que (j^*, h, f^*) es una terna canónica. En particular, por el lema 5.2 (b), (c), tenemos que se cumple:*

$$j^* + h(x) = \min_{a \in A(x)} \left[c(x, a) + \int h(y) Q(dy | x, a) \right] \quad (2.13)$$

$$= c(x, f^*(x)) + \int h(y) Q(dy | x, f^*(x)), \quad x \in X. \quad (2.14)$$

5.3 Iteración de valores (IV)

Considérese las funciones definidas en (2.8). Sean

$$v_n(x) := u_n(x) - u_n(\bar{x}), w_n(x) := u_n(x) - u_{n-1}(x), \quad x \in X, \quad (3.1)$$

donde $\bar{x} \in X$ es un estado fijo. El objetivo del algoritmo de IV (ó aproximaciones sucesivas) es demostrar que:

$$v_n(x) \rightarrow h(x) \quad \text{y} \quad w_n(x) \rightarrow j^*, \quad x \in X, \quad (3.2)$$

donde (j^*, h) es una solución de la ecuación de optimalidad (ver (2.13)).

En esta sección daremos condiciones que garantizan (3.2).

Nota 5.5. En la definición 3.1, Cap. I, se puede considerar también el concepto de política *markoviana*, i.e. se dice que una política es markoviana si existe una sucesión $\{f_t\}$ de funciones $f_t \in F$ tal que $\delta_t(\cdot|h_t)$ está concentrada en $f_t(x_t) \in A(x_t) \forall h_t \in H_t$ y $t \geq 0$. Nótese que, en particular, si $f_t = f \in F \forall t \geq 0$, se obtiene una política estacionaria.

Cuando usamos una política $\delta = \{f_0, f_1, \dots\}$ markoviana, el proceso de estados $\{x_t\}$ resulta ser una cadena de Markov con probabilidades de transición en un paso, definidas por:

$$P^\delta(x_{t+1} \in B | x_t = x) = Q(B|x, f_t) \quad t \geq 0, B \in \mathcal{B}(X).$$

Las probabilidades de transición en n pasos, $n \geq 2$, i.e., $P^\delta(x_{t+n} \in B | x_t = x)$, $n \geq 2$ están dadas recursivamente por:

$$\begin{aligned} Q^n(B|x, (f_t, f_{t+1}, \dots, f_{t+n-1})) \\ = \int Q^{n-1}(B|y, (f_{t+1}, \dots, f_{t+n-1}))Q(dy|x, f_t). \end{aligned}$$

Para una política estacionaria $f \in F$, las probabilidades de transición son homogéneas, i.e.,

$$Q^n(B|x, f) = \int Q^{n-1}(B|y, f)Q(dy|x, f) \quad n \geq 1,$$

con $Q^0(\cdot|x, f)$ la medida de Dirac concentrada en x . Usaremos indistintamente $Q(\cdot|x, f)$ ó $Q(\cdot|x, f(x))$.

CONDICION PARA LA CONVERGENCIA DEL ALGORITMO DE IV (CIV).

Supongamos que las hipótesis **H** y CEO se cumplen. Sea $\{u_n\}$ la sucesión de IV y sea f^* la política CP-óptima obtenida en el lema 2.6. Como condiciones suponemos:

- (a) $\{u_n\}$ es equicontinua;
- (b) Existe una medida de probabilidad P definida en $\mathcal{B}(X)$, independiente de $x \in X$, tal que:

$$Q^n(\cdot|x, f^*(x)) \xrightarrow{d} P \quad \forall x \in X, \quad (3.3)$$

(i.e., $\{Q^n(\cdot|x, f^*(x))\}$ converge débilmente a P), y $P(N) > 0$ para cada vecindad abierta N de x ;

- (c) Sea b la función CEO, y para cada $n \geq 1$, sea $f_n \in F$ un minimizador del lado derecho de (2.9) (su existencia está garantizada por el Teorema 1, Ap. C), i.e.,

$$u_n(x) = c(x, f_n(x)) + \int u_{n-1}(y)Q(dy|x, f_n(x)) \quad \forall x \in X. \quad (3.4)$$

Entonces existe una función $L : X \rightarrow \mathbb{R}$ tal que:

$$\int b(y)Q^n(dy|x, (f_n, \dots, f_1)) \leq L(x) \quad \forall x \in X, \quad n \geq 1. \quad (3.5)$$

A continuación presentamos el resultado más importante de esta sección.

Teorema 5.6. *Supóngase que la condición CIV se cumple. Entonces el algoritmo de IV converge, i.e., (3.2) se cumple, y la convergencia es uniforme en subconjuntos compactos de X .*

Para demostrar el Teorema 5.6 nos basaremos en los lemas 5.7 y 5.8, los cuales requieren de la definición siguiente. Sea (j^*, h, f^*) la terna canónica cuya existencia garantiza el lema 5.2. Definimos:

$$e_n(x) := nj^* + h(x) + N - u_n(x), \quad n \geq 0, \quad x \in X. \quad (3.6)$$

Lema 5.7 *Para cada $x \in X$, $n, m \geq 0$,*

(a) $e_n(x) \geq 0$;

(b) $\int e_m(y)Q^m(dy|x, f^*(x)) \leq e_{n+m}(x)$.

DEMOSTRACIÓN. (a) Dado que $h(\cdot) \geq -N$, de (2.11) obtenemos:

$$nj^* + h(x) \geq -u_n(x) - N \quad \forall x \in X, \quad n \geq 0,$$

de donde concluimos (a).

(b) Esto es trivial en los casos:

- $n \geq 1, m = 0$;
- $n = 0, m \geq 0$, (ver (2.11) y nótese que

$$E_x^{f^*} h(x_n) = \int h(y)Q^n(dy|x, f^*(x)) \quad \forall x \in X, \quad n \geq 0).$$

Ahora consideremos $n \geq 1$ fijo. Demostraremos (b) mediante inducción sobre m . De (2.9) tenemos:

$$u_{n+1}(x) \leq c(x, f^*(x)) + \int u_n(y)Q(dy|x, f^*(x)) \quad \forall x \in X,$$

y también, de (2.14), tenemos:

$$j^* + h(x) - u_{n+1}(x) \geq \int [h(y) - u_n(y)]Q(dy|x, f^*(x)),$$

lo cual es equivalente a:

$$\int e_n(y)Q(dy|x, f^*(x)) \leq e_{n+1}(x) \tag{3.7}$$

como $n \geq 1$ fue arbitrario, concluimos que si $n \geq 1$ y $m = 1$, (b) se cumple. Finalmente si (b) se cumple para algún $m \geq 1$, entonces

$$\begin{aligned} & \int e_n(y)Q^{m+1}(dy|x, f^*(x)) \\ &= \int \int e_n(y)Q^m(dy|z, f^*(z))Q(dz|x, f^*(x)) \\ &\leq \int e_{n+m}(z)Q(dz|x, f^*(x)) \\ &\leq e_{n+m+1}(x) \quad [\text{por (3.7)}]. \end{aligned}$$

con esto se termina la demostración. ■

Lema 5.8 (a) *La sucesión $\{e_n\}$ es equicontinua y puntualmente acotada;*

(b) *Existe una subsucesión $\{e_{n_i}\}$ de $\{e_n\}$ y una constante k tal que $\lim_{i \rightarrow \infty} e_{n_i}(x) = k \forall x \in X$, y esta convergencia es uniforme sobre subconjuntos compactos de X ;*

