UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS
División de Estudios de Posgrado

ASPECTOS TEORICOS DE LOS PROCESOS DE CONTROL DE MARKOV

T E S I S

Que para obtener el grado académico de:
D O C T O R E N C I E N C I A S
(MATEMATICAS)

P r e s e n t a :
JUAN GONZALEZ HERNANDEZ

Director de Tesis: Dr. Onésimo Hernández Lerma

México, D.F. *1998*

Aspectos Teóricos de los Procesos de Control de Markov

Juan González Hernández

Dedico cariñosamente esta tesis

a mi padre (en memoria), por haber escogido llevar una vida jarocha y consecuente consigo mismo,

a mi madre por el amor y apoyo que siempre me ha brindado.

Agradezco

al Dr. Onésimo el haber dirigido e impulsado esta tesis y por su incanzable labor por el desarrollo del control estocástico en México,

al Dr. Miguel Ángel, el apoyo en mi formación como matemático desde la licenciatura hasta el doctorado,

a los sinodales: Dra. Fernández, Dr. Hernández, Dr. Montes de Oca, Dra. Rodrigues y Dr. Rosenbleuth, sus valiosas revisiones de la tesis,

a mis amigos los controleros (César, Guadalupe y Rubén), las motivadoras discusiones que influyeron muy positivamente en la génesis de este trabajo,

y a mi amiga Marisa por haber traducido pacientemente mis garabatos al Scientific Work Place.

Esta tesis aborda algunos aspectos teóricos de los procesos de control de Markov como son: programación lineal infinita, medidas de ocupación y procesos de control de Markov con restricciones. El contenido de la tesis es el siguiente:

El capítulo 1 comprende los conceptos básicos de los procesos de control de Markov tales como la construcción canónica de un proceso de control de Markov, los índices de funcionamiento, la ecuación de optimalidad y las ternas canónicas.

El capítulo 2, por medio de las medidas de ocupación, trata de los procesos de control de Markov en espacios polacos, con índice de funcionamiento el costo descontado, tanto el caso sin restricciones como el caso restringido. La idea de este capítulo es la de ver al índice de costo descontado como una integral de la función de costo sobre una familia de medidas, para esta familia de medidas encontrar una medida representativa (la envolvente superior), de forma tal que si ésta es tensa, entonces la familia entera es tensa. Esto junto con una ecuación que nos identifique a las medidas que provienen de políticas (medidas de ocupación) nos da una condición suficiente de solución al problema de control de Markov.

El capítulo 3, por medio de la programación lineal infinita, trabaja a los procesos de control de Markov en espacios de Borel, con índice de funcionamiento el costo promedio. Aquí también la idea es escribir al índice de funcionamiento de costo promedio como una integral sobre una familia de medidas, lo cual permita plantearlo como un problema de programación lineal.

TOPICS THEORY OF MARKOV CONTROL PROCESSES

This thesis works with some topics theory of Markov control processes such as ocupation measures, infinite linear programming and Markov control processes with restriccions. The content of the thesis is the following:

Chapter 1 contains basic concepts of Markov control processes such as canonical construction, performance criteria, optimality equation and canonical triplets.

Chapter 2, by using ocupation measures, works with Markov control processes in Polish spaces, with discounted cost as performance index, both cases the unrestricted one and the restricted one. The main idea in this chapter is to see to the discounted cost as an integral of the cost function on a family of measures, for this family to find a representative measure (upper envelope), such that if it is tight then the whole family is tight. This together with an equation which says which measures come from policies provides a sifficient condition for a solution to the Markov control problem.

Chapter 3, by using infinite linear programming, works with Markov control processes in Borel spaces, with average cost as performance index. Once again the idea is to see the average cost as an integral of the cost function on a family of measures, this allows to write it as a linear programmin problem.

INTRODUCCIÓN .

Esta tesis aborda algunos aspectos teóricos de los procesos de control de Markov como son: medidas de ocupación, programación lineal infinita y procesos de control de Markov con restricciones. El contenido de la tesis es el siguiente:

El capítulo 1 comprende los conceptos básicos de los procesos de control de Markov tales como la construcción canónica de un proceso de control de Markov, los índices de funcionamiento, la ecuación de optimalidad y las ternas canónicas.

El capítulo 2, por medio de las medidas de ocupación, trata de los procesos de control de Markov en espacios polacos, con índice de funcionamiento el costo descontado, tanto el caso sin restricciones como el caso restringido. La idea de este capítulo es la de ver al índice de costo descontado como una integral de la función de costo sobre una familia de medidas, para esta familia de medidas encontrar una medida representativa (la envolvente superior), de forma tal que si ésta es tensa, entonces la familia entera es tensa. Esto junto con una ecuación que nos identifique a las medidas que provienen de políticas (medidas de ocupación) nos da una condición suficiente de solución al problema de control de Markov. El antecedente a este enfoque es el artículo de Kurano y Kawai [32] quienes usaron en lugar de la envolvente superior otra medida definida por Rogers [35], la cual no les asegura que la familia entera de medidas fuera "compacta". Y de esta forma se hizo posible construir el contraejemplo (2.25). El capítulo abarca el siguiente material: La sección 2.2 se refiere al concepto de tensión y contiene un ejemplo de una cadena de Markov acotada en probabilidad y no tensa. En la sección 2.3 se construyen las envolventes superior e inferior de una familia de medidas. En la sección 2.4 se aplican estos conceptos a los procesos de control de Markov con costo descontado para dar condiciones suficientes para la existencia de políticas óptimas deterministas. Se complementa este enfoque con un contraejemplo a una afirmación de Kurano y Kawai (el Teorema 3.3 de [31]), quienes usaron también medidas de ocupación para tratar de dar condiciones suficientes de existencia de políticas óptimas. La sección 2.6 define los conceptos básicos de los procesos de control de Markov con restricciones. Finalmente, la sección 2.7 aplica las medidas de ocupación a estos procesos para dar condiciones suficientes para la existencia de políticas óptimas estacionarias. La limitación más fuerte de este capítulo es que no se dan métodos para encontrar, al menos aproximadamente, a las políticas óptimas y a la función de valor.

El capítulo 3. por medio de la programación lineal infinita. trabaja a los procesos de control de Markov en espacios de Borel, con índice de funcionamiento el costo promedio. Aquí también la idea es escribir al índice de funcionamiento de costo promedio como una integral sobre una familia de medidas, lo cual permite plantearlo como un problema de programación lineal. Los antecedentes de este enfoque, salvo excepciones ([22], [24], [25] y [26]), abordan solamente espacios finitos o numerables. El capítulo incluye el siguente material: La sección 3.2 presenta los fundamentos de la programación lineal infinita. La sección 3.3 muestra a los programas lineales P y su dual P^* asociados al problema de control de Markov con costo promedio para el caso unicadena. La sección 3.4 propone dos conjuntos de hipótesis que son usadas en las siguientes dos secciones. La sección 3.5, por medio del Teorema Generalizado de Farkas, da condiciones necesarias y suficientes para que el programa lineal P sea consistente (lo cual es un avance). La sección 3.6 muestra que: a) el programa lineal P es resoluble, b) las sucesiones minimizantes para P son débilmente precompactas y c) si existe una sucesión maximizante para P^*, acotada en la norma con peso w_0, entonces el programa dual P^* es resoluble y se satisface la dualidad fuerte, más aún, la ecuación de optimalidad tiene solución. La limitación más fuerte de este capítulo es la misma que la del capítulo anterior.

La mayor parte del material del capítulo 2, hasta la sección 2.5, está contenido en el artículo [19], el ejemplo de la sección 2.2, al que hicimos referencia anteriormente, está presentado en el artículo [20] y la mayor parte del contenido del capítulo 3 está comprendido en el artículo [23].

Contenido

Capítulo 1

CONCEPTOS BÁSICOS

1.1 INTRODUCCIÓN.

En este capítulo daremos la notación, la terminología y los conceptos básicos de los procesos de control de Markov, que serán utilizados en el texto. Comenzando por esta sección en la que hacemos un recuento del contenido del capítulo e indicamos las convenciones sobre terminología y notación. en la sección 1.2 damos el modelo de control de Markov, los conceptos relacionados y la construcción canónica de un proceso de control de Markov. En la sección 1.3 mostramos los índices de funcionamiento que serán usados en este trabajo y mostramos la ecuación de optimalidad para el costo descontado y las ternas canónicas para el costo promedio. Los conceptos relacionados con el índice de funcionamiento de costo descontado serán usados en el capítulo 2, mientras que los conceptos relacionados con el costo promedio serán usados en el capítulo 3.

Terminología y Notación

Dado un espacio topológico $(\mathbb{X}, \tau)$, denotamos por $\mathfrak{B}(\mathbb{X})$ a la σ-álgebra generada por la topología τ, y se le conoce como la σ-álgebra de Borel. Por convención, cuando en el texto nos referimos a conjuntos o funciones "medibles", significa "Borel-medibles", y cuando nos referimos a medidas éstas estarán siempre definidas en la σ-álgebra de Borel $\mathfrak{B}(\mathbb{X})$.

Si $\mathbb{X}$ y $\mathbb{Y}$ son espacios topológicos, un kérnel estocástico en $\mathbb{X}$ dado $\mathbb{Y}$ es una función $K(\cdot/\cdot)$ tal que $K(\cdot/y)$ es una medida de probabilidad en $\mathbb{X}$ para cada $y \in \mathbb{Y}$, y $K(B/\cdot)$ es una función medible en $\mathbb{X}$ para cada $B \in \mathfrak{B}(\mathbb{X})$. Si $\mathbb{X} = \mathbb{Y}$ decimos simplemente que K es un kérnel estocástico en $\mathbb{X}$.

Un espacio polaco es un espacio métrico, completo y separable y un espacio de Borel es un subconjunto medible de un espacio polaco. Por $\overline{\mathbb{R}}$ denotamos al conjunto de los números reales extendidos, $\overline{\mathbb{R}} = \mathbb{R} \bigcup \{\pm\infty\}$.

1.2 PROCESOS DE CONTROL DE MARKOV

Definición 1.1 *Modelo de control de Markov.*

Un modelo de control de Markov es una quíntupla ordenada

$$(\mathbb{E}, \mathbb{A}, \{A(e) : e \in \mathbb{E}\}, K, c)$$

que consiste de:

a) un espacio de Borel $\mathbb{E}$, llamado el espacio de estados cuyos elementos son llamados estados,

b) un espacio de Borel $\mathbb{A}$, llamado el espacio de acciones, cuyos elementos son llamados acciones,

c) una familia $\{A(e) : e \in \mathbb{E}\}$ de subconjuntos medibles de $\mathbb{A}$, no vacíos, donde $A(e)$ denota el conjunto de acciones o controles posibles cuando el sistema está en el estado e, y con la propiedad de que el conjunto de parejas ordenadas

$$\mathbb{G} := \{(e, a) : e \in \mathbb{E}, a \in A(e)\},$$

es decir, la gráfica de la multifunción $e \to A(e)$, es un subconjunto medible de $\mathbb{E} \times \mathbb{A}$,

d) un kérnel estocástico K en $\mathbb{E}$ dado $\mathbb{G}$, llamado la ley de transición, y

e) una función medible $c : \mathbb{E} \to \overline{\mathbb{R}}$, llamada la función de costo por etapa o de costo en un paso.

A las funciones medibles de $\mathbb{E}$ en $\mathbb{A}$ y tales que $f(e) \in A(e)$ para todo estado $e \in \mathbb{E}$ se les llamas selectores medibles. A lo largo de este trabajo supondremos la siguiente hipótesis.

Hipótesis 1.2

a) El conjunto $\mathbb{G}$ contiene a la gráfica de una función medible $f : \mathbb{E} \to \mathbb{A}$ tal que $f(e) \in A(e)$ para todo estado $e \in \mathbb{E}$.

b) La función de costo por etapa c es no negativa.

La razón para pedir que exista al menos un selector medible (Hipótesis 1.2 a) es para asegurar que el conjunto de políticas (Definición 1.4) es no vacío. En el inciso b) de la Hipótesis 1.2, podemos reemplazar c por una función que puede tomar valores negativos siempre y cuando sea acotada por abajo. Al conjunto de todos los selectores medibles lo denotaremos por $\mathfrak{F}$ y al conjunto de kérneles estocásticos ϕ en $\mathbb{A}$ dado $\mathbb{E}$ tales que $\phi(A(e)/e) = 1$ lo denotaremos por Φ.

Observación 1.3

Si v es una función en $\mathbb{G}$ con valores reales y ϕ es un kérnel estocástico en Φ, entonces denotaremos por

$$v(e, \phi) := \int_{\mathbf{A}} v(e, a)\, \phi(da/e).$$

En particular, para un selector medible $f \in \mathfrak{F}$ tenemos

$$v(e, f) := v(e, f(e)).$$

Historias

Denotaremos por $\mathbb{H}_t$ al conjunto de todas las historias posibles hasta el tiempo t para cada t=0,1,2,... , es decir,

$$\begin{aligned} \mathbb{H}_0 &:= \mathbb{E} \\ \mathbb{H}_t &:= \mathbb{G}^t \times \mathbb{E} = \mathbb{G} \times \mathbb{H}_{t-1} \quad \text{para } t = 1, 2, 3, ... \end{aligned}$$

Un elemento genérico de $\mathbb{H}_t$ es un vector de la forma

$$h_t = (e_0, a_0, e_1, a_1, ..., a_{t-1}, e_t)$$

donde $(e_i, a_i) \in \mathbb{G}$.

Definición 1.4 *Política, política estacionaria, política determinista*

a) Una *política de control*, o simplemente una *política*, es una sucesión $\pi = (\pi_t : t = 0, 1, 2, ...)$ de kérneles estocásticos π_t en $\mathbb{A}$ dado $\mathbb{H}_t$ que satisfacen la restricción

$$\pi_t\left(A\left(e_t\right)/h_t\right) = 1 \quad \text{para toda } h_t \in \mathbb{H}_t \tag{1.1}$$

y $t = 0, 1, 2, ...$ A estos kérneles estocásticos los llamaremos reglas de decisión. Al conjunto de todas las políticas lo denotaremos por $\mathcal{P}$.

b) Una política de control π se dice que es *estacionaria aleatorizada*, o simplemente *estacionaria*, si existe un kérnel estocástico $\phi \in \Phi$ tal que

$$\pi_t\left(\cdot/h_t\right) = \phi\left(\cdot/e_t\right)$$

para toda historia $h_t = (e_0, a_0, e_1, a_1, ..., a_{t-1}, e_t)$ y toda $t = 0, 1, 2, ...$ En este caso, escribimos

$$\pi = \phi^{\infty} = (\phi, \phi, \phi, ...) .$$

c) Una política estacionaria se dice que es *determinista* si existe un selector medible $f : \mathbb{E} \to \mathbb{A}$ $(f \in \mathfrak{F})$ tal que ϕ está concentrada en $f(e)$, y en este caso escribimos

$$\pi = f^{\infty} = (f, f, f, ...) .$$

La construcción canónica

Sea $(\Omega, \mathcal{A})$ el espacio medible que consiste del espacio muestral $\Omega = (\mathbb{E} \times \mathbb{A})^{\infty}$ y $\mathcal{A}$ es la σ-álgebra producto correspondiente. Los elementos de Ω son sucesiones de la forma $w = (e_0, a_0, e_1, a_1, ...)$ con $e_t \in \mathbb{E}$ y $a_t \in \mathbb{A}$ para $t = 0, 1, 2, ...$ Observe que Ω contiene al espacio $\mathbb{G}^{\infty}$.

Sean $\pi = (\pi_t)$ una política de control y ν una medida de probablidad en $\mathbb{E}$, conocida como la distribución inicial. Entonces por el teorema de C. Ionescu Tulcea ([26], pág 178) existe una única medida de probabilidad P_ν^π en Ω definida en la σ-álgebra producto $\mathcal{A}$ y que por (1.1) está concentrada en $\mathbb{G}^{\infty}$ y, más aún, para todo $B \in \mathfrak{B}(\mathbb{E})$, $C \in \mathfrak{B}(\mathbb{A})$ y $h_t \in \mathbb{H}_t$

$$\begin{aligned} P_\nu^\pi\left[e_0 \in B\right] &= \nu(B), \\ P_\nu^\pi\left[a_t \in C/h_t\right] &= \pi_t\left(C/h_t\right) \quad \text{y} \\ P_\nu^\pi\left[e_{t+1} \in B/h_t, a_t\right] &= K\left(B/e_t, a_t\right). \end{aligned}$$

Definición 1.5 *Proceso de control de Markov.*

El proceso estocástico $(\Omega, \mathcal{A}, P_\nu^\pi, (e_t))$ se llama un proceso de control de Markov a tiempo discreto, o proceso de decisión de Markov a tiempo discreto.

El proceso (e_t) en la Definición 1.5 depende, por supuesto, de la política π que se está usando y de la distribución inicial ν. De aquí que, en sentido estricto, deberíamos de escribir algo así como $e_t^{\pi,\nu}$ en lugar de e_t. Sin embargo, mantendremos la notación más simple e_t y siempre será claro por el contexto cual es la política π y cual es la distribución inicial ν que se está usando.

Por otro lado, por un abuso de terminología, llamaremos también, a partir de aquí, a la familia

$$\{(\Omega, \mathcal{A}, P_\nu^\pi, (e_t)) : \pi \in \mathcal{P}\}$$

un proceso de control de Markov. El operador esperanza con respecto a P_ν^π se denota por E_ν^π. Si ν está concentrada en e, entonces escribimos P_e^π y E_e^π en su lugar.

1.3 ÍNDICES DE FUNCIONAMIENTO

Consideremos el modelo de control (Definición 1.1) y sea ν una medida de probabilidad en $\mathbb{E}$, la distribución inicial.

Definición 1.6 *Costo descontado, costo promedio.*

a) Sea $\alpha \in (0,1)$. Al índice de funcionamiento de la política π

$$I_d(\pi, \nu) := E_\nu^\pi \left(\sum_{t=0}^{\infty} \alpha^t c(e_t, a_t) \right)$$

se le llama *índice de costo esperado descontado* y α es el factor de descuento. En lo sucesivo diremos simplemente índice de costo descontado o costo descontado. En el caso particular en que ν está concentrada en un estado e, denotamos al índice de costo descontado como $I_d(\pi, e)$.

b) Sea $I_N(\pi,\nu)$ el índice de funcionamiento de la política π del costo total hasta la etapa N-1, i.e.,

$$I_N(\pi,\nu) := E_\nu^\pi\left(\sum_{t=0}^{N-1} c(e_t,a_t)\right).$$

Entonces al índice de funcionamiento

$$I_p(\pi,\nu) = \limsup_{N\to\infty}\frac{1}{N}I_N(\pi,\nu)$$

se le llama *índice de costo esperado promedio.* En lo sucesivo diremos simplemente índice de costo promedio o costo promedio. En el caso particular en que ν esté concentrada en un estado inicial e denotamos a estos índices como $I_N(\pi,e)$ e $I_p(\pi,e)$, respectivamente; así

$$I_p(\pi,e) = \limsup_{N\to\infty}\frac{1}{N}I_N(\pi,e) \quad \text{para } e \in \mathbb{E}.$$

Definición 1.7 *Problema de control de Markov. Función de valor.*

Sea $I(\pi,\nu)$ un índice de funcionamiento, ya sea el de costo descontado, o el de costo total en N etapas, o el de costo promedio, y consideremos el proceso de control de Markov (Definición 1.4). Entonces el problema de control de Markov consiste en encontrar una política $\pi^* \in \mathcal{P}$ tal que

$$I(\pi^*,\nu) = \inf\{I(\pi,\nu) : \pi \in \mathcal{P}\}. \tag{1.2}$$

Si existe tal política π^*, decimos que π^* es ν-óptima ó simplemente óptima. Al lado derecho de la igualdad (1.2) lo denotamos por $I(\nu)$ y le llamamos la función de valor mínimo o simplemente la función de valor.

La familia $\{(\Omega,\mathcal{A},P_\nu^\pi,(e_t)) : \pi \in \mathcal{P}\}$ junto con el índice de funcionamiento a ser optimizado se llama un problema de control de Markov.

Ecuación de optimalidad con costo descontado.

Consideremos el problema de control con índice de funcionamiento el costo descontado (Definiciones 1.5 y 1.6 a).

Definición 1.8 *Ecuación de optimalidad con costo descontado.*

Una función medible v:$\mathbb{E} \to \mathbb{R}$ se dice que es solución de la ecuación de optimalidad con costo descontado si, para todo $e \in \mathbb{E}$,

$$v(e) = \min_{a \in A(e)} \left\{ c(e,a) + \alpha \int_{\mathbb{E}} v(x) K(dx/e,a) \right\} \tag{1.3}$$

donde α es el factor de descuento.

Necesitaremos la siguiente definición, que es similar a la propiedad de Feller "fuerte". (En la definición 2.17 se introduce una versión "débil".)

Definición 1.9 *Kérnel estocástico fuertemente continuo.*

Sean $\mathbb{X}$. $\mathbb{Y}$ espacios topológicos y K un kérnel estocástico en $\mathbb{X}$ dado $\mathbb{Y}$. Decimos que el kérnel estocástico K es fuertemente continuo si la función

$$y \to \int_{\mathbb{X}} h(x) K(dx/y)$$

es continua y acotada en $\mathbb{Y}$ para cualquier función medible y acotada h en $\mathbb{X}$.

Se pueden dar condiciones bastante generales para que la ecuación de optimalidad con costo descontado (Definición 1.8) tenga como solución mínima a la función de valor (Definición 1.7) y, además, para que la ecuación nos proporcione una política determinista óptima, como en el siguiente teorema ([26], págs. 46.47).

Teorema 1.10 .

Supongamos que:

1) La función de costo en una etapa, c, es semicontinua inferiormente.

2) la ley de transición K es fuertemente continua, y

3) existe una política π tal que $I_d(\pi, e) < \infty$ para cada $e \in \mathbb{E}$.

Entonces:

a) La función de valor $I_d(e) := \inf_\pi I_d(\pi, e)$ es la mínima solución de la ecuación de optimalidad con costo descontado, i. e.,

$$I_d(e) = \min_{a \in A(e)} \left\{ c(e,a) + \alpha \int_{\mathbb{E}} I_d(x) K(dx/e,a) \right\} \tag{1.4}$$

para todo $e \in \mathbb{E}$, y si u es otra solución de la ecuación (1.3) entonces $u \geq I_d$.

b) Existe un selector medible $f \in \mathfrak{F}$ tal que $f(e)$ alcanza el mínimo en (1.4), i. e.,

$$I_d(e) = c(e, f) + \alpha \int_{\mathbb{E}} I_d(x) K(dx/e, f)$$

y la política determinista f^∞ es e-óptima para todo estado inicial $e \in \mathbb{E}$. Recíprocamente si una política estacionaria determinista es e-óptima para todo estado inicial $e \in \mathbb{E}$, entonces el correspondiente índice de funcionamiento satisface la ecuación de optimalidad (1.3).

En el problema de control con restricciones (Definición 2.26) podemos tener una situación bastante diferente ya que puede darse el caso de que la función de valor no sea solución de la ecuación de optimalidad y/o que la política óptima no sea estacionaria, por ejemplo, puede depender del estado inicial o que el óptimo no se alcance dentro de las políticas deterministas (Ejemplo 2.28).

Ternas canónicas.

Recordemos el problema de control de Markov con índice de costo promedio (Definiciones 1.6b y 1.7). Consideraremos únicamente para el caso en que el proceso de control de Markov tiene una sola clase ergódica, el así llamado caso unicadena.

Sea $h : \mathbb{E} \to \mathbb{R}$ una función medible dada y sea $I_N(\pi, e, h)$ el costo esperado total en N etapas cuando se usa la política π, el estado inicial es e y el costo terminal es h, a saber:

$$\begin{aligned} I_0(\pi, e, h) &: = h(e) \quad \text{y} \\ I_N(\pi, e, h) &: = E_e^\pi \left[\sum_{t=0}^{N-1} c(e_t, a_t) + h(e_N) \right] \\ &: = I_N(\pi, e) + E_e^\pi (h(e_N)), \qquad N \geqslant 1. \end{aligned}$$

La función de valor correspondiente la definimos como

$$I_N(e, h) := \inf_{\pi \in \mathcal{P}} I_N(\pi, e, h).$$

Definición 1.11 *Terna canónica.*

Sean $h : \mathbb{E} \to \mathbb{R}$ una función medible, ρ un número real y f un selector medible dado. Entonces se dice que (ρ, h, f) es una terna canónica si

$$I_N(f^\infty, e, h) = I_N(e, h) = N\rho + h(e)$$

para todo $e \in \mathbb{E}$ y $N = 0, 1, \ldots$. Un selector medible f se dice que es canónico si entra en alguna terna canónica.

Teorema 1.12

(ρ, h, f) es una terna canónica si y sólo si, para todo estado $e \in \mathbb{E}$,

a) (ρ, h) satisface

$$\rho + h(e) = \inf_{a \in A(e)} \left\{ c(e, a) + \int_{\mathbb{E}} h(x) K(dx/e, a) \right\};$$

b) $f \in \mathfrak{F}$ alcanza el mínimo en (a), es decir, (tomando en cuenta la Observación 1.3)

$$\rho + h(e) = c(e, f) + \int_{\mathbb{E}} h(x) K(dx/e, f).$$

Note que si (ρ, h, f) es una terna canónica, entonces también lo es $(\rho, h + k, f)$ para cualquier constante k ([26], pág. 78).

Hipótesis 1.13

$$\lim_{N \to \infty} \frac{E_e^\pi h(e_N)}{N} = 0$$

para toda $\pi \in \mathcal{P}$ y $e \in \mathbb{E}$.

Bajo la Hipótesis 1.13 la política canónica f^∞ en el Teorema 1.12 es óptima y

$$\rho = I_p(f^\infty, e) = I_p(e)$$

para todo estado inicial $e \in \mathbb{E}$.

Teorema 1.14

Sea (ρ^*, h, f) una terna canónica y supongamos la Hipótesis 1.13. Entonces f^∞ es e-óptima para todo $e \in \mathbb{E}$ y ρ^* es la función de valor; de hecho,

$$I_p(e) = \rho^* = I_p(f^\infty, e) = \lim_{N \to \infty} \frac{I_N(f^\infty, e)}{N} \quad \forall e \in \mathbb{E}.$$

([26], págs. 80-81).

COMENTARIO FINAL AL CAPÍTULO 1

El problema de control (Definición 1.7) consiste en encontrar una política π que sea óptima con respecto a algún criterio de funcionamiento y el enfoque más usual para resolverlo es por medio de la programación dinámica. En los siguientes dos capítulos veremos dos enfoques alternativos de solución a dicho problema de control, en el capítulo 2, vía las medidas de ocupación para el costo descontado, y, en el capítulo 3, vía la programación lineal para el costo promedio en el caso unicadena. Además, en la sección 2.6 ilustraremos con un ejemplo que en el caso restringido para el costo descontado, no siempre funciona la ecuación de optimalidad (1.3) para obtener políticas óptimas.

Capítulo 2

MEDIDAS DE OCUPACIÓN Y PROBLEMAS DESCONTADOS

2.1 INTRODUCCIÓN

Un hecho bien conocido en la teoría del control es que muchos problemas de control, deterministas o estocásticos, pueden ser transformados en problemas de minimización sobre conjuntos de medidas, véanse por ejemplo [4], [8], [21], [34], [39] y [40]. En este caso el problema de control se reduce típicamente a la forma:

$$\text{minimizar } \int_{\mathbf{X}} c(x)\,\mu(dx) \quad \text{sujeto a } \mu \in \Gamma \tag{2.1}$$

donde Γ es una cierta familia de medidas definidas en $\mathfrak{B}(\mathbb{X})$ y c denota al costo corriente del problema de control. Más aún, bajo hipótesis relativamente simples sobre c y $\mathbb{X}$, y considerando a la familia de medidas con la topología débil, resulta que la función $\mu \mapsto \int_{\mathbf{X}} c(x)\,\mu(dx)$ es semicontinua inferiormente y, por lo tanto, una forma directa de demostrar que existe una solución óptima para (2.1) es mostrar que Γ es secuencialmente débilmente compacto. A su vez, por el Teorema de Prohorov (Teorema 2.8), para probar esto es suficiente mostrar que Γ es tensa, que es una de las principales cuestiones con las que estamos interesados en este capítulo.

El contenido del capítulo es el siguiente: En la sección 2.2 recordamos algunos conceptos y teoremas relacionados con el concepto de tensión (como el Teorema 2.3 y el Teorema de Prohorov), y construimos un ejemplo de una cadena de Markov que es acotada en probabilidad; pero que no es tensa (Ejemplo 2.6). En la sección 2.3 construimos las envolventes superior e inferior de una familia arbitraria de medidas (Teoremas 2.11 y 2.12), damos algunos casos de interés (Observación 2.13 y Ejemplo 2.15) y relacionamos tales envolventes con el concepto de tensión (Teorema 2.14). En la sección 2.4 aplicamos estos conceptos a los procesos de control de Markov (Lema 2.21 y Teorema 2.22) para dar condiciones suficientes para la existencia de una política óptima determinista cuando el índice de funcionamiento es el costo descontado. En la sección 2.5 damos un contraejemplo a una afirmación de Kurano y Kawai, el Teorema 3.3 de [32]. En la sección 2.6 presentamos los conceptos básicos de los procesos de control de Markov con restricciones y con índice de funcionamiento el costo descontado; en el (Ejemplo 2.28) mostramos las diferencias que pueden existir con el caso no restringido, como pueden ser, que el control óptimo no necesariamente se alcance dentro de la clase de las políticas deterministas que satisfacen la ecuación de optimalidad (Definición 1.8). Por último, en la sección 2.7 damos condiciones (Hipótesis 2.29), al problema de control de Markov con restricciones (Definición 2.27), para que el valor óptimo del costo descontado se alcance dentro de la clase de las políticas estacionarias aleatorizadas (Teorema 2.31).

Terminología y Notación.

La siguiente notación será usada en el capítulo. Dado un espacio topológico $\mathbb{X}$, denotamos por $2^{\mathbb{X}}$ al conjunto potencia de $\mathbb{X}$, por A^c al complemento del conjunto $A \in 2^{\mathbb{X}}$ y por $\overline{A}$ a la cerradura del conjunto. Además denotamos por $\mathbb{R}_+ := \{r \in \mathbb{R} : r \geq 0\}$ a los números reales no negativos y por $\overline{\mathbb{R}}_+ := \mathbb{R}_+ \cup \{+\infty\}$ a los números reales extendidos no negativos. Así mismo, denotamos por $C_a(\mathbb{X})$ al conjunto de funciones reales continuas y acotadas en $\mathbb{X}$.

2.2 TENSIÓN

En esta sección y en la siguiente $\mathbb{X}$ denota un espacio topológico.

Definición 2.1 *Medida tensa, familia de medidas tensa.*

a) Una *medida* finita γ en $\mathbb{X}$ se dice que es *tensa* si para cada $\varepsilon > 0$ existe un conjunto compacto C tal que $\gamma(C^c) < \varepsilon$.

b) Una *familia de medidas* finitas Γ en $\mathfrak{B}(\mathbb{X})$ se dice que es *tensa* si para cada $\varepsilon > 0$ existe un conjunto compacto C tal que $\gamma(C^c) < \varepsilon$ para toda $\gamma \in \Gamma$.

Por ejemplo, si $\mathbb{X}$ es un espacio $\sigma-$compacto entonces cada medida finita en $\mathbb{X}$ es *tensa* ([9] *teorema* 1.4). La propiedad de tensión para una familia Γ está muy relacionada con la existencia de una función inf-compacta (Definición 2.2).