(c) *Cada subsucesión de $\{e_n\}$ tiene una (sub-) subsucesión que converge a k uniformemente sobre compactos y, por tanto, $\{e_n\}$ también converge a k uniformemente sobre compactos.*

DEMOSTRACIÓN. (a) De (3.6) tenemos:

$$|e_n(x) - e_n(y)| \leq |h(x) - h(y)| + |u_n(x) - u_n(y)|,$$

$\forall x, y \in X, n \geq 0$. Por tanto la equicontinuidad de $\{e_n\}$ se sigue de CIV (a) y de la continuidad de h (ver lema 5.4). Ahora de (2.13) y (3.4) tenemos que para $x \in X$ y $n \geq 1$:

$$\begin{aligned} j^* + h(x) &\leq c(x, f_n(x)) + \int h(y)Q(dy|x, f_n(x)) \\ &= u_n(x) + \int [h(y) - u_{n-1}(y)]Q(dy|x, f_n(x)), \end{aligned}$$

lo cual es equivalente a:

$$e_n(x) \leq \int e_{n-1}(y)Q(dy|x, f_n(x)). \quad (3.8)$$

Iterando (3.8) tenemos:

$$e_n(x) \leq \int e_0(y)Q^n(dy|x, (f_n, f_{n-1}, \dots, f_1)).$$

Como $e_0(x) := h(x) + N \leq b(x) + N$, (3.5) y el lema 5.7 (a) dan como resultado:

$$0 \leq e_n(x) \leq L(x) + N \quad \forall x \in X, \quad (3.9)$$

lo que finaliza la demostración de (a).

(b) A partir de (a) y del Teorema de Ascoli-Arzelá (ver Ap. B), existe una subsucesión $\{e_{n_i}\}$ de $\{e_n\}$ que converge uniformemente sobre compactos a una función continua $\theta : X \rightarrow \mathbf{R}$. En el lema 5.7 (b), tómesese $m \rightarrow \infty$ a través de valores tales que $e_{n+m}(x) \rightarrow \theta(x)$. Entonces CIV (b) nos lleva a:

$$\begin{aligned} \theta(x) &\geq \liminf_{m \rightarrow \infty} \int e_n(y) Q^m(dy|x, f^*(x)) \\ &\geq \int e_n(y) P(dy) \end{aligned} \quad (3.10)$$

lo cual implica, por el lema de Fatou (recuerde que $e_n(\cdot) \geq 0$), que:

$$\theta(x) \geq \liminf_{i \rightarrow \infty} \int e_{n_i}(y) P(dy) \geq \int \theta(y) P(dy),$$

i.e., $\theta(x) \geq \int \theta(y) P(dy)$. Puesto que esta última desigualdad se cumple para cada $x \in X$, definiendo $k := \inf_x \theta(x)$ se obtiene:

$$k \geq \int \theta(y) P(dy) \geq k,$$

i.e.,

$$\int \theta(y) P(dy) = k,$$

ó, equivalentemente:

$$\int (\theta(y) - k) P(dy) = 0,$$

lo que implica (obsérvese que $\theta(\cdot) - k \geq 0$):

$$\theta(x) = k \quad P - \text{casi seguramente.}$$

Sea $B := \{x \in X | \theta(x) > k\}$ y nótese que, como θ es continua, B es abierto. Por tanto, de CIV (b) se concluye que B es vacío, i.e., $\theta(x) = k, \forall x \in X$. Con esto se termina la demostración de (b).

(c) Sea $\{e_{n_j}\}$ una subsucesión de $\{e_n\}$. Entonces de la misma manera que en la demostración de (b), existe una subsucesión de $\{e_{n_j}\}$, la cual

denotaremos nuevamente como $\{e_{nj}\}$, que converge uniformemente sobre compactos a una constante k' , si $j \rightarrow \infty$. Por tanto, de (3.10) y el lema de Fatou, tenemos que $k \geq k'$. Un argumento similar demuestra que $k \leq k'$. De donde $k = k'$. Finalmente, supóngase que la sucesión $\{e_n\}$ no converge uniformemente sobre compactos a k . Entonces existen: un compacto $K \subset X$, $\epsilon > 0$, una subsucesión $\{e_{n_\ell}\}$ de $\{e_n\}$, y una sucesión $\{y_\ell\}$ en K tal que:

$$|e_{n_\ell}(y_\ell) - k| \geq \epsilon \quad \text{si } \ell \geq 1.$$

Luego, cada subsucesión de $\{e_{n_\ell}\}$ no converge uniformemente sobre compactos a k , lo cual contradice lo demostrado al inicio de este inciso. Esto implica que $\{e_n\}$ converge uniformemente sobre compactos a k . Con esto se concluye la demostración de (c). ■

DEMOSTRACIÓN DEL TEOREMA 5.6. Nótese que $\forall x \in X$ y $n \geq 1$:

$$w_n(x) = j^* - e_n + e_{n-1}(x), \quad \text{y}$$

$$v_n(x) = h(x) - e_n(x) + e_n(\bar{x}).$$

Por lo tanto, del lema 3.5 se sigue el teorema 5.6. ■

Terminaremos esta sección con un corolario al Teorema 5.6.

Primero daremos la definición siguiente.

Definición 5.9 Una política $\delta^* \in \Delta$ es llamada:

(a) *fuertemente óptima* (ó asintóticamente óptima) si

$$J(\delta^*, x) \leq \liminf n^{-1} J_n(\delta, x) \quad \forall \delta \in \Delta, x \in X; \quad (3.11)$$

(b) *F-fuertemente óptima* (ó fuertemente óptima en el sentido de Flynn)

si

$$\lim_{n \rightarrow \infty} n^{-1} [J_n(\delta^*, x) - u_n(x)] = 0 \quad \forall x \in X. \quad (3.12)$$

Nótese que, como $J_n(\delta, \cdot) \geq u_n(\cdot) \quad \forall \delta \in \Delta$ y $n \geq 0$, (3.12) es equivalente a:

$$\limsup_{n \rightarrow \infty} n^{-1} [J_n(\delta_n^*, x) - u_n(x)] = 0 \quad \forall x \in X.$$

Corolario 5.10 *Bajo las hipótesis del teorema 5.6:*

- (a) $\lim_{n \rightarrow \infty} u_n(x)/n = j^* \quad \forall x \in X$;
- (b) f^* es F -fuertemente óptima;
- (c) $J(f^*, x) := \limsup_{n \rightarrow \infty} n^{-1} J_n(f^*, x) = \lim_{n \rightarrow \infty} n^{-1} J_n(f^*, x) = j^*, \quad \forall x \in X$.
- (d) f^* es fuertemente óptima.