Definición 2.2 *Función ínf-compacta.*

Una función medible f en un espacio topológico $\mathbb{X}$ y con valores en los números reales extendidos $\overline{\mathbb{R}}$ se dice que es inf-compacta si el conjunto de nivel es $\{x \in \mathbb{X}: f(x) \leqslant r\}$ es un conjunto compacto para todo número real $r \in \mathbb{R}$.

Teorema 2.3

Sea Γ una familia de medidas en $\mathfrak{B}(\mathbb{X})$. Si Γ es acotada[1], entonces Γ es tensa si y sólo si existe una constante k y una función ínf.-compacta f tal que $f \geq 1$ y

$$\int_{\mathbf{X}} f(x)\,\gamma(dx) \leqslant k \qquad \forall\gamma \in \Gamma.$$

Si Γ es una familia de medidas de probabilidad en $\mathfrak{B}(\mathbb{X})$, la condición $f \geq 1$ puede ser reemplazada por $f \geq 0$ (por ejemplo, en [6] o en [11] pág. 109).

Como una consecuencia del Teorema 2.3, se puede ver que si el espacio $\mathbb{X}$ es compacto, entonces cualquier familia acotada de medidas en $\mathfrak{B}(\mathbb{X})$ es tensa. (Tómese $f(\cdot) \equiv$constante.)

[1] Existe una constante k tal que $\gamma(\mathbf{X}) \leqslant k$ para toda $\gamma \in \Gamma$.

Definición 2.4 *Cadena de Markov tensa, cadena de Markov acotada en probabilidad.*

a) Sean M una cadena de Markov en $\mathbb{X}$ y K^n la probabilidad de transición en n pasos asociada a M, $n = 0, 1, 2, ...$ Decimos que la *cadena de Markov* M es *tensa* si para cada $x \in \mathbb{X}$ y cada $\varepsilon > 0$ existe un conjunto compacto C tal que $K^n(C^c/x) < \varepsilon$ para toda $n \in \mathbb{N}$. Es decir, M es tensa, si para cada $x \in \mathbb{X}$ la familia $\Gamma_x := \{K^n(\cdot/x) : n = 1, 2, ...\}$ es tensa.

b) Y decimos que la *cadena de Markov* es *acotada en probabilidad* si para cada $x \in \mathbb{X}$ y cada $\varepsilon > 0$ existe un conjunto compacto C tal que $\limsup_{n\to\infty} K^n(C^c/x) < \varepsilon$.

Observe que la condición $K^n(C^c/x) < \varepsilon$ en la Definición 2.4 a) es equivalente a $K^n(C/x) \geq 1 - \varepsilon$ y que la condición $\limsup_{n\to\infty} K^n(C^c/x) < \varepsilon$ en la Definición 2.4 b) es equivalente a $\liminf_{n\to\infty} K(C/x) \geq 1 - \varepsilon$.

En general la condición de que una cadena de Markov sea tensa implica la condición de que sea acotada en probabilidad. Y es fácil dar condiciones suficientes para que se cumpla la equivalencia entre ambas definiciones, como en la siguiente Proposición.

Proposición 2.5

Sea M una cadena de Markov en $\mathbb{X}$ tal que para cada $x \in \mathbb{X}$ y cada $n \in \mathbb{N}$ la probabilidad de transición en n pasos $K^n(\cdot/x)$ es tensa. Entonces M es tensa si y sólo si M es acotada en probabilidad.

Demostración:

Hicimos notar que si M es una cadena tensa, entonces es acotada en probabilidad. Sólo nos resta probar la implicación recíproca.

($\Longleftarrow$) Supóngase que la cadena M es acotada en probabilidad, y sean $x \in \mathbb{X}$ y $\epsilon > 0$. Entonces existen un conjunto compacto $C = C(x, \epsilon)$ y un número entero $N = N(x, \epsilon)$ tales que $K^n(C/x) \geq 1 - \epsilon$ para toda $n \geq N$. Por otro lado, como cada probabilidad de transición es tensa, existen conjuntos compactos C_j tales que $K^j(C_j/x) \geq 1 - \epsilon$ para $j = 1, 2, ..., N - 1$. Así definiendo $C_* := C_1 \cup ... \cup C_{N-1} \cup C$ obtenemos que este conjunto C_* satisface $K^n(C_*/x) \geq 1 - \epsilon$ para toda $n = 1, 2, ...$ Como x y ϵ son arbitrarias, M es tensa.▲

Una condición suficiente para que se cumpla la hipótesis de la Proposición 2.5 es que en el espacio $\mathbb{X}$ cualquier medida de probabilidad sea tensa. Así observamos que si $\mathbb{X}$ satisface cualquiera de las siguientes condiciones:

- $\mathbb{X}$ es un espacio polaco,
- $\mathbb{X}$ es un espacio métrico separable localmente compacto,
- $\mathbb{X}$ es σ-compacto,

entonces, ambas definiciones son equivalentes. Sin embargo, en general ambas definiciones no son equivalentes como lo muestra el siguiente ejemplo.

Ejemplo 2.6 *Una cadena de Markov acotada en probabilidad y no tensa.*

a) Sea $\mathbb{X}_1$ un conjunto no numerable y sea

$$\tau_1 = \{A \subseteq \mathbb{X}_1 : A^c \text{ es numerable}\} \cup \{\emptyset\}$$

la topología de todos los conjuntos de complemento numerable. Definamos la medida de probabilidad p_1 en $\mathfrak{B}(\mathbb{X}_1)$ como

$$p_1(A) = \begin{cases} 1 & \text{si } A^c \text{ es un conjunto numerable} \\ 0 & \text{si } A \text{ es un conjunto numerable} \end{cases}$$

donde

$$\mathfrak{B}(\mathbb{X}_1) = \{A \subseteq \mathbb{X}_1 : A \in \tau_1 \text{ o } A^c \in \tau_1\}$$

es la σ-álgebra de Borel

Entonces:

1. El espacio topológico $(\mathbb{X}_1, \tau_1)$ no es localmente compacto y si A es un subconjunto infinito de $\mathbb{X}_1$ entonces A no es compacto, y

2. p_1 es una medida de probabilidad y p_1 no es tensa.

Demostración:

1) Sea $A \subseteq \mathbb{X}_1$ un conjunto infinito, si A es numerable tómese $B = A$; en otro caso tómese $B = \{b_1, b_2, b_3 ...\} \subseteq A$. Definamos $A^n := B^c \cup \{b_1, b_2, ..., b_n\}$. Entonces $\{A_n\}$ es una cubierta abierta de A sin subcubiertas finitas. Por lo tanto, A no es un conjunto compacto. En particular, $\mathbb{X}_1$ no es un conjunto

compacto. Además, si C es un conjunto no vacío en τ_1, entonces su cerradura es el espacio total, es decir, $\overline{C} = \mathbb{X}_1$. Por lo tanto $(\mathbb{X}, \tau_1)$ no es localmente compacto.

2) La prueba de la σ-aditividad de p_1 se hace analizando por casos, primero cuando todos los subconjuntos son conjuntos numerables, y despues cuando existe solamente uno que sea de complemento numerable, ya que no pueden existir dos conjuntos ajenos de complemento numerable en $\mathfrak{B}(\mathbb{X}_1)$. Finalmente, p_1 no es tensa porque, por 1) y la definición de p_1, los conjuntos compactos tienen probabilidad cero.▲

b) Sean a y b dos elementos cualesquiera que no pertenecen a $\mathbb{X}_1$ y definamos

$$\mathbb{X} := \mathbb{X}_1 \cup \{a, b\}$$

con la topología

$$\tau := \{A, A \cup \{a\}, A \cup \{b\}, A \cup \{a, b\} : A \in \mathfrak{B}(\mathbb{X}_1)\}.$$

Consideremos una cadena de Markov M sobre $\mathbb{X}$ con probabilidad de transición K de la siguiente manera:

$$\begin{aligned} K[A/a] &:= p_1(A) \qquad \text{para } A \in \mathfrak{B}(\mathbb{X}_1) \\ K[x/a] &:= 0 \qquad \text{para } x \in \{a, b\} \\ K[b/x] &:= 1 \qquad \text{para } x \in \mathbb{X}_1 \cup \{b\} \\ K[\cdot/\cdot] &:= 0 \qquad \text{en otro caso.} \end{aligned}$$

entonces la cadena de Markov M definida por K es acotada en probabilidad y no es tensa.

Demostración:

Que M no es tensa se sigue del inciso a), mientras que la propiedad de ser acotada en probabilidad se sigue del hecho de que $K^n[b/a] = p_1(\mathbb{X}_1) = 1$ para todo $n \geq 2$.▲

Terminaremos esta sección recordando el Teorema de Prohorov 2.8, que será usado en el Lema 2.21 y en el Teorema 2.22, para lo cual necesitamos la siguiente definición:

Definición 2.7 *Convergencia débil*

Sea $\mathbb{X}$ un espacio topológico y sean γ_n, $n = 1, 2, ...,$ y γ medidas finitas definidas en $\mathfrak{B}(\mathbb{X})$. Decimos que γ_n converge débilmente a γ si

$$\int_{\mathbb{X}} f(x) \gamma_n(dx) \longrightarrow \int_{\mathbb{X}} f(x) \gamma(dx)$$

para toda función f en $C_a(\mathbb{X})$.

Teorema 2.8 *Teorema de Prohorov*

Sea $\mathbb{X}$ un espacio polaco y sea Γ una familia de medidas de probabilidad definidas en $\mathfrak{B}(\mathbb{X})$. Entonces Γ es tensa si y sólo si Γ es secuencialmente débilmente precompacta (por ejemplo, [9], [11], [16] y [31]).

Observación 2.9

Es fácil ver que el teorema de Prohorov se extiende al caso en que Γ es una familia acotada de medidas ([16], [40]).

2.3 ENVOLVENTES DE MEDIDAS

Definición 2.10 *Medida exterior, premedida.*

Sea $\mathbb{X}$ un espacio topológico, $\mathcal{A} \subseteq 2^{\mathbb{X}}$ una familia de subconjuntos de $\mathbb{X}$, y $\lambda : \mathcal{A} \to \overline{\mathbb{R}}_+$ una función definida en la familia de conjuntos $\mathcal{A}$ a los reales extendidos no negativos. Se dice que λ es:

a) *monótona* si para cualesquiera A, B en $\mathcal{A}$ tales que $A \subseteq B$ se cumple que $\lambda(A) \leqslant \lambda(B)$,

b) *subaditiva* si para cualquier sucesión $\{A_n\}$ en $\mathcal{A}$ de conjuntos ajenos y cuya unión es $A \in \mathcal{A}$ se cumple que

$$\lambda(A) \leqslant \sum_{n=1}^{\infty} \lambda(A_n),$$

c) *superaditiva* si para cualquier sucesión $\{A_n\}$ como en (b)

$$\lambda(A) \geqslant \sum_{n=1}^{\infty} \lambda(A_n),$$

d) una *medida exterior* si, con $\mathcal{A} = 2^{\mathbf{X}}$, $\lambda(\emptyset) = 0$ y λ es monótona y subaditiva, y

e) una *premedida* en $\mathcal{A} = \mathfrak{B}(\mathbb{X})$ si, $\lambda(\emptyset) = 0$ y λ es monótona y superaditiva.

Una sucesión $\{A_n\}$ como en el inciso b) será llamada una partición de A. Por otra parte, diremos que $\{A_n\}$ es una cubierta de un conjunto B si $\cup_n A_n \supset B$.

Teorema 2.11 *(Envolvente superior de una familia de medidas).*

Sea Γ una familia de medidas definidas en $\mathfrak{B}(\mathbb{X})$ y sean σ y γ^s las funciones de conjuntos definidas en $\mathfrak{B}(\mathbb{X})$ por

$$\sigma(A) := \sup\{\gamma(A) : \gamma \in \Gamma\} \tag{2.2}$$

y

$$\gamma^s(A) := \sup\left\{\sum_{n=1}^{\infty} \sigma(A_n) : \{A_n\} \text{ es una partición de } A \text{ en } \mathfrak{B}(\mathbb{X})\right\}, \tag{2.3}$$

respectivamente. Entonces

a) $\sigma(\emptyset) = 0$ y σ es monótona y subaditiva

b) γ^s es una medida, y más aún,

c) i. $\gamma^s \geqslant \sigma$, y

ii. si μ es una medida en $\mathbb{X}$ tal que $\mu \geqslant \sigma \Rightarrow \mu \geqslant \gamma^s$.

En virtud de los incisos b) y c) del Teorema 2.11 nos referiremos a γ^s como la envolvente superior de Γ.

Demostración:

a) Las propiedades $\sigma(\emptyset) = 0$ y que σ es monótona son obviamente ciertas. Para probar la subaditividad de σ, sean $A \in \mathfrak{B}(\mathbb{X})$ y $\{A_n\}$ una partición de A en $\mathfrak{B}(\mathbb{X})$. Entonces dado que

$$\gamma(A) = \sum_n \gamma(A_n) \qquad \forall \gamma \in \Gamma,$$

las propiedades del supremo implican que

$$\sigma(A) = \sup_{\gamma} \sum_{n} \gamma(A_n) \leqslant \sum_{n} \sup_{\gamma} \gamma(A_n) = \sum_{n} \sigma(A_n).$$

b) Por a) la función γ^s satisface $\gamma^s(\emptyset) = 0$ y es monótona, de aquí que solamente resta mostrar que γ^s es σ-aditiva. Sean $A \in \mathfrak{B}(\mathbb{X})$ y $\{A_n\}$ una partición de A en $\mathfrak{B}(\mathbb{X})$. Consideraremos dos casos: $\gamma^s(A)$ finito y $\gamma^s(A)$ infinito.

Caso 1: $\gamma^s(A) < \infty$.

En este caso la monotonicidad de γ^s implica que $\gamma^s(A_n) < \infty$ para toda n.

Sea $\varepsilon > 0$. Entonces, por la definición de de γ^s dada en (2.3), para cada n existe una partición $\{B_{n_k} : k = 1, 2, ...\}$ de A_n tal que

$$\gamma^s(A_n) \leqslant \sum_{k} \sigma(B_{n_k}) + \frac{\varepsilon}{2^n}. \tag{2.4}$$

Luego como $\{B_{n_k} : k = 1, 2, ... \text{ y } n = 1, 2, 3, ...\}$ es una partición de A_n en $\mathfrak{B}(\mathbb{X})$, de (2.3) y (2.4) obtenemos

$$\sum_{n} \gamma^s(A_n) \leqslant \sum_{n} \sum_{k} \sigma(B_{n_k}) + \varepsilon \leqslant \gamma^s(A) + \varepsilon.$$

Así, dado que $\varepsilon > 0$ es arbitrario, obtenemos

$$\sum_{n} \gamma^s(A_n) \leqslant \gamma^s(A). \tag{2.5}$$

Por otro lado, fijando $\varepsilon > 0$ otra vez y usando (2.3) existe una partición $\{C_k\}$ de A en $\mathbb{X}$ tal que

$$\gamma^s(A) \leqslant \sum_{k} \sigma(C_k) + \varepsilon. \tag{2.6}$$

Por lo tanto, definiendo $C_{n_k} := C_k \cap A_n$ obtenemos que $\{C_{n_k} : k = 1, 2, 3, ...\}$ es una partición de A_n en $\mathfrak{B}(\mathbb{X})$ y, además, se cumple que $C_k = \cup_n C_{n_k}$ para toda k. Así, de a) y (2.6),

$$\gamma^s(A) \leqslant \sum_{k} \sum_{n} \sigma(C_{n_k}) + \varepsilon = \sum_{n} \sum_{k} \sigma(C_{n_k}) + \varepsilon \leqslant \sum_{n} \gamma^s(A_n) + \varepsilon.$$

Luego, como $\varepsilon > 0$ es arbitario,

$$\gamma^s(A) \leqslant \sum_n \gamma^s(A_n)$$

y junto con (2.5) obtenemos la igualdad.

Caso 2: $\gamma^s(A) = \infty$.

Sea $M > 0$. Entonces, por (2.3), existe una partición $\{C_k\}$ de A en $\mathbb{X}$ tal que

$$\sum_k \sigma(C_k) \geq M.$$

Para cada n y k definamos $C_{kn} := C_k \cap A_n$. Entonces $\{C_{kn} : n = 1, 2, 3, ...\}$ es una partición de C_k y $\{C_{kn} : k = 1, 2, 3, ...\}$ es una partición de A_n. De aquí que

$$\gamma^s(A_n) \geq \sum_k \sigma(C_{kn})$$

y, por lo tanto

$$\begin{aligned}\sum_n \gamma^s(A_n) &\geq \sum_n \sum_k \sigma(C_{kn}) = \sum_k \sum_n \sigma(C_{kn}) \\ &\geq \sum_k \sigma(C_{kn}) \quad \text{ya que } \sigma \text{ es subaditiva por el inciso a)} \\ &\geq M.\end{aligned}$$

Como M es arbitrario, se sigue que $\gamma^s(A) = \sum_n \gamma^s(A_n)$.

c) i. Dado que $\{A, \emptyset, \emptyset, \emptyset, ...\}$ es una partición de A en $\mathfrak{B}(\mathbb{X})$, por definición (2.3) de γ^s, obtenemos $\gamma^s(A) \geqslant \sigma(A)$

c) ii. Sea μ una medida definida en $\mathfrak{B}(\mathbb{X})$ tal que $\mu \geqslant \sigma$, y sea A en $\mathfrak{B}(\mathbb{X})$. Supongamos primero que $\gamma^s(A) < \infty$. Entonces, dado $\varepsilon > 0$, existe una partición $\{A_n\}$ de A tal que

$$\gamma^s(A) \leqslant \sum_n \sigma(A_n) + \varepsilon$$

Así que, como $\mu \geqslant \sigma$ y μ es una medida,

$$\gamma^s(A) \leqslant \sum_n \mu(A_n) + \varepsilon = \mu(A) + \varepsilon.$$

Luego, como $\varepsilon > 0$ es arbitrario obtenemos $\gamma^s(A) \leqslant \mu(A)$.

Ahora supongamos que $\gamma^s(A) = \infty$. Entonces dado $M > 0$ existe una partición $\{A_n\}$ de A en $\mathfrak{B}(\mathbb{X})$ tal que

$$\sum_n \sigma(A_n) \geqslant M.$$

Por lo tanto,

$$\mu(A) = \sum_n \mu(A_n) \geqslant \sum_n \sigma(A_n) \geqslant M$$

y como $M > 0$ es arbitraria, obtenemos $\mu(A) = \infty = \gamma^s(A)$ ▲.

Ahora construiremos un envolvente inferior para una familia de medidas Γ. Dicha construcción está relacionada con la construcción de una medida a partir de una premedida [35] (Definición 2.10 e).

Teorema 2.12 *(Envolvente inferior de una familia de medidas.)*

Sea Γ una familia de medidas definidas en $\mathfrak{B}(\mathbb{X})$ y sean ι y η las funciones definidas en $\mathfrak{B}(\mathbb{X})$ y en $2^{\mathbb{X}}$, respectivamente, por

$$\iota(A) := \inf\{\gamma(A) : \gamma \in \Gamma\} \qquad \text{para } A \in \mathfrak{B}(\mathbb{X}), \tag{2.7}$$

y

$$\eta(A) := \inf\left\{\sum_{n=1}^{\infty} \iota(A_n) : \{A_n\} \text{ es una cubierta de } A \text{ en } \mathfrak{B}(\mathbb{X})\right\} \tag{2.8}$$

para $A \in 2^{\mathbb{X}}$. Además, sea γ^i la restricción de η a la $\sigma-$álgebra $\mathfrak{B}(\mathbb{X})$, i.e.,

$$\gamma^i := \eta_{|\mathfrak{B}(\mathbb{X})}. \tag{2.9}$$

Entonces

a) ι es una premedida en $\mathfrak{B}(\mathbb{X})$,

b) η es una medida exterior,

c) γ^i es una medida en $\mathfrak{B}(\mathbb{X})$ y más aún

d) **i.** $\gamma^i \leqslant \iota$, y

ii. si μ es una medida en $\mathfrak{B}(\mathbb{X})$ tal que $\mu \leqslant \iota$, entonces $\mu \leqslant \gamma^i$.

Nos referimos a γ^i como la envolvente inferior de Γ.

Omitiremos la demostración ya que es similar a la del Teorema 2.3. De hecho, la prueba de las partes b) y c) del Teorema 2.12 son similares a la prueba del Teorema de Extensión de Carathéodory ([3], pág 20). Ahora mencionaremos brevemente algunos casos especiales de interés de las envolventes superior e inferior.

Observación 2.13

a) Si $\mathbb{Y}$ es un subconjunto de Borel de $\mathbb{X}$, entonces la conclusión del Teorema 2.11 inciso c) ii. se satisface en $\mathbb{Y}$. Esto es. supóngase que $\mathbb{Y} \in \mathfrak{B}(\mathbb{X})$ y sea μ una medida en $\mathbb{X}$ tal que $\mu(A) \geqslant \sigma(A)$ para todo subconjunto de Borel de $\mathbb{Y}$. Entonces $\mu \geqslant \gamma^s$ en $\mathfrak{B}(\mathbb{Y})$.

Para ver esto, defínase una nueva medida μ' en $\mathfrak{B}(\mathbb{X})$ como

$$\mu'(A) := \begin{cases} \mu(A) & \text{si } A \in \mathfrak{B}(\mathbb{Y}) \\ \infty & \text{en otro caso} \end{cases} \tag{2.10}$$

Así. μ' es una medida en $\mathbb{X}$ y $\mu' \geqslant \sigma$. De aquí que. por el Teorema 2.11 inciso c). $\mu' \geqslant \gamma^s$ en $\mathbb{X}$. En particular, si $A \in \mathfrak{B}(\mathbb{Y})$. entonces.

$$\mu(A) = \mu'(A) \geqslant \gamma^s(A),$$

i.e..

$$\mu' \geqslant \gamma^s \text{ en } \mathfrak{B}(\mathbb{Y}).$$

Una observación similar se satisface para el inciso d) ii) del Teorema 2.12: a saber. si $\mathbb{Y} \in \mathfrak{B}(\mathbb{X})$ y μ es una medida en $\mathbb{X}$ de forma tal que $\mu \leqslant \iota$ en $\mathbb{Y}$. entonces $\mu \leqslant \gamma^i$ en $\mathfrak{B}(\mathbb{Y})$.

Para ver esto. en lugar de μ' en (2.10) tómese $\mu'(A) := \mu'(A \cap \mathbb{Y})$. para $A \in \mathfrak{B}(\mathbb{X})$.

b) Si $\Gamma = \{\gamma_k\}$ es un conjunto numerable de medidas en $\mathfrak{B}(\mathbb{X})$. las envolventes γ^i y γ^s se pueden construir en forma más fácil. Por ejemplo. para cada $n = 1, 2, 3, \ldots$ definamos

$$\mu_n := \max\{\gamma_1, \gamma_2, \ldots, \gamma_n\}.$$

esto es,

$$\mu_1 := \gamma_1,$$

y para $n \geqslant 2$ (por ejemplo,[36], pág. 237), μ_n se puede obtener recursivamente como

$$\mu_n := \max\left\{\mu_{n-1}, \gamma_n\right\} = \frac{1}{2}\left\{\mu_{n-1} + \gamma_n + \left|\mu_{n-1} - \gamma_n\right|\right\} \tag{2.11}$$

Entonces las medidas μ_n forman una sucesión creciente que converge a una medida ([16], página 31) que se puede verificar fácilmente que es γ^s. Alternativamente, en lugar de definir σ como en (2.2), la función de conjuntos

$$\sigma := \limsup_{k\to\infty} \gamma_k = \lim_n \sup_{k\geqslant n} \gamma_k \tag{2.12}$$

satisface el inciso a) del Teorema 2.11 y, entonces, suponiendo que la sucesión Γ es acotada [2], la envolvente superior γ^s se obtiene como en la ecuación (2.3). Este resultado requiere que Γ sea acotada porque la sucesión $\sup_{k\geqslant n}\gamma_k$ en (2.12) es decreciente, en contraste con μ_n en (2.11) que forma una sucesión creciente.

Similarmente, si la sucesión de medidas Γ es acotada, entonces la sucesión $\eta_n := \min\left\{\gamma_1, \gamma_2, \ldots, \gamma_n\right\}$ decrece a γ^i. Y alternativamente a definir ι como en (2.7), podemos definir, aún si Γ no es acotada,

$$\iota := \liminf_{k\to\infty} \gamma_k = \lim_n \inf_{k\geqslant n} \gamma_k.$$

Entonces ι es una premedida y obtenemos a γ^i como en (2.8) y (2.9).

c) Sea $\Gamma = \{\gamma_k\}$ una sucesión de medidas en $\mathfrak{B}(\mathbb{X})$, y supóngase que existe una medida en $\mathfrak{B}(\mathbb{X})$, σ-finita, tal que, para todo k, γ_k es absolutamente continua con respecto a λ; en símbolos: $\gamma_k << \lambda$ para todo k. Sea f_k la densidad (o derivada de Radon-Nikodym) de γ_k respecto a λ. Como en el inciso b), sea

$$\mu_n := \max\left\{\gamma_1, \gamma_2, \ldots \gamma_n\right\}$$

y definamos

$$\begin{aligned} g_1 &:= f_1 \\ g_n &:= \max\{f_1, f_2, \ldots, f_n\} = \max\{g_{n-1}, f_n\} \quad \text{para } n = 2, 3, 4, \ldots \end{aligned}$$

[2]Existe una constante c tal que $\gamma_k(\mathbb{X}) \leqslant c$ para toda k.

Entonces, como en b), $\lim_n \mu_n = \gamma^s$ y, además

i. g_n es la densidad de μ_n con respecto a λ, y

ii. la función $f^s := \lim_n g_n = \sup_k f_k$ es la densidad de γ^s con respecto a λ.

En realidad, i) se sigue de (2.11), mientras que ii) se sigue de i) y del Teorema de Convergencia Monótona. La parte ii) puede ser interpretada como una variante del teorema de Vitali-Hahn-Sacks ([16] y [27]). En forma resumida, c) establece que $\gamma_k << \lambda$ para todo k implica que $\gamma^s << \lambda$, y, por supuesto, la implicación recíproca es trivialmente cierta por el inciso c) del Teorema 2.11. De hecho, esto se satisface para una familia general Γ, no necesariamente numerable. En efecto:

d) Sea Γ una familia de medidas en $\mathbb{X}$, sea γ^s la envolvente superior de esta familia, y sea λ una medida en $\mathbb{X}$. Entonces $\gamma^s << \lambda$ si y sólo si $\gamma << \lambda$ para toda $\gamma \in \Gamma$.

La implicación $\Rightarrow$ es obvia por el inciso c) del Teorema 2.11. Supongamos entonces que $\gamma << \lambda$ para toda $\gamma \in \Gamma$ y definamos la medida λ' en $\mathbb{X}$ como

$$\lambda'(A) = \begin{cases} 0 & \text{si } \lambda(A) = 0 \\ \infty & \text{en otro caso} \end{cases}$$

entonces, con σ como en (2.2), $\lambda' \geqslant \sigma$. Esto implica, por el inciso c) del Teorema 2.11, que $\lambda' \geqslant \gamma^s$. Así

$$\lambda(A) = 0 \Rightarrow \lambda'(A) = 0 \Rightarrow \gamma^s(A) = 0,$$

es decir, $\gamma^s << \lambda$.

Teorema 2.14

Sea Γ una familia de medidas en $\mathbb{X}$, y sean γ^s y γ^i las envolventes superior e inferior de Γ respectivamente. Entonces

a) γ^s tensa $\Rightarrow \Gamma$ tensa $\Rightarrow \gamma^i$ tensa

b) Si Γ es tensa y γ^s es una medida finita, entonces γ^s es tensa.

La conclusión del Teorema 2.14 b) puede fallar si γ^s no es finita. Por ejemplo, supóngase que $\mathbb{X} = \mathbb{R}$, y sea $\Gamma = \{\gamma_k\}$ la familia numerable de medidas de probabilidad definidas como

$$\gamma_k(B) := \sum_{n=1}^{\infty} 2^{-n}\delta_{n+\frac{1}{k}}(B) \qquad \text{para } B \in \mathfrak{B}(\mathbb{R})$$

donde δ_r denota a la medida de Dirac concentrada en $r \in \mathbb{R}$. Entonces Γ es tensa pero γ^s no lo es, porque para cualquier conjunto compacto C en $\mathbb{R}$ existe un intervalo $I = [m, m+1]$ contenido en C^c y $\gamma^s(I) \geqslant \sigma(I) = \infty$, donde σ y γ^s son como en el Teorema 1.4.

Demostración del Teorema 2.14:

a) Esta parte es trivial por la definición de tensión junto con los Teoremas 2.11 c) y 2.12 d).

b) Por la tensión de Γ, para cada $n = 1, 2, \ldots$, existe un conjunto compacto C_n tal que $\gamma(C_n^c) < \frac{1}{n}$ para toda $\gamma \in \Gamma$, i.e., $\sigma(C_n^c) < \frac{1}{n}$, con σ como en (2.2). Sea $B := \cup C_n$, que es el límite de una sucesión creciente de conjuntos compactos $C_n' \uparrow B$ con $C_n' = \cup_{i=1}^n C_i$. Entonces, por el inciso a) de la Observación 2.13, $\gamma^s(B^c) = 0$. Así resulta que $\gamma^s(B) = \gamma^s(\mathbb{X}) < \infty$. Entonces $\gamma^s(C_n') \uparrow \gamma^s(B) = \gamma^s(\mathbb{X})$. Por lo tanto, $\lim_{n\to\infty}\gamma^s(C_n'^c) = 0$, i.e., γ^s es tensa.▲

Ejemplo 2.15

a) Sea $h \in L^1 := L^1(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \lambda)$ (i.e., $\int_{\mathbb{R}} |h(x)|\,\lambda(dx) < \infty$, donde λ es la medida de Lebesgue en $\mathbb{R}$), tal que $h \geqslant 0$ y h es no decreciente en $(-\infty, 0]$ y h es no creciente en $[0, \infty)$. Sea $f \in L^1$ tal que $0 \leqslant f \leqslant h$ y $\int_{\mathbb{R}} f(x)\,\lambda(dx) = 1$.

Sea $g(x; \mu, \sigma_0) := \frac{1}{\sigma_0} f\left(\frac{x-\mu}{\sigma_0}\right)$, donde $\mu \in \mathbb{R}$ es un parámetro de localización y $\sigma_0 > 0$ es un parámetro de dispersión. Además, sean γ_μ las medidas definidas por las densidades $g(x; \mu, \sigma_0)$ con $\mu \in C \subseteq \mathbb{R}$, a saber, $\gamma_\mu(A) = \int_A g(x; \mu, \sigma_0)\,\lambda(dx)$ para $A \in \mathfrak{B}(\mathbb{R})$. Se puede probar que si el conjunto C es acotado, entonces, por medio de la función h, podemos acotar por arriba a la familia de densidades $\{g(\cdot; \mu, \sigma_0) : \mu \in C\}$ por una función en L^1. Así, en este caso resulta que la envolvente superior γ^s de Γ es una medida finita, y por lo tanto tensa (y entonces el Teorema 2.14 implica que la familia de medidas Γ es tensa). Así, obtenemos el recíproco de la siguiente afirmación:

La envolvente superior γ^s es tensa si y sólo si el conjunto C es acotado.