DEMOSTRACIÓN. (a) Esto es consecuencia de (3.6) y del lema 5.8 (c).

(b) De (3.6) y (2.10):

$$\begin{aligned} e_n(x) &= J_n(f^*, x) + E_x^{f^*} h(x_n) + N - u_n(x) \\ &\geq J_n(f^*, x) - u_n(x) \quad [\text{por (2.12)}]. \end{aligned}$$

Entonces (b) se sigue del lema 5.8 (c).

Finalmente, (c) se sigue de (a) y (b), mientras que (d) se sigue de (a), (b) y (c) pues:

$$\begin{aligned} J(f^*, x) &= \lim_{n \rightarrow \infty} \frac{u_n(x)}{n} = \liminf u_n(x)/n \\ &\leq \liminf_{n \rightarrow \infty} \frac{J_n(\delta, x)}{n} \quad \forall \delta \in \Delta \text{ y } x \in X. \end{aligned}$$

■

Es claro que la optimalidad fuerte y la F -optimalidad implican CP-optimalidad, pero el recíproco en general no se cumple (ver [19]). Es decir, las hipótesis CIV conducen a criterios de optimalidad más fuertes.

5.4 Construcción de una política CP-óptima

En esta sección demostraremos el resultado siguiente.

Teorema 5.11 *Supongamos que las hipótesis del Teorema 5.6 se cumplen. Además, supongamos que el conjunto A de controles es localmente compacto, y que existe una constante K tal que:*

$$v_n(x) \geq -K \quad \forall x \in X, \quad n \geq 0, \quad (4.1)$$

donde $v_n(\cdot)$ está definida en (3.1). Sean $f_n \in F$ como en (3.4). Entonces existe una política estacionaria $\hat{f} \in F$ tal que $\hat{f}(x)$ es un punto de acumulación de $\{f_n(x)\}$ para cada $x \in X$ y, además \hat{f} es CP-óptima.

DEMOSTRACIÓN. A partir de la definición de $v_n(\cdot)$ y $w_n(\cdot)$ en (3.1), podemos reescribir (2.9) y (3.4) como:

$$\begin{aligned} w_n(\bar{x}) + v_n(x) &= \min_{a \in A(x)} \left[c(x, a) + \int v_{n-1}(y) Q(dy|x, a) \right] \\ &= (x, f_n(x)) + \int v_{n-1}(y) Q(dy|x, f_n(x)), \end{aligned} \quad (4.2)$$

$\forall x \in X, n \geq 1$. Sea $\hat{A} := A \cup \{\infty\}$ la compactificación en un punto de A (recuérdese que A es, por hipótesis, un espacio de Borel localmente compacto), y sea ϕ la multifunción constante de X a \hat{A} definida por $\phi(x) := \hat{A} \quad \forall x \in X$. Obviamente ϕ es medible con valores compactos. Entonces, como \hat{A} es metrizable (ver [8] p. 247), separable, y $f_n(x) \in \hat{A} \quad \forall x \in X, n \geq 1$, el Teorema 2, Ap. C, garantiza la existencia de una función medible $\hat{f} : X \rightarrow \hat{A}$ tal que para cada $x \in X, \hat{f}(x) \in \hat{A}$ es un punto de acumulación de $\{f_n(x)\}$. Para completar la demostración del teorema, debemos demostrar que:

- (a) $\hat{f} \in F$, i.e., $\hat{f}(x) \in A(x) \quad \forall x \in X$; y
- (b) \hat{f} es CP-óptima.

Como X es separable y localmente compacto, existe una sucesión de conjuntos abiertos $G(\ell), \ell \geq 1$, cada uno con cerradura $\overline{G(\ell)}$ compacta, y $G(\ell) \uparrow X$. Fijemos $x \in X$ y escójase $G(\ell)$ de forma tal que $x \in G(\ell)$

(nótese que $x \in G(\ell') \forall \ell' \geq \ell$). Sea $\epsilon > 0$ dado. Del Teorema 5.6, existe $T_\ell > 0$ tal que:

$$h(y) - \epsilon/3 \leq h_n(y) \leq h(y) + \epsilon/3 \quad \forall n \geq T_\ell; \quad y \in \overline{G(\ell)}, \quad (4.3)$$

y

$$w_n(\bar{x}) \leq j^* + \epsilon/3 \quad \forall n \geq T_\ell. \quad (4.4)$$

Entonces, como $x \in G(\ell)$, (4.3) y (4.4) nos conducen a

$$\begin{aligned} j^* + h(x) + 2/3\epsilon &\geq v_n(x) + w_n(\bar{x}) \\ &= c(x, f_n(x)) + \int v_{n-1}(y)Q(dy|x, f_n(x)) \end{aligned}$$

$\forall n \geq T_\ell$. Sumamos K a ambos lados de la desigualdad anterior usando (4.3) y que $v_{n-1}(\cdot) + K \geq 0$, para obtener

$$\begin{aligned} j^* + (h(x) + K) + \epsilon \\ \geq c(x, f_n(x)) + \int_{G(\ell)} (h(y) + K)Q(dy|x, f_n(x)) \quad \forall n \geq T_\ell. \end{aligned} \quad (4.5)$$

Definimos:

$$\begin{aligned} D(x) &:= \{a \in A(x) | c(x, a) + \int_{G(\ell)} (h(y) + K)Q(dy|x, a) \\ &\leq j^* + (h(x) + K) + \epsilon\}. \end{aligned}$$

Nótese que $f(x) \in D(x) \forall n \geq T_\ell$ y que, de la hipótesis $H(a)$, $D(x)$ es compacto. Por tanto, el complemento de $D(x)$ en \hat{A} resulta ser abierto en \hat{A} , i.e., $D(x)$ es cerrado en \hat{A} . Puesto que:

$$\lim_{i \rightarrow \infty} f_{n_i}(x) = \hat{f}(x) \quad (4.6)$$

para alguna subsucesión $n_i = n_i(x)$, entonces $\hat{f}(x) \in D(x)$. Por tanto $\hat{f}(x) \in A(x)$. Esto demuestra (a).