Demostración:

$\Rightarrow$) Supongamos que C es no acotado superiormente. Sea $\mu_1 \in C \Longrightarrow \exists a_1, b_1 \in \mathbb{R}$ tales que $\gamma_{\mu_1}([a_1, b_1]) > 0$. Como C es no acotado $\exists \mu_2 \in C$ tal que $a_1 := \mu_2 + a_1 - \mu_1 > b_1$ y análogamente definamos a $b_2 := \mu_2 + b_1 - \mu_1$. Repitiendo este argumento encontramos intervalos ajenos $[a_1, b_1], [a_2, b_2], [a_3, b_3]$,... tales que $\gamma_{\mu_1}([a_1, b_1]) = \gamma_{\mu_2}([a_2, b_2]) = \gamma_{\mu_3}([a_3, b_3]) = ...$ Por lo tanto

$$\gamma^s(\cup_i [a_i, b_i]) \geq \sum_i \gamma_{\mu_i}([a_i, b_i]) = \infty.$$

b) Tomemos la siguiente función en L^1:

$$\begin{aligned} h(x) &:= \delta_0 \min\left\{\frac{1}{|x|^{1+\delta_1}}, \frac{1}{|x|^{1-\delta_2}}\right\} \qquad \text{para } x \neq 0, \text{ y} \\ h(x) &:= \infty \qquad \text{para } x = 0 \end{aligned}$$

donde δ_0, δ_1 y δ_2 son constantes positivas, f como en el inciso a). Y definamos a las densidades $g(x; \mu_0, \sigma) := \frac{1}{\sigma} f\left(\frac{x-\mu_0}{\sigma}\right)$ para $\sigma \in C$ donde C es un subconjunto de números positivos y μ_0 es una constante real, y definamos a las medidas $\gamma_\sigma(A) := \int_A g(x; \mu_0, \sigma)\, \lambda(dx)$ para $A \in \mathfrak{B}(\mathbb{R})$ y $\sigma \in C$. Sean $\Gamma := \{\gamma_\sigma : \sigma \in C\}$ y γ^s y γ^i las envolventes superior e inferior de Γ respectivamente. Entonces:

1. $0 < \inf C$ y $\sup C < \infty$ si y sólo si γ^s es tensa.

2. Si $\gamma^i \neq 0$ entonces $0 < \inf C$ y $\sup C < \infty$.

Demostración:

1. $\Rightarrow$ Cálculos directos nos permiten acotar a la familia de densidades $\{g(\cdot; \mu_0, \sigma) : \sigma \in C\}$ por la función $h' \in L^1$ definida como

$$h'(x) := \begin{cases} \delta_0 \frac{s^{\delta_1}}{|x-\mu_0|^{1+\delta_1}} & \text{si } |x - \mu_0| \geq 1 \\ \delta_0 \frac{1}{i^{\delta_2}|x-\mu_0|^{1-\delta_2}} & \text{si } |x - \mu_0| < 1 \end{cases}$$

donde $i = \inf C$ y $s = \sup C$. Así el Teorema 2.11 c) implica que $\gamma^s(A) \leqslant \int_A h'(x)\, \lambda(dx)$ para $A \in \mathcal{B}(\mathbb{R})$. Por lo cual γ^s resulta ser una medida finita en $\mathbb{R}$, y por lo tanto tensa.

$\Longleftarrow$ Supongamos $s = \infty$, y sea $\sigma_1 \in C$. Entonces $\exists a_1, b_1 \in \mathbb{R}$ tales que $\gamma_{\sigma_1}([a_1, b_1]) > 0$. Sin pérdida de generalidad podemos suponer que $a_1 > \mu_0$. Entonces $\exists \sigma_2 \in C$ tal que $b_2 := \sigma_1 \frac{b_1 - \mu_0}{\sigma_2} + \mu_0 < a_1$, y definamos análogamente $a_2 := \sigma_1 \frac{a_1 - \mu_0}{\sigma_2} + \mu_0$, y así sucesivamente, y repetimos el procedimiento de la demostración del inciso a).

c) Sea $\mathbb{X}$ un espacio topológico $\sigma-$compacto y sean Π y μ_n medidas de probabilidad definidas en $\mathcal{B}(\mathbb{X})$, $n = 1, 2, 3, ...$, tales que

$$\|\mu_n - \Pi\| \leqslant a_n \text{ para } n = 1, 2, 3, ...,$$

donde $\|\cdot\|$ es la norma en variación total y la sucesión de números reales $\{a_n\}$ es tal que $\sum a_n < \infty$. Entonces γ^s la envolvente superior de $\Gamma := \{\mu_n : n \in \mathbb{N}\}$, es una medida tensa y, por lo tanto, por el Teorema 2.14, Γ es tensa.

Demostración:

Sea $n \in \mathbb{N}$ y definamos $\mu^{(n)} := \max\{\mu_i : i = 1, 2, ..., n\}$, y

$\mu_i^n := \max\{\mu_j : j = 1, 2, ..., n \text{ y } j \neq i\}$.

Sea A_0 tal que $(\Pi - \mu^{(n)})^+ (\mathbb{X}) = \Pi(A_0) - \mu^{(n)}(A_0)$.

Sea B_i tal que $(\mu_i - \mu_i^n)^+ (\mathbb{X}) = \mu_i(B_i) - \mu_i^n(B_i)$, para $i = 1, 2, ..., n$.

Definamos $A_i := B_i \cap B_{i-1}^c \cap ... \cap B_1^c \cap A_0^c$, para $i = 1, 2, ..., n$.

Como

$$\begin{aligned}
\mu^{(n)}(\mathbb{X}) &= \sum_{i=0}^{n} \mu^{(n)}(A_i) \\
&\leqslant \Pi(A_0) + \Pi(A_1) + a_1 + ... + \Pi(A_n) + a_n \\
&= 1 + a_1 + a_2 + ... + a_n.
\end{aligned}$$

Entonces

$$\gamma^s(\mathbb{X}) = \lim_{n \to \infty} \mu^{(n)}(\mathbb{X}) \leqslant 1 + \sum_{i=1}^{\infty} a_i < \infty.$$

Así resulta que, γ^s es una medida finita en un espacio $\sigma-$compacto, y por lo tanto, tensa.

2.4 PROBLEMAS CON COSTO DESCONTADO

En el resto del capítulo, $\mathbb{E}$ y $\mathbb{A}$ denotarán espacios polacos. Consideremos el modelo de control de Markov (Definición 1.1) con índice de costo descontado (Definición 1.6 a). En esta sección consideraremos a la función de costo por etapa c definida en todo el conjunto $\mathbb{E}\times\mathbb{A}$, por ejemplo, tomando $c(e,a) := \infty$ para todo $(e,a) \notin \mathbb{G}$. Además, usaremos la convención $0\cdot\infty = 0$.

Recordemos que el índice de funcionamiento de costo descontado es

$$I_d(\pi,\nu) := E_\nu^\pi\left(\sum_{t=0}^{\infty}\alpha^t c(e_t,a_t)\right) = \sum_{t=0}^{\infty}\alpha^t E_\nu^\pi\left(c(e_t,a_t)\right), \tag{2.13}$$

donde ν es la distribución inicial, y $\alpha\in(0,1)$ es el factor de descuento. El problema de control de Markov en el que estamos interesados en esta sección consiste en encontrar una política ν-óptima, esto es una política π^* tal que

$$I_d(\pi^*,\nu) = \inf_\pi I_d(\pi,\nu) := I_d(\nu)$$

(véase Definición 1.7).

Como mencionamos en la introducción de este capítulo, nos gustaría describir al criterio de funcionamiento I_d en forma integral como en (2.1). Para hacer esto, nótese que si la función c en (2.13) es reemplazada por la función indicadora $\mathbf{1}_B$ de un conjunto B obtenemos una medida en $\mathbb{E}\times\mathbb{A}$ de la forma siguiente.

Definición 2.16 *Medida de ocupación.*

Sean $\pi\in\mathcal{P}$ una política arbitraria, ν la distribución inicial y $\alpha\in(0,1)$ el factor de descuento. Entonces a la medida

$$\mu^\pi(B) := \sum_{n=0}^{\infty}\alpha^n P_\nu^\pi\left[(e_n,a_n)\in B\right] \qquad B\in\mathfrak{B}(\mathbb{E}\times\mathbb{A})$$

se le llama la medida de ocupación esperada de la política π.

Algunos autores llaman a μ^π una "medida estratégica" porque está determinada por la "estrategia" de control π y la distribución inicial ν. El hecho importante es que podemos escribir (2.13) en la forma

$$I_d(\pi,\nu) = \int_{\mathbb{E}\times\mathbb{A}} c d\mu^\pi. \tag{2.14}$$

Nótese que $\mu^\pi(\mathbb{E}\times\mathbb{A}) = \frac{1}{1-\alpha}$ para toda $\pi \in \mathcal{P}$. De aquí que $(1-\alpha)\mu^\pi$ es una medida de probabilidad en $\mathbb{E}\times\mathbb{A}$, que por la restricción (1.1), está concentrada en el conjunto $\mathbb{G}$. Esto significa en particular que la integral en (2.14) es, en realidad, una integral sobre $\mathbb{G}$.

Necesitaremos la siguiente definición (compare con la Definición 1.9).

Definición 2.17 *Kérnel estocástico débilmente continuo.*

Sean $\mathbb{X}$, $\mathbb{Y}$ espacios topológicos y K un kérnel estocástico en $\mathbb{X}$ dado $\mathbb{Y}$. Decimos que el kérnel estocástico K es débilmente continuo si la función

$$y \mapsto \int_{\mathbb{X}} h(x) K(dx/y)$$

es continua y acotada en $\mathbb{Y}$ para cualquier función h continua y acotada en $\mathbb{X}$.

Observación 2.18

Sea Γ el conjunto de todas las medidas de ocupación μ^π tales que

$$I_d(\pi,\nu) = \int c d\mu^\pi < \infty.$$

En las Hipótesis 2.19, que serán usadas en el Lema 2.21 y en el Teorema 2.22, se requiere que Γ sea un conjunto tenso, no vacío. Más aún, con $\hat{\mu}^\pi(C) := \hat{\mu}^\pi(C\times\mathbb{A})$ para $C \in \mathfrak{B}(\mathbb{E})$, la medida marginal (o proyección) de μ^π en $\mathbb{E}$, en la prueba del inciso a) del Lema 2.21 usamos el hecho de que una medida μ^π en Γ está caracterizada por la relación siguiente (por ejemplo, [26] Teorema 6.3.7)

$$\hat{\mu}^\pi(C) = \nu(C) + \alpha \int_{\mathbb{E}\times\mathbb{A}} K(C/e,a)\, \mu(d(e,a)) \qquad \forall C \in \mathfrak{B}(\mathbb{E})$$

que es equivalente a

$$\int_{\mathbb{E}} h(e)\,\hat{\mu}^{\pi}(de) = \int_{\mathbb{E}} h(e)\,\nu(de) + \alpha \int_{\mathbb{E}\times\mathbb{A}} \int_{\mathbb{E}} h(e')\,K(de'/e,a)\,\mu^{\pi}(d(e,a))$$

para toda función h continua y acotada en $\mathbb{E}$. La implicación recíproca también es cierta; véase el Lema 2.4 b).

Hipótesis 2.19

a) Existe una política $\pi \in \mathcal{P}$ tal que $I_d(\pi,\nu) = \int c d\mu^{\pi} < \infty$,

b) la función c de costo por etapa es no negativa y semicontinua inferiormente,

c) el conjunto Γ, definido en la Observación 2.18, es tenso, y

d) la ley de transición K es débilmente continua.

La Hipótesis 2.19 a) asegura que el problema de control es no trivial; si a) no se satisfaciera, entonces cualquier política sería óptima. Una condición suficiente para a) es que c sea acotada; en este caso a) se cumple para toda política π. Por otro lado, una condición suficiente para la Hipótesis 2.19 b) es

b') La función de costo c es inf.-compacta (Definición 2.2).

Si además se cumple que

c') $\sup \int c(x)\,\mu^{\pi}(dx) < \infty$, donde el supremo es tomado en las políticas que satisfacen a), i.e. en Γ,

entonces, por el Teorema 2.3, obtenemos c). Alternativamente, por el inciso a) del Teorema 2.14, la hipótesis c) se cumple si la envolvente superior γ^s de Γ es tensa, por ejemplo, si γ^s es una medida finita (Ver el Ejemplo 2.15 a). Obsérvese que también se obtiene la hipótesis c) si $\mathbb{E}$ es compacto. Finalmente, d) es la propiedad de Feller, la cual se cumple para muchas clases de sistemas de control. Por ejemplo, si consideramos un sistema no lineal de la forma

$$e_{n+1} = F(e_n, a_n, w_n) \qquad n = 0, 1, 2, \ldots$$

con e_0 dado, y donde $\{w_n\}$ es una sucesión de vectores aleatorios independientes e idénticamente distribuidos con valores en un espacio polaco $\mathbb{W}$, y con distribución común ξ. Además, $F : \mathbb{E} \times \mathbb{A} \times \mathbb{W} \to \mathbb{E}$ es una función medible dada y el estado inicial es independiente de $\{w_n\}$. Aún más, si

$(e,a) \to F(e,a,w)$ es una función continua en $\mathbb{G}$ para todo $w \in \mathbb{W}$, entonces la hipótesis d) se cumple ya que

$$(e,a) \to \int_{\mathbb{E}} h(e') K(de'/e,a) = \int_{\mathbb{W}} h(F(e,a,w))\, \xi(dw)$$

es una función continua y acotada siempre y cuando $h : \mathbb{E} \mapsto \mathbb{R}$ lo sea. Como otro ejemplo, trivial, observe que todas las hipótesis de 2.19 se cumplen cuando $\mathbb{E}$ y $\mathbb{A}$ son conjuntos finitos.

En la demostración del Lema 2.21 necesitaremos ciertos resultados, que resumimos en el siguiente lema (véase [26]). El conjunto Φ que aparece en el siguiente lema se definió en el párrafo que sigue a la Hipótesis 1.2.

Lema 2.20

a) Si μ es una medida de probabilidad definida en $\mathfrak{B}(\mathbb{E} \times \mathbb{A})$ y concentrada en $\mathbb{G}$, entonces existe un kérnel estocástico $\phi \in \Phi$ tal que

$$\mu(C \times D) = \int_C \phi(D/e)\, \hat{\mu}(de) \qquad \forall C \in \mathfrak{B}(\mathbb{E}), D \in \mathfrak{B}(\mathbb{A})$$

donde $\hat{\mu}(C) := \mu(C \times \mathbb{A})$ denota la marginal (o proyección) de μ en $\mathbb{E}$. ([18] pág. 88 o [28], pág. 89).

b) Sea μ una medida en $\mathfrak{B}(\mathbb{E} \times \mathbb{A})$ y concentrada en $\mathbb{G}$ tal que $\int c d\mu < \infty$, $(1-\alpha)\mu(\mathbb{E} \times \mathbb{A}) = 1$ y

$$\hat{\mu}(C) = \nu(C) + \alpha \int_{\mathbb{E}\times\mathbb{A}} K(C/e,a)\, \mu(d(e,a)) \qquad \forall C \in \mathfrak{B}(\mathbb{E}) \qquad (2.15)$$

donde $\hat{\mu}$ es la marginal de μ en $\mathbb{E}$. Desintegre la medida de probabilidad $(1-\alpha)\mu$ como en a), i.e.,

$$\mu(C \times D) = \int_C \phi(D/e)\, \hat{\mu}(de) \qquad \forall C \in \mathfrak{B}(\mathbb{E}), D \in \mathfrak{B}(\mathbb{A})$$

para algún kérnel estocástico $\phi \in \Phi$. Entonces μ es la medida de ocupación (Definición 2.16) correspondiente a la política estacionaria aleatorizada ϕ^∞, esto es $\mu = \mu^{\phi^\infty}$ y $\int c d\mu = I_d(\phi^\infty, \nu)$ ([26], Teorema 6.3.7 b).

c) Si ϕ es un kérnel estocástico en Φ y $v : \mathbb{G} \to \mathbb{R}$ es una función medible tal que $e \to \int v(e,a)\,\phi(da/e)$ es una función (finita) de $\mathbb{E}$ en $\mathbb{R}$, entonces existe un selector medible $f \in \mathfrak{F}$ tal que

$$\int_{\mathbb{A}} v(e,a)\,\phi(da/e) \geqslant v(e, f(e)) \qquad \text{para todo } e \in \mathbb{E}$$

([18], pág. 97).