Para demostrar (b), en (4.5) sustituimos n por n_i y hacemos $i \rightarrow \infty$. Entonces, como $I_{G(\ell)}(\cdot)(h(\cdot) + K)$ es medible y acotada (con $I_B =$ función indicadora de B), las hipótesis H y el lema de Fatou nos llevan a:

$$\begin{aligned} j^* + (h(x) + K) + \epsilon \\ \geq c(x, \hat{f}(x)) + \int_{G(\ell)} (h(y) + K)Q(dy|x, \hat{f}(x)). \end{aligned}$$

tomando $\ell \uparrow \infty$, $\epsilon \rightarrow 0$ y cancelando la constante K obtenemos:

$$j^* + h(x) \geq c(x, \hat{f}(x)) + \int h(y)Q(dy|x, \hat{f}(x)). \quad (4.7)$$

Como x era arbitrario, se sigue que (4.7) se cumple $\forall x \in X$. De aquí, usando que $h(\cdot) \geq -N$, se tiene (ver demostración del Teorema 3.2 (c), Cap. III) que \hat{f} es CP-óptima. Esto demuestra el teorema. ■

5.5 Comentarios y conclusiones del capítulo

Comenzaremos con algunos comentarios acerca de la condición CIV.

Nota 5.12. La hipótesis CIV (a) se cumple trivialmente si X , el espacio de estados, es un conjunto numerable con la topología discreta. También se cumple en las situaciones siguientes:

- (i) En PCM's *convexas*, i.e., en los que X es un subconjunto convexo de \mathbb{R}^n y se dan condiciones para que las funciones u_n sean convexas y continuas.
- (ii) En PCM's de *Lipschitz*, i.e., PCM's en los que las u_n son funciones localmente de Lipschitz.

Nota 5.13. (3.3) se cumple trivialmente en el caso en el que el PCM es *fuertemente ergódico* en el sentido siguiente: para cada política estacionaria f existe una medida de probabilidad P_f definida en $\mathcal{B}(X)$ tal que:

$$\|Q^n(\cdot|x, f) - P_f(\cdot)\| \rightarrow 0 \text{ si } n \rightarrow \infty, \forall x \in X, \quad (5.1)$$

donde $\|\cdot\|$ es la norma de la variación total para medidas signadas finitas. Hay condiciones suficientes bien conocidas para (5.1) (ver [10,15]) ó pueden ser deducidas de resultados generales para cadenas de Markov. Por ejemplo, considere un sistema con ruido aditivo:

$$x_{t+1} = F(x_t, a_t) + \epsilon_t, \quad t \geq 0, \quad X = \mathbb{R}^d, \quad (5.2)$$

con $\{\epsilon_t\}$ una sucesión de variables aleatorias *i.i.d.* Entonces usando resultados de series de tiempo no lineales (ver [7,24]) obtenemos condiciones para que (5.1) se cumpla, y también para que $P_f(N) > 0$ siempre que $N \subset \mathbb{R}^d$ sea abierto. Esto conduce a CIV (b).

Nota 5.14. Definimos $\bar{c}(x) := \inf_{a \in A(x)} c(x, a)$, y suponga que para cada $r \geq 0$ existe un conjunto compacto $K_r \subset X$ tal que:

$$\bar{c}(x) \geq r \quad \forall x \notin K_r \quad (5.3)$$

Suponga también que se cumplen las hipótesis del lema 2.6. Entonces, del Teorema en [17] ó verificando las hipótesis del Teorema 2 en [40], existe una medida invariante P para el kernel estocástico $Q(\cdot|x, f^*)$. Si además el kernel es Harris-recurrente y aperiódico, un Teorema de Orey [29] (c.f. [15]) conduce a (5.1) para $f = f^*$ y $P_f = P_{f^*}$.

Conclusiones del capítulo

En este capítulo básicamente se propone una condición (condición CIV) que junto con las hipótesis **H**, nos dá:

- (a) Convergencia uniforme sobre compactos del algoritmo de iteración de valores;
- (b) La política CP-óptima f^* (obtenida por medio de una solución de la igualdad de optimalidad con costo promedio) es también fuertemente óptima y F-fuertemente óptima;
- (c) Si, además se tiene que el conjunto A de controles es localmente compacto, y las funciones $v_n(\cdot)$, $n \geq 1$ (ver (3.1)), están acotadas por abajo por una constante, entonces se puede construir una política estacionaria CP-óptima que resulta ser (puntualmente) un punto de acumulación de una sucesión de políticas óptimas para problemas con horizonte finito.

Cuarta Parte

EJEMPLOS
Y
CONCLUSIONES

Resumen de Hipótesis

Hipótesis H

- (a) $c = c(x, a)$ es semicontinua inferiormente e inf-compacta en \mathbb{K} ;
- (b) Q es fuertemente continuo.

Sean:

$$\begin{aligned} h_\alpha(x) &:= V_\alpha(x) - V_\alpha(\bar{x}), \quad y \\ g_\alpha(x) &:= V_\alpha(x) - m_\alpha, \\ \text{con } m_\alpha &= \inf_{x'}(x'), \quad x \in X \text{ y } \alpha \in (0, 1). \end{aligned}$$

CONDICION C1. Existen constantes no negativas N y M , una función $b : X \rightarrow \mathbb{R}$ no negativa y no necesariamente medible, y $\alpha_0 \in (0, 1)$ tal que:

- (a) $V_\alpha(x) < \infty$ para todo $x \in X$ y $\alpha \in (0, 1)$;
- (b) $(1 - \alpha)V_\alpha(\bar{x}) \leq M$ para todo $\alpha \in [\alpha_0, 1)$;
- (c) $-N \leq h_\alpha(x) \leq b(x)$ para todo $x \in X$ y $\alpha \in [\alpha_0, 1)$.

CONDICION C2. Existen una constante N no negativa, una función b definida en X , medible y no negativa, $\alpha_0 \in (0, 1)$, y una política estacionaria $\hat{f} \in F$ tal que:

- (a) $V_\alpha(x) < \infty$ para todo $x \in X$ y $\alpha \in (0, 1)$;
- (b) $h_\alpha(x) \geq -N$ para todo $x \in X$ y $\alpha \in [\alpha_0, 1)$;
- (c) $h_\alpha(x) \leq b(x)$ y $\int b(y)Q(dy|x, \hat{f}(x)) < \infty$ para todo $x \in X$ y $\alpha \in [\alpha_0, 1)$.

CONDICION C3. (a) Existe una política $\hat{\delta} \in A$ y un estado inicial $\hat{x} \in X$ tal que $J(\hat{\delta}, \hat{x}) < \infty$;

(b) Existe $\beta \in [0, 1)$ tal que: $\sup_{\beta < \alpha < 1} g_\alpha(x) < \infty$ para cada $x \in X$.