Lema 2.21

Supóngase que las Hipótesis 2.19 se satisfacen. Entonces

a) Existe una política estacionaria aleatorizada $\pi^* = \phi^\infty$que es ν-óptima, i. e.,

$$I_d(\pi^*, \nu) = I_d(\nu).$$

b) Para cualquier política estacionaria aleatorizada ϕ^∞ tal que $I_d(\phi^\infty, \nu) < \infty$ existe una política estacionaria determinista f^∞ (con $f \in \mathfrak{F}$ un selector medible) tal que

$$I_d(\phi^\infty, \nu) \geqslant I_d(f^\infty, \nu).$$

Demostración:

a) Para cada $k = 1, 2, ...$, sea π^k una política de control tal que

$$I_d(\nu) \leqslant I_d(\pi^k, \nu) \leqslant I_d(\nu) + \frac{1}{k} \tag{2.16}$$

y sea $\mu_k := \mu^{\pi^k}$ la correspondiente medida de ocupación esperada (Definición 2.16). Por la Hipótesis 2.19 c) y el Teorema de Prohorov (Teorema 2.8)[3], existe una subsucesión $\{\mu_m\}$ de $\{\mu_k\}$ y una medida μ en $\mathbb{E} \times \mathbb{A}$ concentrada en $\mathbb{G}$, con

$$\mu(\mathbb{E} \times \mathbb{A}) = \frac{1}{1-\alpha} \tag{2.17}$$

y tal que μ_m converge débilmente a μ.

[3] Véase la observación 2.9.

Dado que c es semicontinua inferiormente [Hipótesis 2.19 b)] y no negativa, existe una sucesión creciente $\{g_j\}$ de funciones continuas y acotadas en $C_a(\mathbb{E}\times\mathbb{A})$ tales que $g_j \uparrow c$. Así que, en particular

$$\int c(x)\,\mu_m(dx) \geqslant \int g_j(x)\,\mu_m(dx) \qquad \forall m, j.$$

Luego por la convergencia débil $\mu_m \to \mu$ y por el Teorema de Convergencia Monótona,

$$\liminf_{m\to\infty} \int c(x)\,\mu_m(dx) \geqslant \int c(x)\,\mu(dx),$$

que combinada con (2.16) implica

$$\int c(x)\,\mu(dx) \leqslant I_d(\nu). \tag{2.18}$$

Ahora, desintegremos μ como $\mu(d(a,e)) = \varphi(da/s)\,\hat{\mu}(ds)$. Así para completar la prueba es suficiente mostrar que μ y $\hat{\mu}$ satisfacen (2.15), porque entonces el Lema 2.4 b) conduciría a $\int c(x)\,\mu(dx) = I_d(\phi^\infty,\nu)$ y la conclusión deseada, 2.21 a), seguiría de (2.18). La verificación de (2.15) se sigue de la convergencia débil de μ_m a μ y de la Hipótesis 2.19 d). En efecto, por la Observación 2.18, podemos escribir

$$\int_{\mathbf{E}} h(e)\,\hat{\mu}(de) = \int_{\mathbf{E}} h(e)\,\nu(de) + \alpha \int_{\mathbf{E}\times\mathbf{A}} \int_{\mathbf{E}} h(e')\,K(de'/e,a)\,\mu_m(d(e,a))$$

para toda función h continua y acotada en $\mathbb{E}$. Entonces haciendo tender m a infinito, la Hipótesis 2.19 d) y la convergencia débil de μ_m a μ conducen a

$$\int_{\mathbf{E}} h(e)\,\hat{\mu}(de) = \int_{\mathbf{E}} h(e)\,v(de) + \alpha \int_{\mathbf{E}\times\mathbf{A}} \int_{\mathbf{E}} h(e')\,K(de'/e,a)\,\mu(d(e,a))$$

Así, como esta igualdad se cumple para toda función h continua y acotada en $\mathbb{E}$, obtenemos (2.15)

b) Sea $\varphi \in \Phi$ tal que la política estacionaria ϕ^∞ satisface que

$$I_d(\phi^\infty,\nu) = \int_{\mathbf{E}} I_d(\phi^\infty,e)\,\nu(de) < \infty \tag{2.19}$$

donde

$$I_d(\phi^\infty, e) := E_e^{\phi^\infty}\left[\sum_{n=0}^{\infty}\alpha^n c(e_n, a_n)\right] \tag{2.20}$$

es el costo esperado descontado cuando se está usando la política estacionaria y el estado inicial es e.

Por (2.19), $I_d(\phi^\infty, e) < \infty$ para toda $e \notin N$ donde N es un conjunto de medida ν-nula. Redefiniendo a $I_d(\phi^\infty, \cdot)$ si fuera necesario, podemos asumir, y lo haremos, que $I_d(\phi^\infty, e) < \infty$ para todo estado inicial $e \in \mathbb{E}$. Ahora, expandiendo y usando la propiedad de Markov, obtenemos

$$\begin{aligned} I_d(\phi^\infty, e) &= E_e^{\phi^\infty}\left[c(e_0, a_0) + \alpha\sum_{n=1}^{\infty}\alpha^{n-1}c(e_n, a_n)\right] \\ &= \int_{\mathbb{A}}\left\{c(e,a) + \alpha\int_{\mathbb{E}} I_d(\phi^\infty, e') K(de'/e, a)\right\}\varphi(da/e) \end{aligned}$$

Entonces, por el Lema 2.4 c), existe un selector medible $f \in \mathfrak{F}$ tal que

$$I(\phi^\infty, e) \geqslant c(e, f)) + \alpha\int_{\mathbb{E}} I_d(\phi^\infty, e') K(de'/e, f)$$

para todo $e \in \mathbb{E}$[4]. Iterando esta expresión y denotando por K^k a la probabilidad de transición en k pasos cuando se está usando la política determinista f^∞, obtenemos

$$\begin{aligned} I(\phi^\infty, e) &\geqslant E_e^{\phi^\infty}\left[\sum_{n=0}^{k-1}\alpha^n c(e_n, f(e_n)) + \alpha^k\int_{\mathbb{E}} I_d(\phi^\infty, e') K^k(de'/e, f(e))\right] \\ &\geqslant E_e^{f^\infty}\left[\sum_{n=0}^{k-1}\alpha^n c(e_n, f(e_n))\right] \qquad \text{para toda } k = 1, 2, 3, \ldots \end{aligned}$$

esta última desigualdad es porque $I_d(\phi^\infty, \cdot) \geqslant 0$.

Así, haciendo tender k a infinito, obtenemos $I(\phi^\infty, e) \geqslant I(f^\infty, e)$ para todo $e \in \mathbb{E}$, e integración con respecto a la distribución inicial ν completa la prueba ▲.

El siguiente teorema es el principal resultado en esta sección y cuya demostración se sigue del Lema 2.21.

[4]Ver Observación 1.3.

Teorema 2.22

Bajo las Hipótesis 2.19 existe una política ν-óptima determinista $\pi^* = f^\infty$ con $f \in \mathfrak{F}$ un selector medible, i.e.,

$$I_d(\pi^*, \nu) = I_d(\nu).$$

En lugar de la Hipótesis 2.19 c) se puede pedir simplemente que esta propiedad se cumpla para las políticas tales que el índice de funcionamiento sea menor o igual que alguna β, i.e., que exista un número real $\beta > I_d(\nu)$ tal que el conjunto de medidas de ocupación μ^π tales que $I_d(\pi, \nu) \leqslant \beta$ sea un conjunto tenso. Aún más, basta con pedir que exista una sucesión de políticas π_n tales que $I_d(\pi_n, \nu) \to I_d(\nu)$ cuando $n \to \infty$ y que la sucesión de medidas de ocupación $\{\mu^{\pi_n}\}$ sea tensa. En cualquiera de estos casos se obtiene la conclusión del Teorema 2.22.

2.5 UN CONTRAEJEMPLO

En esta sección daremos un contraejemplo a una afirmación de Kurano y Kawai ([32], teorema 3.3), quienes usando también medidas de ocupación trataron de dar un teorema de existencia de políticas óptimas, cuya demostración no es del todo correcta. Dichos autores, en lugar de usar la envolvente superior γ^s que introducimos en los Teoremas 2.11 y 2.14, usan otra función de conjuntos desarrollada por Rogers [35].

Necesitaremos la siguiente notación usada en el artículo ([32], pág. 99). Sea $(\overline{e}, \overline{a}) \in \mathbb{G}$, $\overline{n}$ un entero positivo y $\delta > 0$. Entonces $\Pi_{\delta,\overline{e}}^{\overline{n},\overline{a}}$ denota al conjunto de todas las políticas π tales que

$$P_e^\pi\left[e_k = \overline{e} \text{ y } a_k = \overline{a} \text{ para algún } k, \text{ con } 0 \leqslant k \leqslant \overline{n}\,\right] \geqslant \delta$$

para todo estado inicial $e \in \mathbb{E}$.

Hipótesis 2.23

Existen $(\overline{e}, \overline{a}) \in \mathbb{G}$, $\overline{n} \geqslant 0$, $\delta > 0$ tales que para cualquier política $\pi \in \mathcal{P}$ existe una política $\pi' \in \Pi_{\delta,\overline{e}}^{\overline{n},\overline{a}}$ satisface que

$$I_d(\pi', \nu) \leqslant I_d(\pi, \nu).$$

Afirmación 2.24 *(Kurano y Kawai):*

Si se satisface la Hipótesis 2.23 entonces existe una política ν-óptima.

Ejemplo 2.25 *(Contraejemplo a la afirmación 2.24.)*

Considérese el siguiente modelo de control de Markov.
Sean

$$\begin{aligned} \mathbb{E} &: = \{0,1,2,...\} \quad \text{el espacio de estados;} \\ \mathbb{A} &: = \{1,2,3,...\} = A(e) \quad \text{para todo } e \in \mathbb{E}. \end{aligned}$$

La ley de transición:

$$\begin{aligned} K(a/0,a) &= K(0/1,a) = 1 \qquad \text{para toda } a \in \mathbb{A}, \text{ y} \\ K(1/e,a) &= 1 \qquad \text{para toda } a \in \mathbb{A}, \text{ todo } e > 1. \end{aligned}$$

Costo por etapa[5]:

$$\begin{aligned} c(0,a) &= q^a \qquad \text{para algún } 0 < q < 1, \text{ y} \\ c(e,a) &= 0 \qquad \text{si } e > 0. \end{aligned}$$

Sea $\alpha \in (0,1)$ el factor de descuento.
Sea ν la distribución inicial concentrada en $e_0 = 0$. Es claro que tenemos

$$I_d(\pi,0) > 0 \qquad \text{para cualquier política } \pi \in \mathcal{P}. \tag{2.21}$$

Ahora, para cada acción $a \in \mathbb{A}$, defina los siguientes selectores medibles $f_a \in \mathfrak{F}$ como

$$f_a(0) := a, \quad \text{y} \quad f_a(e) := 1 \qquad \text{si } e > 0.$$

Entonces, tomando $(\overline{e},\overline{a}) = (1,1)$, $\overline{n} = 2$ y $\delta = 1$, observamos que las políticas f_a^∞ satisfacen la Hipótesis 2.23, y más aún

$$I_d(f_1^\infty,0) = \frac{q}{1-\alpha^2}, \quad \text{y} \quad I_d(f_a^\infty,0) = \frac{q^a}{1-\alpha^3} \quad \text{si } \ a \geqslant 2.$$

De aquí que $I_d(0) = \inf_\pi I_d(\pi,0) = 0$, y que por (2.21) muestra que la Afirmación 2.24 no es válida.

Kurano y Kawai afirman también en el mismo artículo que si se satisface la Hipótesis 2.23 y el espacio de estados es numerable, entonces existe una política óptima determinista (teorema 3.4 de [32]). Claramente, aquí también funciona el contraejemplo.

[5]Note que c no es una función inf.-compacta.

2.6 PROBLEMAS RESTRINGIDOS CON COSTO DESCONTADO

En la literatura sobre las aplicaciones de los procesos de control de Markov es común encontrar problemas con restricciones. En esta sección y la siguiente discutiremos sobre los procesos de control de Markov con restricciones y con índice de funcionamiento el costo descontado, aplicando los resultados sobre las envolventes de medidas.

Sean V un espacio vectorial sobre $\mathbb{R}$ y $\mathbb{C}$ un cono en V con vértice en el orígen, es decir, $\mathbb{C}$ es un subconjunto no vacío y cerrado bajo la suma de vectores en $\mathbb{C}$ y la multiplicación de un escalar no negativo por un vector en $\mathbb{C}$. Con un cono podemos definir un orden en V en la forma usual, a saber, $u \preceq v$ sí y sólo sí $v - u \in \mathbb{C}$. Consideraremos en este capítulo a $\mathbb{R}^l$ con el orden definido por un cono cerrado $\mathbb{C}$.[6]

Definición 2.26 *Un modelo de control de Markov con restricciones.*

Un modelo de control de Markov con restricciones es la séxtupla ordenada

$$(\mathbb{E}, \mathbb{A}, \{A(e) : e \in \mathbb{E}\}, K, c, d)$$

donde los primeros cinco elementos son los mismos que en la definición de un modelo de control de Markov sin restricciones (Definición 1.1) y el sexto elemento, d, es una función medible en $\mathbb{E} \times \mathbb{A}$, con valores en $\mathbb{R}^l$, y representa el costo sobre el que se imponen restricciones.

Sean I_d, J_d los índices de funcionamiento de la política π con los factores de descuento α, β para los costos c, d, respectivamente, y con distribución inicial ν, i.e.,

$$I_d(\pi, \nu) := E_\nu^\pi \left(\sum_{t=0}^{\infty} \alpha^t c(e_t, a_t) \right) \quad \text{y} \tag{2.22}$$

$$J_d(\pi, \nu) := E_\nu^\pi \left(\sum_{t=0}^{\infty} \beta^t d(e_t, a_t) \right). \tag{2.23}$$

[6]Un cono cerrado $\mathbb{C}$ es un cono que es un subconjunto cerrado $\mathbb{R}^l$ con la topología usual.

Definición 2.27 *Problema de control de Markov con costo descontado y restricciones.*

Consideremos el modelo de control de Markov con restricciones (Definición 2.26). Entonces el problema de control de Markov con costo descontado y restricciones consiste en encontrar una política $\pi^* \in \mathcal{P}$ tal que

$$I_d(\pi^*, \nu) = \inf\{I_d(\pi, \nu) : \pi \in \mathcal{P} \text{ y } 0 \preceq J_d(\pi, \nu) \preceq k\}$$

donde ν es la distribución inicial, k es un vector dado en $\mathbb{R}^l_+$ y $\preceq$ es el orden definido por un cono $\mathbb{C}$.

Observe que una restricción de la forma $k_1 \preceq J_d \preceq k_2$ con $k_1, k_2 \in \mathbb{R}^l$ tales que $k_1 \preceq k_2$, puede ser reducida a la forma $0 \preceq J_d \preceq k$, tomando la siguiente función $d'(e,a) := d(e,a) - (1-\beta)k_1$ y $k := k_2 - k_1$.