CONDICION CEO.

- (a) C1 se cumple;
- (b) Sea $\alpha_n \uparrow 1$ como en el lema 3.1, Cap. III, y defínase $h_n(x) := h_{\alpha_n}(x)$, $x \in X$. Se supone que $\{h_n\}$ es equicontinua;
- (c) La función b en C1 es medible, y $\int b(y)Q(dy|x, a) < \infty$ para cada $a \in A(x)$ y $x \in X$;
- (d) X es localmente compacto.

CONDICION CIV. Supongamos que la hipótesis H y CEO se cumplen. Sea $\{u_n\}$ la sucesión de IV, y sea f^* la política CP-óptimo obtenida en el lema 5.4 Cap. V, como condiciones suponemos:

- (a) $\{u_n\}$ equicontinua;
- (b) Existe una medida de probabilidad P definida en $\mathcal{B}(x)$, independiente de x tal que:

$$Q^n(\cdot|x, f^*(x)) \xrightarrow{d} P \quad \forall x \in X,$$

y $P(N) > 0$ para cada vecindad abierta N de x .

- (c) Sea b la función en CEO, y para cada $n \geq 1$, sean $f_n \in F$ políticas estacionarias tales que:

$$u_n(x) = c(x, f_n(x)) + \int u_{n-1}(y)Q(dy|x, f_n(x)) \quad \forall x \in X.$$

Entonces existe una función $L : X \rightarrow \mathbf{R}$ tal que:

$$\int b(y)Q^n(dy|x, (f_n, \dots, f_1)) \leq L(x) \quad \forall x \in X, n \geq 1.$$

Sea $v_n(x) := u_n(x) - u_n(\bar{x})$, $x \in X$, tenemos también las condiciones de una política CP-óptima por medio del algoritmo de IV, las cuales para referirlas las denotaremos por: CIV (d) y CIV (e).

CIV (d): El conjunto A de controles es localmente compacto;

CIV (e): Existe una constante K tal que $v_n(x) \geq -K \forall x \in X, n \geq 1$.

6 Un ejemplo: El problema lineal con costo cuadrático

6.1 Introducción.

6.2 El problema lineal con costo cuadrático.

6.3 Verificación de hipótesis.

6.1 Introducción

En [37], Sennott presenta algunos ejemplos en sistemas de colas para el caso numerable en los que todas nuestras hipótesis se cumplen. Aquí daremos un ejemplo donde también se cumplen tales hipótesis, pero X no es numerable. En las páginas (41-42), se ha incluido un resumen de las hipótesis propuestas en la tesis, esto con la idea de facilitar la lectura de este capítulo y del siguiente.

6.2 El problema lineal con costo cuadrático (PLCC)

Consideremos un sistema lineal:

$$x_{t+1} = \gamma x_t + \beta a_t + \epsilon_t, \quad t \geq 0 \quad (2.1)$$

con costo cuadrático

$$c(x, a) = qx^2 + ra^2. \quad (2.2)$$

Aquí $X = A(\cdot) = \mathbf{R}$ y suponemos las hipótesis siguientes:

Hipótesis 2.1. (a) $0 < \gamma < 1$, $\beta \neq 0$, $q, r > 0$;

(b) Las variables aleatorias $\epsilon_0, \epsilon_1, \dots$ se suponen *i.i.d.* con distribución común $\mu \sim N(0, 1)$ e independientes del estado inicial $x_0 = x$.

Es conocido (ver por ejemplo [3]) que el PLCC bajo las hipótesis 2.1. satisface lo siguiente.

6.1.

$$(i) \quad V_\alpha(x) = k(\alpha)x^2 + \frac{k(\alpha)\alpha}{1-\alpha} \quad \forall x \in \mathbf{R}, \quad \alpha \in (0, 1), \quad (2.3)$$

donde $k(\alpha)$ es la única solución positiva de la ecuación

$$k = [1 - \alpha k \beta^2 (r + \alpha k \beta^2)^{-1}] \alpha k \gamma^2 + q. \quad (2.4)$$

Además $k(\alpha) \rightarrow \bar{k}$ si $\alpha \uparrow 1$, donde \bar{k} es la única solución positiva de la ecuación obtenida de (2.4) cuando $\alpha \uparrow 1$.

(ii) Las funciones de IV u_n , $n \geq 1$, están dadas por:

$$u_n(x) := S_n x^2 + \sum_{i=0}^{n-1} S_i \quad \forall x \in X, \quad n \geq 1, \quad (2.5)$$

donde los S_n son números positivos, $n \geq 1$, $S_0 = 0$, y están definidos recursivamente por una ecuación de forma similar a la ecuación (2.4). Además, existe $S > 0$ tal que

$$S_n \rightarrow S, \quad n \rightarrow \infty. \quad (2.6)$$

(iii) Los minimizadores $f_n \in F$ en (3.4), Cap. V, y la política f^* CP-óptima están dadas por:

$$f_{n+1}(x) = -(r + S_n \beta^2)^{-1} S_n \gamma \beta x, \quad n \geq 0, \quad f_0(x) := 0, \quad x \in X, \quad (2.7)$$

$$f^*(x) = -(r + \bar{k} \beta^2)^{-1} \bar{k} \gamma \beta x, \quad x \in X. \quad (2.8)$$

6.3 Verificación de hipótesis

Para el PLCC planteado en la sección 2 las hipótesis: H (a), C1 (a), CEO (d) y CIV (e) se cumplen trivialmente. La hipótesis H (b) ya se verificó en una situación más general en el ejemplo 5.2, Cap. I. Para el resto de las hipótesis tenemos:

C1 (b) y C1 (c). Tomando $\bar{x} = 0$ en (2.3) tenemos:

$$(1 - \alpha) V_\alpha(0) = k(\alpha) \alpha \quad \forall \alpha \in (0, 1), \quad y \quad (3.1)$$

$$h_\alpha(x) = k(\alpha) x^2, \quad \forall x \in X, \quad \alpha \in (0, 1). \quad (3.2)$$

Sea $\epsilon > 0$ dado. Escójase $\alpha_0 \in (0, 1)$ tal que:

$$k(\alpha) \leq \bar{k} + \epsilon \quad \forall \alpha \in [\alpha_0, 1).$$

Por lo tanto definiendo:

$$M := (\bar{k} + \epsilon) \text{ y } b(x) := (\bar{k} + \epsilon)x^2, \quad x \in X, \quad (3.3)$$

concluimos que

$$(1 - \alpha)V_\alpha(0) \leq M, \text{ y } 0 \leq h_\alpha(x) \leq b(x) \quad \forall x \in X, \alpha \in [\alpha_0, 1).$$

Por tanto C1 (b) y C1 (c) se cumplen.

Sea $\alpha_n \uparrow 1$ la sucesión obtenida en el lema 3.1, Cap. III (recuérdese que C1 se cumple).

CEO (b). De (3.2) se tiene que:

$$h_n(x) := h_{\alpha_n}(x) = k(\alpha_n)x^2, \quad x \in X, \quad n \geq 1.$$

Como $k(\alpha_n) \rightarrow \bar{k}$ si $n \rightarrow \infty$, sea $m_1 > 0$ tal que:

$$0 \leq k(\alpha_n) \leq m_1 \quad \forall n \geq 1.$$

Luego,

$$\begin{aligned} |h_n(x) - h_n(y)| &\leq k(\alpha_n)|x^2 - y^2| \\ &\leq m_1|x^2 - y^2|, \quad \forall x, y \in X. \end{aligned}$$

De aquí que la equicontinuidad de $\{h_n\}$ se sigue de la continuidad de $g(x) = x^2, x \in X$.

CEO (c). De (3.3) tenemos que $b(x) := (\bar{k} + \epsilon)x^2, x \in X$. b es obviamente medible y un cálculo sencillo comprueba que

$$\int b(y)Q(dy|x, a) = (\bar{k} + \epsilon) [(\gamma x + \beta a)^2 + 1],$$

el cual es finito, para cada $x \in X$ y $a \in A(x)$. Por tanto se concluye *CEO (c)*.

CIV (a). La demostración de CIV (a) es análoga a la dada para CEO (b), usando (2.5) y (2.6)

CIV (b) poniendo $x_t = x$ y $a_t = f^*(x)$ en (2.1), y usando (2.8), obtenemos:

$$x_{t+1} = F^{t+1}x + \sum_{j=0}^t F^j \epsilon_{t-j}, \quad t \geq 0. \quad (3.3)$$

donde $F := \gamma r / (r + \bar{k}\beta^2)$. Nótese que $0 < F \leq \gamma < 1$. Calculando la función característica ϕ_{t+1} de x_{t+1} y tomando $t \rightarrow \infty$ obtenemos:

$$\lim_{t \rightarrow \infty} \phi_{t+1}(s) = \exp \left[-S^2 / 2(1 - F^2) \right], \quad s \in \mathbf{R}.$$

Por tanto, aplicando el Teorema de Levy (ver [1] p. 332) concluimos CIV (b) con P teniendo distribución $N(0, (1 - F^2)^{-1})$.

CIV (c). De (3.3) tenemos que $b(x) = (\bar{k} + \epsilon)x^2 = mx^2$. Sea $F_n := \gamma r / (r + S_n \beta^2)$ si $n \geq 0$, y $F_{-1} := \gamma$. Nótese que: $0 < F_n \leq \gamma < 1$. Sea:

$$L(x) := mx^2 + m(1 - \gamma^2) \quad x \in \mathbf{R}.$$

Se puede demostrar por inducción que:

$$\begin{aligned} & \int b(y) Q^n(dy | x, (f_n, \dots, f_1)) \\ &= m F_0^2 F_1^2 \dots F_{n-1}^2 x^2 + \\ & \quad m(F_0^2 \dots F_{n-2}^2 + F_0^2 \dots F_{n-3}^2 + F_0 + 1) \\ & \leq m \gamma^{2n} x^2 + m(\gamma^{2(n-1)} + \dots + \gamma^2 + 1) \\ & \leq mx^2 + m(1 - \gamma^{2n}) / (1 - \gamma^2) \\ & \leq L(x), \quad \forall x \in X, \quad n \geq 1; \end{aligned}$$

por tanto, CIV (c) se cumple.

Finalmente, nótese que, de (2.5), poniendo $\bar{x} = 0$ y $k = 0$ tenemos que:

$$\begin{aligned} v_n(x) &= u_n(x) - u_n(0) \\ &= S_n x^2 \geq 0 \quad \forall x \in X, \quad n \geq 0; \end{aligned}$$

por tanto, CIV (d) se cumple.

7 Conclusiones y problemas abiertos

- 7.1 Introducción.
- 7.2 Conclusiones.
- 7.3 Problemas abiertos.

7.1 Introducción

En este trabajo hemos propuesto una serie de condiciones que, por medio del enfoque del factor de descuento desvaneciente, dan respuesta al problema 2.4, planteado en el capítulo II.

Este capítulo consta de tres secciones: ésta, la sección 2, donde se presenta una recopilación de las conclusiones obtenidas en los capítulos III, IV y V, y la sección 3 en la que se proponen algunos problemas abiertos emanados de esta tesis.

7.2 Conclusiones

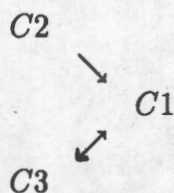
En este punto cabe observar que durante el desarrollo de la tesis siempre hemos supuesto que las hipótesis H se cumplen. Bajo este supuesto, considérense las condiciones $C1$, $C2$, $C3$ y la notación dada a continuación. Sean:

- (A) : CEO (b), CEO (c) y CEO (d);
- (B) : CIV (a), CIV (b) y CIV (c);
- (C) : CIV (d), CIV (e).

Entonces como *conclusiones* tenemos:

- (i) $C1$ implica la existencia de una solución de la desigualdad de optimalidad con costo promedio, la cual lleva a la existencia de políticas estacionarias CP-óptimas.

(ii) C1 está relacionada con C2 y C3 de la manera siguiente:



donde " \rightarrow " denota implicación y " \leftrightarrow " equivalencia;

- (iii) C1+(A) implica la existencia de una solución de la igualdad de optimalidad con costo promedio y de políticas estacionarias CP-óptimas;
- (iv) C1+(A)+(B) implica la convergencia uniforme sobre compactos del algoritmo de iteración de valores y que la política CP-óptima, obtenida a partir de C1+(A) es fuertemente óptima y F -fuertemente óptima; y
- (v) C1+(A)+(B)+(C) implica que se puede construir una política estacionaria CP-óptima, la que resulta ser un punto de acumulación (puntual) de una sucesión de políticas óptimas para problemas con horizonte finito.

7.3 Problemas abiertos

Tenemos los problemas siguientes:

(i) Sabemos que C1 y C3 son equivalentes y que C2 implica C1 (ver teorema 3.1, Cap. III); la pregunta abierta es: ¿C1 \Rightarrow C2?

Existen otras condiciones bajo las cuales se ha estudiado el problema 2.4, Cap. II. A continuación describiremos dos de ellas, las que denotaremos por C4 y C5.