Con el siguiente ejemplo exhibimos la diferencia en los problemas de control con costo descontado sin restricciones y con restricciones ya que en este último caso podemos tener las siguientes situaciones (véase Teorema 1.10):

1. que la política óptima dependa del estado inicial,

2. que la función de valor no cumpla la ecuación de optimalidad (Definición 1.8) y

3. que no exista una política óptima determinista aunque sí exista una política óptima estacionaria aleatorizada.

Ejemplo 2.28

Sean

$\mathbb{E} = \{0, 1, 2, ...\}$ el espacio de estados,

$\mathbb{A} = A(e) = \{0, 1\}$ $\forall e \in \mathbb{E}$ el espacio de acciones,

$K(e+1/e, a) = 1$ para $e \in \mathbb{E}$, $a \in \mathbb{A}$, el kérnel estocástico,

$c(e,0) = 0$, $c(e,1) = 1$ para $e \in \mathbb{E}$, el costo por etapa y

$d(e,0) = 1$, $d(e,1) = 0$ para $e \in \mathbb{E}$, el costo sobre el que se imponen restricciones. También tenemos

$$\alpha = \beta \in (0,1)\,, \text{ los factores de descuento.} \tag{2.24}$$

Entonces para los índices de funcionamiento I_d, J_d en (2.22) y (2.23) con $l = 1$ y el orden típico de los números reales, se tiene la relación

$$I_d(\pi, e) = \frac{1}{1-\alpha} - J_d(\pi, e) \tag{2.25}$$

por lo que para minimizar I_d debemos de maximizar J_d.

Observe que este ejemplo satisface las hipótesis del Teorema 1.10.

a) PROBLEMA 1.

$$\begin{aligned} &\text{Minimizar } I_d(\pi, e), \\ \text{sujeto a } : \quad & 0 \leqslant J_d(\pi, e) \leqslant 1, \\ \text{con } \pi \in \mathcal{P}. \end{aligned}$$

Si e es el estado inicial, entonces una política óptima está dada por el selector medible $f_e \in \mathfrak{F}$ definido como:

$$f_e(x) = \begin{cases} 0 & \text{si } x = e, \\ 1 & \text{si } x \neq e. \end{cases}$$

Así, la política óptima $\pi^* = f_e^\infty = (f_e, f_e, ...)$ depende del estado inicial e. Además, el valor óptimo es

$$I_d(e) = I_d(\pi^*, e) = \frac{\alpha}{1-\alpha} \tag{2.26}$$

con

$$J_d(\pi^*, e) = 1.$$

b) La ecuación de optimalidad para el costo descontado es (Definición 1.8)

$$v(e) = \min_{a \in A(e)} \left\{ c(e, a) + \alpha \int_{\mathbf{E}} v(x) K(dx/e, a) \right\}$$

que en nuestro caso, sustituyendo la función de valor (2.26), obtenemos

$$\frac{\alpha}{1-\alpha} = \min_{a \in A(e)} \left\{ c(e, a) + \alpha \frac{\alpha}{1-\alpha} \right\}.$$

Sin embargo, la única solución de esta ecuación es $\alpha = 0$, lo cual contradice (2.24), pues $\alpha \in (0, 1)$ Así pues, en este caso la función de valor no es solución de la ecuación de optimalidad, i.e., no se satisface el principio de optimalidad de Bellman.

c) PROBLEMA II.

Sea $\alpha = \frac{1}{4}$.

$$\begin{aligned} &\text{Minimizar } I_d(\pi, e), \\ \text{sujeto a } : \quad &0 \leqslant J_d(\pi, e) \leqslant \frac{1}{2}, \\ \text{con } \pi \in &\ \mathcal{P}. \end{aligned} \tag{2.27}$$

Si nos restringimos por el momento a políticas deterministas con $\pi = f^{\infty}$ con $f \in \mathfrak{F}$ un selector medible, entonces la relación (2.25) y la restricción (2.27) implican que el valor óptimo en la clase de las políticas deterministas es

$$I_d(\pi, e) = 1 \text{ con } J_d(\pi, e) = \frac{1}{3}. \tag{2.28}$$

Ahora consideremos la política estacionaria no determinista $\pi^* = \phi^{\infty}$ con $\phi \in \Phi$ tal que

ϕ toma la acción 0 con probabilidad $\frac{1}{2}$, si $x = e$,
ϕ toma la acción 1 con probabilidad $\frac{1}{2}$, si $x = e$,
ϕ toma la acción 1 con probabilidad 1, si $x \neq e$.

Entonces, la relación (2.25) implica que π^* es una política estacionaria óptima, donde

$$I_d(\pi^*, e) = \frac{5}{6} \quad \text{con} \quad J_d(\pi^*, e) = \frac{1}{2}. \tag{2.29}$$

Así, el óptimo no se alcanza dentro de las políticas deterministas.

d) Finalmente, veamos que en el caso del problema II pueden existir una cantidad no numerable de políticas óptimas.

Sea B un subconjunto no vacío de números naturales, y definamos

$$C(B) := \sum_{n \in B} \left(\frac{1}{4}\right)^n.$$

Consideremos las políticas estacionarias aleatorizadas dadas por las reglas de decisión $\phi_B(a/x)$ tal que

ϕ_B toma la acción 0 con probabilidad $\frac{1}{2} - C(B)$, si $x = e$,
ϕ_B toma la acción 1 con probabilidad $\frac{1}{2} + C(B)$, si $x = e$,
ϕ_B toma la acción 0 con probabilidad 1, si $x \in B + e$,

ϕ_B toma la acción 1 con probabilidad 1, si $x \notin B + e$ y $x \neq e$.

Entonces, para cualquier conjunto $B \subset \mathbb{N}$, la política $\pi_B^* = (\phi_B, \phi_B, ...)$ satisface (2.29), i.e.,

$$I_d(\pi_B^*, e) = \frac{5}{6} \qquad \text{con } J_d(\pi_B^*, e) = \frac{1}{2}.$$

Es decir, todas las políticas π_B^* son óptimas y, por lo tanto, existe una cantidad no numerable de políticas óptimas.

2.7 EXISTENCIA DE POLÍTICAS ÓPTIMAS

En esta sección discutimos sobre la existencia de políticas óptimas para el problema restringido y con índice de funcionamiento el costo descontado. Comenzamos proponiendo las Hipótesis 2.29, que son comentadas en la Observación 2.30. En el Teorema 2.31 demostramos la existencia de políticas óptimas estacionarias aleatorizadas, haciendo uso de los siguientes resultados vistos en el capítulo: el Teorema 2.3, el Teorema de Prohorov (Teorema 2.8), los Teoremas sobre las envolventes de medidas 2.11 y 2.14 y los Lemas 2.4 y 2.21.

Las siguientes hipótesis, que serán usadas en el Teorema 2.31, son las análogas para el caso restringido a las Hipótesis 2.19 que fueron usadas en el Teorema 2.22.

Hipótesis 2.29

Supóngase que k es una constante en $\mathbb{R}^l$, ν es la distribución inicial y $\preceq$ es el orden definido por un cono cerrado $\mathbb{C}$. Además:

a) Existe una política $\pi \in \mathcal{P}$ tal que $0 \preceq J_d(\pi, \nu) \preceq k$ e $I_d(\pi, \nu) < \infty$.

b) La función de costo c en un paso es no negativa y semicontinua inferiormente, la función de costo d es continua y acotada.

c) La ley de transición K es débilmente continua (Definición 2.17).

d) El conjunto Γ de medidas de ocupación (Definición 2.16) tales que $0 \preceq I_d(\pi, \nu) \preceq k$ e $I_d(\pi, \nu) < \infty$ es tenso.

Observación 2.30

a) La Hipótesis 2.29 a) es para asegurar que el problema de control esté bien definido; si no se cumpliera esta propiedad, entonces todas las políticas que cumplieran la restricción $0 \preceq J_d(\pi, \nu) \preceq k$ serían óptimas.

b) En lugar de tomar a la función de costo c no negativa se puede tomar a c acotada por abajo, y en lugar de pedir que la restricción para la función de costo d sea de la forma $0 \preceq J_d(\pi, \nu) \preceq k$ se puede tomar en su lugar una restricción de la forma $k_1 \preceq J_d(\pi, \nu) \preceq k_2$ con $k_1, k_2 \in \mathbb{R}^l$, y $k_1 \preceq k_2$.

c) Las Hipótesis 2.29 al ser análogas a las Hipótesis 2.19 nos aseguran que podemos aplicar los tres incisos del Lema 2.4. Así podemos seguir la argumentación del Lema 2.4 para encontrar una política óptima π^* y que al mismo tiempo satisfaga la restricción $0 \preceq J_d(\pi^*, \nu) \preceq k$.

d) Una condición suficiente para que la Hipótesis 2.29 d) se cumpla es que la función de costo c sea inf$-$compacta, y si además el $\sup_{\pi \in \mathcal{P}} I_d(\pi, \nu) < \infty$ sobre la familia Γ definida en la Hipótesis 2.29 d), entonces por el Teorema 2.3 obtenemos la Hipótesis 2.29 d). La Hipótesis 2.29 d) también se cumple si el espacio de estados $\mathbb{E}$ es compacto. Para más comentarios véase el párrafo que sigue a las Hipótesis 2.19.

Teorema 2.31

Bajo las Hipótesis 2.29, existe una política óptima estacionaria aleatorizada $\pi^* = \phi^\infty$.

Demostración.
Sea $I_d(\nu) := \inf_{\pi \in \mathcal{P}} \{I_d(\pi, \nu) : 0 \preceq J_d(\pi, \nu) \preceq k\}$
Para cada $n = 1, 2, 3, ...$, sea π^n una política de control en Γ (definida en la Hipótesis 2.29 d) tal que $I_d(\nu) \leqslant I_d(\pi^n, \nu) \leqslant I_d(\nu) + \frac{1}{n}$.
Sea μ^n la medida de ocupación de π^n (Definición 2.16).
Como Γ es tensa, por el Teorema de Prohorov (Teorema 2.8), existe una subsucesión $\{\mu^j\}$ de $\{\mu^n\}$ y existe una medida μ definida en $\mathfrak{B}(\mathbb{E} \times \mathbb{A})$ tal que $\mu^j \underset{j \to \infty}{\longrightarrow} \mu$ débilmente (Definición 2.7).

Desintegremos a la medida μ como en el Lema 2.4 a), es decir $\mu(d(e,a)) = \phi(da/e)\,\hat{\mu}(de)$, donde $\hat{\mu}$ es la medida marginal de μ en $\mathbb{E}$ y $\phi \in \Phi$. Así, como en la demostración del Lema 2.21, resulta que μ es la medida de ocupación de la política estacionaria aleatorizada $\pi^* = \phi^\infty$. Esto junto con la convergencia débil de μ^j a μ y la Hipótesis 2.29 b) implican que $J_d(\pi^j,\nu) \underset{j\to\infty}{\longrightarrow} J_d(\pi^*,\nu)$ con $0 \preceq J_d(\pi^j,\nu) \preceq k$ para todo $j = 1,2,3,\ldots$. De aquí resulta que $J_d(\pi^*,\nu)$ es un punto de acumulación del "intervalo" $\left\{x \in \mathbb{R}^l : 0 \preceq x \preceq k\right\}$ que por ser un conjunto cerrado contiene al punto $J_d(\pi^*,\nu)$.

También obtenemos, siguiendo la demostración del Lema 2.21, que la política estacionaria aleatorizada $\pi^* = \phi^\infty$, con $\phi \in \Phi$, es óptima, i.e.

$$\begin{aligned} I_d(\pi^*,\nu) &= I_d(\nu) \\ 0 &\preceq J_d\left(\pi^{j'},\nu\right) \preceq k. \blacktriangle \end{aligned}$$

COMENTARIO FINAL AL CAPÍTULO 2

En este capítulo encontramos condiciones suficientes, vía las medidas de ocupación, para que el problema de control de Markov tenga soluciones óptimas estacionarias cuando el índice de funcionamiento es el costo descontado. En el caso no restringido, dicha soluciones son estacionarias, mientras que en el caso restringido resultan ser aleatorizadas. Una de las hipótesis básicas, la Hipótesis 2.19 c) se puede verificar usando la envolvente superior de una familia de medidas, construida en la Sección 2.3. Además, dimos un contraejemplo a la solución dada por Kurano y Kawai al mismo problema de control de Markov no restringido. En el Ejemplo 2.28 mostramos las diferencias entre los dos casos. En el siguiente capítulo veremos otro enfoque para estudiar el problema de control vía la programación lineal, cuando el índice de funcionamiento es el costo promedio.

Capítulo 3

PROGRAMACIÓN LINEAL INFINITA

3.1 INTRODUCCIÓN

La programación lineal es una técnica común para el estudio de diferentes clases de problemas de control (por ejemplo [14], [26], [37], [38]). En particular, como se muestra en los comentarios históricos en [2], [26], [30] la programación lineal ha sido usada desde principios de los años sesenta para el estudio de los procesos de control de Markov (PCMs), pero, con excepción de unos pocos artículos (véase [24] y [25] y para el índice de costo descontado [22]), todos ellos tratan con espacios numerables, principalmente espacios finitos. En este capítulo usamos la programación lineal infinita para estudiar los procesos de control de Markov en espacios de Borel con costo promedio para el caso unicadena.

De hecho, en un trabajo reciente [23] también estudiamos el caso "multicadena", pero no lo incluimos aquí, en parte porque el caso unicadena basta para ilustrar el enfoque de la programación lineal y en parte porque el problema multicadena requiere muchos preliminares técnicos. (Hasta donde nosotros sabemos, el trabajo [23] es el único que estudia el problema multicadena en espacios de Borel no-numerables; para el caso numerable, véase [29].)

El contenido del capítulo es el siguiente: En la sección 3.2 presentamos los conceptos básicos de la programación lineal infinita y mencionamos los teoremas relacionados que usaremos en las demostraciones, en particular, el

Teorema Generalizado de Farkas (Teorema 3.4) y el Teorema de Banach-Alaoglu-Bourbaki (Teorema 3.5). En la sección 3.3 presentamos el Programa Lineal P asociado al problema de control (Definición 1.7) con costo promedio, para el caso unicadena, y también presentamos a su Programa Dual P^*. En la sección 3.4 proponemos dos conjuntos de Hipótesis, 3.8 y 3.9, y las discutimos en la Observación 3.10; dichas hipótesis las usaremos en los resultados más importantes de este capítulo, Teoremas 3.13 y 3.15. En la sección 3.5, usando el Teorema Generalizado de Farkas (Teorema 3.4), damos condiciones necesarias y suficientes para que el Programa Lineal P sea consistente, (Teorema 3.13). En la sección 3.6 demostramos en el Teorema 3.15 que: a)El Programa Lineal P es resoluble y que no hay abertura dual; b) las sucesiones minimizantes para P son débilmente precompactas y c) si existe una sucesión maximizante para P^* acotada en la norma con peso w_0, entonces el Programa Dual P^* es resoluble y se satisface la dualidad fuerte, y más aún, existen una función $h \in F(\mathbb{E})$ y un selector medible $f \in \mathfrak{F}$ que satisfacen las ecuaciones de optimalidad (3.2) y (3.3).

Terminología y Notación

Si $\mathbb{X}$ es un espacio métrico, $C_a(\mathbb{X})$ denota el espacio de Banach de las funciones continuas y acotadas con la norma del supremo $\|u\| := \sup_x |u(x)|$, y denotamos por $C_0(\mathbb{X})$ al subespacio de las funciones continuas que se anulan en el infinito. Si $\mathbb{X}$ es compacto, entonces $C_a(\mathbb{X}) = C_0(\mathbb{X})$.

En cualquier espacio vectorial, al vector nulo lo denotaremos por "0". Así $\mu(\cdot) = 0$ es la medida trivial y $u(\cdot) = 0$ es la función nula.

3.2 CONCEPTOS BÁSICOS

En esta sección daremos las definiciones y resultados de programación lineal infinita que usaremos en este capítulo.