Para cada $\alpha \in (0, 1)$ sea $\delta_\alpha \in \Delta$ una política dada. Definimos:

$$H_\alpha(T, x) := J(\delta_\alpha, x) - \inf_{t \geq T} (t+1)^{-1} E_x^{\delta_\alpha} \left[\sum_{k=0}^t c(x_k, a_k) \right],$$

donde $T \geq 0$ y $x \in X$. Consideremos:

CONDICION 4 (C4) [23]. Existe una sucesión de factores de descuento $\alpha_n \uparrow 1$, y políticas δ_{α_n} y δ tales que, para cada n , δ_{α_n} es α_n -óptima y

$$(a) \limsup_{n \rightarrow \infty} J(\delta_{\alpha_n}, x) \geq J(\delta, x) \quad \forall x \in X.$$

$$(b) \lim_{T \rightarrow \infty} \sup_n H_{\alpha_n}(T, x) = 0 \quad \forall x \in X.$$

CONDICION 5. (C5) [12].

(a) Existe una política $\hat{\delta}$ y un estado inicial \hat{x} tal que $J(\hat{\delta}, \hat{x}) < \infty$;

(b) Existen sucesiones de compactos $A_n \uparrow A$ y $X_n \uparrow A$ tales que: $X_n \times A_n$ es un subconjunto de \mathbb{K} , y

$$\liminf_{n \rightarrow \infty} \{c(x, a) | (x, a) \notin X_n \times A_n\} = \infty.$$

En [28] se han dado contraejemplos que garantizan que:

- C4 \Rightarrow Ci para $i=1,2,3$;
- C5 \Rightarrow Ci para $i=1,2,3$;
- C5 $\not\Rightarrow$ C4;
- Ci $\not\Rightarrow$ C5 para $i=1,2,3,4$,

quedando como problema abierto:

(ii) ¿ C1 \Rightarrow C4?, y ¿ C2 \Rightarrow C4?

En algunos trabajos previos (por ejemplo en [10], Cap. III, donde se considera el PCOCP teniendo la función costo acotada), se ha resuelto el PCOCP imponiendo condiciones de ergodicidad (ver sección 3.3 en [10]) en la ley de transición Q , i.e., se han impuesto condiciones en el modelo de control (X, A, Q, c) para el estudio del PCOCP. Puesto que nuestras hipótesis están dadas en función del PCOCD, queda como problema abierto:

(iii) Plantear condiciones sobre las componentes del modelo control: (X, A, Q, c) que permitan analizar el problema 2.4, formulado en el capítulo II.

Apéndice

A. Notación y Terminología

1. $\mathcal{B}(W)$: σ -álgebra de Borel de un espacio topológico W , i.e., la σ -álgebra más pequeña que contiene los conjuntos abiertos W .
2. *El espacio de Borel*: un subconjunto de Borel de un espacio métrico, separable y completo.
Sean X y Z espacios de Borel.
3. *Kernel estocástico*: un kernel estocástico sobre X dada Z es una función $Q(\cdot|\cdot)$, ó $Q(dx|z)$, tal que:
 - (a) $Q(\cdot|z)$ es una medida de probabilidad en X para cada $z \in Z$,
y
 - (b) $Q(B|\cdot)$ es una función medible sobre Z para cada $B \in \mathcal{B}(X)$.
4. *Abreviaturas*:
 - CEO: condición para la ecuación de optimalidad.
 - CIV: condición para la convergencia del algoritmo de iteración de valores.
 - CP-óptimo: costo-promedio óptimo.
 - PCOCD: el problema de control óptimo con costo total descontado.
 - PCOCP: el problema de control óptimo con costo promedio.
 - PLCC: el problema lineal con costo cuadrático.
 - PCM: proceso de control de Markov.

B. Algunos resultados de análisis

- *Teorema Tauberiano*. Sea $\{c_n\}$ una sucesión de números reales no negativos y $\alpha \in (0, 1)$. Entonces:

$$\begin{aligned} \liminf_{N \rightarrow \infty} N^{-1} \sum_{m=0}^{N-1} c_m &\leq \liminf_{\alpha \uparrow 1} (1 - \alpha) \sum_{n=0}^{\infty} \alpha^n c_n \\ &\leq \limsup_{\alpha \uparrow 1} (1 - \alpha) \sum_{n=0}^{\infty} \alpha^n c_n \leq \limsup_{N \rightarrow \infty} N^{-1} \sum_{m=0}^{N-1} c_m. \end{aligned}$$

La demostración de este teorema se puede encontrar en [39].

- *El Teorema de Ascoli-Arzelá.* Sean Y un espacio métrico separable, Z un espacio métrico y \mathcal{F} una familia equicontinua de funciones de Y a Z . Sea $\{g_n\}$ una sucesión en \mathcal{F} tal que: para cada x en Y la cerradura del conjunto $\{g_n(x) : 0 \leq n < \infty\}$ es un conjunto compacto. Entonces existe una subsucesión $\{g_{n_k}\}$ que converge puntualmente a una función continua g , y dicha convergencia es uniforme en subconjuntos compactos de Y .

La demostración de este Teorema se puede encontrar en [33] pág. 169.

C. Teoremas de selección medible

- *Teorema 1.* Sea (X, A, Q, c) un modelo de control, y supóngase que las hipótesis **H** se cumplen. Sea $u : X \rightarrow \mathbf{R}$ una función medible y no negativa. Entonces la función \bar{u} definida como:

$$\bar{u}(x) := \inf_{a \in A(x)} \left[c(x, a) + \int u(y)Q(dy|x, a) \right], \quad x \in X$$

es medible, y existe $f \in F$ tal que:

$$\bar{u}(x) = c(x, f(x)) + \int u(y)Q(dy|x, f(x)) \quad \forall x \in X.$$

La demostración de este Teorema se puede encontrar en [30].

- *Teorema 2.* Sean X un espacio de Borel y A un espacio métrico separable. Sea $\phi : X \rightarrow A$, una multifunción Borel-medible y con valores compactos. Si $\{f_n\}$ es una sucesión de selectores medibles para ϕ (i.e. para cada $n \geq 1$, $f_n : X \rightarrow A$ es una función medible tal que $f_n(x) \in \phi(x) \forall x \in X$), entonces existe un selector medible f^* para ϕ tal que, para cada $x \in X$, $f^*(x) \in \phi(x)$ es un punto de acumulación de $\{f_n(x)\}$.

La demostración de este Teorema se puede encontrar en [34].

Bibliografía

- [1] Ash, R.B.: *Real Analysis and Probability*, Academic Press, New York, 1972.
- [2] Arapostathis A., Borkar, V., Fernández-Gaucherand, E. Ghosh, M.K. and Marcus, S.I.: *Controlled Markov processes with an average cost criterion: a survey*, SIAM J. Control Optim. 31 (1993), 282-344.
- [3] Bertsekas, D. P.: *Dynamic Programming: Deterministic and Stochastic Models*, Prentice-Hall, Englewood Cliffs, N.J., 1987.
- [4] Cavazos-Cadena, R. *Weak conditions for the existence of optimal stationary policies in average Markov decision chains with unbounded costs*. Kybernetika, 25 (1989), 145-156.
- [5] Cavazos-Cadena, R., *Solution to the optimality equation in a class of average Markov decision chains with unbounded cost*. Kybernetika (Prague), 27 (1991), 23-37.
- [6] Cavazos-Cadena R., *A counterexample on the optimality equation Markov decision chains with average cost criterion*, System Control Lett. 16 (1991) 387-392.
- [7] Diebolt, J. and Guégan, D. *Probabilistic properties of the general nonlinear markovian process of order one and applications to time series modeling*, Rapport Technique No. 125, Laboratoire de Statistique Théorique et Appliqué, CNRS-URA 1321, Université Paris VI, 1990.
- [8] Dugundji J. *Topology*, Allyn and Bacon, Boston, 1966.
- [9] Fernández-Gaucherand, E., Marcus, S.I. and Arapostathis, A. *Convex stochastic control problem*, Proc. 31st. IEEE-CDC, Tucson Az., Dec. 1992.
- [10] Hernández-Lerma, O, *Adaptive Markov Control Processes*, Springer-Verlag, New York, 1989.
- [11] Hernández-Lerma, O., *Average optimality in dynamic programming on Borel spaces unbounded costs and controls*, Syst. Control Lett. 17 (1991), 237-242.

- [12] Hernández-Lerma O., *Existence of average optimal policies in Markov control processes with strictly unbounded costs*, Kybernetika (Prague) 29 (1993), 1-17.
- [13] Hernández-Lerma O. and Marcus, S.I., *Discretization procedures for adaptive Markov control processes*, J. Math. Anal. Appl., 137 (1989), 485-514.
- [14] Hernández-Lerma O and J.B. Lasserre, *Average cost optimal policies for Markov control processes with Borel state space and unbounded costs*, Syst, Control Lett. 15 (1990), 349-356.
- [15] Hernández-Lerma O, Montes de Oca R. and Cavazos-Cadena, R., *Recurrence conditions for Markov decision processes with Borel state space, a survey*, Ann. Oper. Res. 28 (1991), 29-46.
- [16] Hernández-Lerma, O. and Muñoz de Ozak, M., *Discrete-time Markov control processes with discontinued unbounded costs: optimality criteria*, Kybernetika (Prague) 28 (1992), 191-212.
- [17] Hernández-Lerma, O. and Lasserre, J.B., *Invariant probabilities for Feller-Markov chains*, Reporte Interno No. 122, CINVESTAV-IPN, México, 1993. (Submitted.)
- [18] Hernández-Lerma, O. and Runggaldier, J.W., *Monotone approximations for convex stochastic control problems*, J. Math. Syst., Estimation, and Control 4 (1994), 99-140.
- [19] Hernández-Lerma, O. and Lasserre, J.B. *Linear programming and average optimality of Markov control processes on Borel spaces-unbounded costs*, SIAM J. Control Optim. 32 (1994), 480-500.
- [20] Hernández-Lerma O. and J.B. Lasserre, *Average optimality in Markov control processes via discounted cost problems and linear programming* Report No. 93060, LAAS-CNRS, Toulouse, 1993, (submitted).
- [21] Hernández-Lerma O. and Lasserre J.B. *Discrete-time Markov Control Processes*. (Book in preparation).
- [22] Hinderer, K. *Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter*, Lecture Notes Oper. Res. 33, Springer, New York, 1970.

- [23] A. Hordijk and F.A. Van der Duyn Schouten, *Average optimal policies in Markov decision drift processes with applications to a queuing and replacement model*, Adv. Appl. Prob. 15 (1983), 274-303.
- [24] Mokkadem, A. Sur un modèle autorégressif nonlinéaire: ergodicité et ergodicité géométrique, J. Time Series Anal. 8 (1987), 195-205.
- [25] R. Montes-de-Oca, R. *The average cost optimality equation for Markov control processes on Borel spaces*, Syst. Contr. Lett., to appear.
- [26] 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), 1-19.
- [27] R. Montes-de-Oca and O. Hernández-Lerma, *Value iteration in average cost Markov control processes on Borel spaces*, Reporte interno No. 134, Departamento de Matemáticas CINVESTAV-IPN, 1993, (submitted).
- [28] R. Montes-de-Oca, J.A. Minjarez Sosa and O. Hernández-Lerma, *Conditions for average optimality in Markov control processes on Borel spaces*, Reporte interno No. 145, Depto. de Matemáticas, CINVESTAV-IPN, 1993, to appear, Boletín de la Sociedad Matemática Mexicana.
- [29] Orey, S., *Limit Theorems for Markov Chains Transition Probabilities*, Van Nostrand Reinhold, London, 1971.
- [30] U. Rieder, *Measurable selection theorems for optimization problems*, Manuscripta Mathematica 24 (1978), 11-131.
- [31] R. K. Ritt and L.I. Sennott, *Optimal stationary policies in general state Markov decision chains with finite action set*, Mathematics of Operations Research 17, (1992), 901-909.
- [32] D.R. Robinson, *Markov decision chains with unbounded costs and applications to the control of queues*, Adv. Appl. Prob. 8 (1976), 159-176.
- [33] Royden, H.L., *Real Analysis*, 3rd. Ed., Macmillan, New York, 1988.

- [34] Schäl, M., *Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal*, Zeit. Wahrs. Verw. Geb. 32 (1975), 179-196.
- [35] M. Schäl, *Average optimality in dynamic programming with general state space*, Math. Oper. Res. 18 (1993), 163-172.
- [36] L.I. Sennott, *Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs*, Oper. Res. 37 (1989), 626-633.
- [37] Sennott, L.I., *Value iteration in countable state average cost Markov decision processes with unbounded costs*, Ann. Oper. Res. 29 (1991), 261-271.
- [38] Sennott, L.I., *The average cost optimality equation and critical number policies*, Prob. Eng. Inform. Sci. 7 (1993), 47-67.
- [39] R. Sznajder and J.A. Filar, *Some comments on a theorem of Hardy and Lettewood*, J. Optim. Theory Appl. 75 (1992), 201-208.
- [40] Tweedie, R.L., *Invariant measures for Markov chains with no irreducibility assumptions*, J. Appl. Prob. 25 A (1988), 275-285.
- [41] R.R. Weber and S.S. Stidham Jr., *Optimal control of service rates in networks of queues*, Adv. Appl. Prob. 15 (1987), 202-218.
- [42] White, D.J., *Dynamic programming Markov chains, and the method of successive approximations*, J. Math. Anal. Appl. 6 (1963), 373-376.