Sean $\mathbb{X}, \mathbb{Y}$ dos espacios vectoriales reales con una forma bilineal $\langle \cdot, \cdot \rangle$ definida en $\mathbb{X} \times \mathbb{Y}$, tal que $\forall x \in \mathbb{X}$ con $x \neq 0$ $\exists y \in \mathbb{Y}$ tal que $\langle x, y \rangle \neq 0$, y $\forall y \in \mathbb{Y}$ con $y \neq 0$ $\exists x \in \mathbb{X}$ tal que $\langle x, y \rangle \neq 0$. Entonces decimos que $(\mathbb{X}, \mathbb{Y})$ es un par dual de espacios vectoriales con producto interno $\langle x, y \rangle$ para $x \in \mathbb{X}$, $y \in \mathbb{Y}$. En este caso denotamos por $\sigma(\mathbb{X}, \mathbb{Y})$ a la topología débil en $\mathbb{X}$. Es decir, si $\{x_n\}$ es una sucesión en el sentido usual o una sucesión generalizada, decimos que $\{x_n\}$ converge a $x \in \mathbb{X}$ en la topología débil $\sigma(\mathbb{X}, \mathbb{Y})$, o que "$x_n \to x$ débilmente" si $\langle x_n, y\rangle \to \langle x, y\rangle$ $\forall y \in \mathbb{Y}$. Si $\mathbb{Y}$ en un espacio de Banach y $\mathbb{X} = \mathbb{Y}^*$, el dual topológico, entonces $\sigma(\mathbb{X}, \mathbb{Y})$ es llamada la topología débil-* en $\mathbb{X}$ ([1], [12] y [17]).

Sea $(\mathbb{X}, \mathbb{Y})$ un par dual y $\mathfrak{C}$ un cono convexo en $\mathbb{X}$, es decir, un subconjunto no vacío de $\mathbb{X}$ cerrado bajo la suma de vectores en $\mathfrak{C}$ y la multiplicación de un escalar no negativo por un vector en $\mathfrak{C}$. Entonces el cono dual $\mathfrak{C}^*$ en $\mathbb{Y}$ es el conjunto $\{y \in \mathbb{Y} : \langle x, y\rangle \geqslant 0 \quad \forall x \in \mathfrak{C}\}$.

Sean $(\mathbb{X}, \mathbb{Y})$, $(\mathbb{Z}, \mathbb{W})$ dos pares duales reales, $\mathfrak{C}$ un cono convexo en $\mathbb{X}$, $G : \mathbb{X} \to \mathbb{Z}$ una función lineal débilmente continua, $b \in \mathbb{Z}$, y $c \in \mathbb{Y}$ dos vectores dados. Considere el Programa Lineal:

$$\begin{aligned} P \ &: \ \text{minimizar } \langle x, c\rangle \\ \text{sujeto a } Gx \ &= \ \text{b} \\ x \ &\in \ \mathfrak{C}. \end{aligned}$$

El programa lineal dual de P es

$$\begin{aligned} P^* \ &: \ \text{maximizar } \langle b, w\rangle \\ \text{sujeto a } c - G^* w \ &\in \ \mathfrak{C}^* \\ w \ &\in \ \mathbb{W}, \end{aligned}$$

donde G^* es la función lineal dual de G.

Decimos que $x \in \mathfrak{C}$ es una solución factible para P si satisface que $Gx = b$. Al ínfimo de $\langle x, c\rangle$ sobre todas las soluciones factibles para P se le llama el valor de P y lo denotamos por $\inf P$. Además, cuando este ínfimo se alcanza decimos que el Programa Lineal P es soluble, y en este caso usamos la convención usual de reemplazar al ínfimo por el mínimo. Se definen análogamente: solución factible para P^*, $\sup P^*$ y $\max P^*$.

Si existen x factible para P y w factible para P^*, entonces se puede ver que $\langle b, w\rangle \leqslant \langle x, c\rangle$. De aquí se sigue la propiedad de dualidad débil: $\sup P^* \leqslant \inf P$. Si se cumple la desigualdad estricta se dice que hay abertura dual. Cuando se cumple la igualdad, es decir, cuando no existe abertura de dualidad, y además P^* y P son resolubles, es decir, $\max P^* = \min P$ se dice que hay dualidad fuerte.

Definición 3.1

Sea $\mathbb{H}$ el subconjunto de $\mathbb{Z} \times \mathbb{R}$ definido como

$$\mathbb{H} := \{(Gx, \langle x, c\rangle + r) : x \in \mathbb{C}, r \geqslant 0\}$$

El programa P se dice que es subconsistente si existe algún $r \in \mathbb{R}$ con $(b, r) \in \overline{\mathbb{H}}$ (la cerradura débil de $\mathbb{H}$). Si P es subconsistente, se define su subvalor como el ínfimo de todos los reales $r \in \mathbb{R}$ para los cuales $(b, r) \in \overline{\mathbb{H}}$ ([1] página 40).

Teorema 3.2

P es subconsistente con un subvalor finito r si y sólo si el dual P^* es consistente con un valor finito $\sup P^*$ ([1] teorema 3.3).

Teorema 3.3

Si P es consistente con un valor finito y $\mathbb{H}$ es débilmente cerrado, entonces P es resoluble y no hay abertura dual para P; es decir, $\min P = \sup P^*$ ([1] teoremas 3.9 y 3.22).

Usaremos el siguiente teorema para dar condiciones suficientes y necesarias para que P sea consistente.

Teorema 3.4 *Teorema Generalizado de Farkas*

Sean $(\mathbb{X}, \mathbb{Y})$ y $(\mathbb{Z}, \mathbb{W})$ dos pares duales, sea $\mathfrak{C}$ un cono convexo en $\mathbb{X}$, y sea $G : \mathbb{X} \to \mathbb{Z}$ una función lineal débilmente continua. Si $G(\mathfrak{C})$ es débilmente cerrado, entonces las siguientes condiciones sobre $b \in \mathbb{Z}$ son equivalentes (donde G^* denota al adjunto de G) :

a) La ecuación $Gx = b$ tiene una solución $x \in \mathfrak{C}$.

b)$G^*w \in \mathfrak{C}^* \Rightarrow \langle b, w\rangle \geqslant 0$.

([13] teorema 2).

El siguiente resultado en combinación con la Observación 3.12 será usado en la demostración de los Teoremas 3.13 y 3.15.

Teorema 3.5 *Alaoglu o Banach-Alaoglu-Bourbaki*

Sea $\mathbb{X}$ un espacio de Banach con dual topológico $\mathbb{X}^*$ y sea $\mathbb{U}$ la esfera unitaria cerrada en $\mathbb{X}^*$. Entonces $\mathbb{U}$ es compacta con la topología debil-* $\sigma(\mathbb{X}^*, \mathbb{X})$. Más aún, si $\mathbb{X}$ es separable, entonces la topología débil-* de $\mathbb{U}$ es metrizable ([12] y [17]).

3.3 PROBLEMAS CON COSTO PROMEDIO

En esta sección presentaremos los programas lineales infinitos P y su dual P^* asociados al problema de control (Definición 1.7), con índice de funcionamiento el costo promedio (Definición 1.6 b)) para el caso unicadena.

Recordemos primero que el índice de funcionamiento de costo promedio, Definición 1.6 b), es

$$\begin{aligned} I_p(\pi,\nu) \; &:= \limsup_{N\to\infty} \frac{1}{N} I_N(\pi,\nu) \\ &= \limsup_{N\to\infty} E_\nu^\pi \left(\sum_{t=0}^{N-1} c(e_t,a_t) \right), \end{aligned}$$

donde ν es la distribución inicial. El problema de control de Markov (Definición 1.7) en el que estamos interesados en este capítulo es encontrar una política ν-óptima, esto es, una política π^* tal que

$$I_p(\pi^*,\nu) = \inf_\pi I_p(\pi,\nu) =: I_p(\nu) \qquad (=\rho^*), \tag{3.1}$$

donde la función de valor $I_p(\nu)$ es igual a una constante ρ^* (Teorema 1.12). El enfoque usual de la programación dinámica consiste en encontrar una terna canónica (ρ^*, h, f), tal que el par (ρ^*, h) satisface la ecuación de optimalidad con costo promedio (3.2) y el selector medible $f \in \mathfrak{F}$ alcanza el mínimo en (3.2) para todo estado inicial $e \in \mathbb{E}$, y tomando en cuenta la notación de la Observación 1.3, se obtienen las siguientes ecuaciones para todo $e \in \mathbb{E}$ (Teorema 1.12)

$$\rho^* + h(e) = \inf_{a\in A(e)} \left\{ c(e,a) + \int_{\mathbb{E}} h(y)\, K(dy/e,a) \right\} \tag{3.2}$$

$$\rho^* + h(e) = c(e,f) + \int_{\mathbb{E}} h(y)\, K(dy/e,f). \tag{3.3}$$

Finalmente, si h satisface la Hipótesis (1.13), es decir, $\limsup_{n\to\infty} E_e^\pi(h(e_n))/n = 0$ para toda $\pi \in \mathcal{P}$ y $e \in \mathbb{E}$, entonces f^∞ es

e-óptima para toda $e \in \mathbb{E}$ (Teorema 1.14). Ahora veremos como el problema de control de Markov se puede plantear como un problema de programación lineal.

Introduciremos los pares duales $(M(\mathbb{G}), F(\mathbb{G}))$ y $(M(\mathbb{E}), F(\mathbb{E}))$ (como en los artículos [23] y [24]; véanse también [25] o [26]) en la Definición 3.6, donde usamos las siguientes funciones de peso o ponderación

$$w(e,a) := 1 + c(e,a) \tag{3.4}$$

$$w_0(e) := \inf\{w(e,a) : a \in A(e)\} \tag{3.5}$$

Dado que c es no negativa, Hipótesis 1.2 b), tenemos que

$$1 \leqslant w_0(e) \leqslant w(e,a) \qquad \forall (e,a) \tag{3.6}$$

y, ya sea bajo las Hipótesis 3.8 ó 3.9, w_0 es medible.

Definición 3.6 *Los pares duales.*

a) $M(\mathbb{G})$ denota al espacio vectorial normado de medidas signadas finitas μ definidas en $\mathbb{G}$ con la norma ponderada

$$\|\mu\|_w := \int_{\mathbf{G}} w(x)\,|\mu|(dx) < \infty,$$

donde $|\mu| = \mu^+ + \mu^-$ es la medida de variación total de μ, y $F(\mathbb{G})$ denota al espacio vectorial normado de las funciones medibles u: $\mathbb{G} \to \mathbb{R}$ con la norma ponderada

$$\|u\|_w := \sup{}_{(e,a)} \frac{|u(e,a)|}{w(e,a)} < \infty.$$

$(M(\mathbb{G}), F(\mathbb{G}))$ es un par dual con respecto a la forma bilineal

$$\langle \mu, u \rangle := \int_{\mathbf{G}} u d\mu$$

con $\mu \in M(\mathbb{G}), u \in F(\mathbb{G})$.

b) Similarmente, reemplazando $\mathbb{G}$ y w por $\mathbb{E}$ y w_0, respectivamente obtenemos el par dual $(M(\mathbb{E}), F(\mathbb{E}))$.

De hecho, consideraremos varias topologías en $M(\mathbb{G})$ (véase la Observación 3.12) pero a menos que se diga lo contrario, consideraremos siempre a $M(\mathbb{G})$ con la topología débil $\sigma(M(\mathbb{G}), F(\mathbb{G}))$ y análogamente para $M(\mathbb{E})$. Denotaremos por $M_+(\mathbb{G})$ y $M_+(\mathbb{E})$ a los conos convexos de medidas no negativas definidas en $\mathbb{G}$ y $\mathbb{E}$ respectivamente .

Ahora definiremos dos operadores lineales L_1 y L que resultarán ser débilmente continuos bajo las Hipótesis 3.8 o 3.9.

$$\begin{aligned} L_1 &: M(\mathbb{G}) \to M(\mathbb{E}), \\ \mu &\mapsto L_1\mu(B) := \hat{\mu}(B) - \int K(B/x,a)\,\mu(d(x,a)), \end{aligned}$$

y

$$\begin{aligned} L &: M(\mathbb{G}) \to \mathbb{R}\times M(\mathbb{E}), \\ \mu &\mapsto L\mu := (\langle \mu, 1\rangle, L_1\mu) \end{aligned} \tag{3.7}$$

donde $\langle \mu, 1\rangle = \int\limits_{\mathbf{G}} d\mu = \mu(\mathbb{G})$ y $\hat{\mu}$ es la marginal de μ en $\mathbb{E}$.

Los correspondientes operadores adjuntos son $L_1^* : F(\mathbb{E}) \to F(\mathbb{G})$ tal que

$$\begin{aligned} (L_1^* g)(e,a) &: = g(e) - \int_{\mathbf{G}} g(y)\,K(dy/e,a) \\ &\text{y } L^* : \mathbb{R}\times F(\mathbb{E}) \to F(\mathbb{G}) \text{ tal que } L^*(\rho,g) := \rho + L_1^* g. \end{aligned}$$

[1]

Con esta notación, el programa lineal primal asociado al problema de control con costo promedio para el caso unicadena es

$$P : \text{minimizar } \langle \mu, c\rangle$$

$$\text{sujeto a } L\mu = (1,0) \text{ con } \mu \in M_+(\mathbb{G}). \tag{3.8}$$

El programa dual de P es:

$$P^* : \text{maximizar } \langle (1,0), (\rho,h)\rangle$$

$$\text{sujeto a } L^*(\rho,h) \leqslant c, \quad (\rho,h) \in \mathbb{R}\times F(\mathbb{E}). \tag{3.9}$$

[1]Donde $(\rho,g) \mapsto (\rho,h)$ como en (3.9).

Observación 3.7

a) Las igualdades y desigualdades se entienden componente a componente. Por ejemplo, en (3.8), $L\mu = (1, 0)$, significa $\langle \mu, 1 \rangle = 1$ y $L_1\mu = 0$. Observe que si μ es factible para P, entonces μ es una medida de probabilidad.

b) Una función $h \in F(\mathbb{E})$ será identificada con la función en $F(\mathbb{G})$ constante en las acciones, es decir $h(e, a) \equiv h(e)$. Así, podemos escribir $F(\mathbb{E}) \subseteq F(\mathbb{G})$. Esto puede ser hecho porque, por (3.6), $\|h\|_w \leqslant \|h\|_{w_0} < \infty$ si $h \in F(\mathbb{E})$. De aquí que también podemos escribir $\langle \mu, h \rangle = \langle \hat{\mu}, h \rangle$ para toda función $h \in F(\mathbb{E})$ y $\mu \in M(\mathbb{G})$, donde $\hat{\mu}$ es la marginal de μ en $\mathbb{E}$.

c) La definición de las funciones de peso w y w_0 (3.4) y (3.5) es útil porque se satisface automáticamente que $c \in F(\mathbb{G})$ y también la desigualdad (3.6). Sin embargo, es importante observar que en el desarrollo de este capítulo se pueden tomar cualesquiera funciones de peso w y w_0 siempre y cuando c esté en $F(\mathbb{G})$ y se cumplan (3.6) y también la Hipótesis 3.8c) (que es la misma que la Hipótesis 3.9d)). Similarmente la Hipótesis 3.8 a) a ser presentada más adelante puede ser reemplazada por la condición de que c sea semicontinua inferiormente y que la función de peso w sea inf-compacta. Esta condición junto con la Hipótesis 3.8 d) asegura, por ejemplo, que el conjunto de medidas μ que satisfacen (3.8) y $\int c d\mu < \infty$ sea tenso (hipótesis 5.1 c. en [24]).

3.4 LAS HIPÓTESIS

En esta sección proponemos dos conjuntos de Hipótesis 3.8 y 3.9 las cuales serán usadas en los resultados más importantes de este capítulo, Teoremas 3.13 y 3.15.

Hipótesis 3.8

a) La función de costo c es inf-compacta.

b) La ley de transición K es débilmente continua (Definición 2.17).

c) $\int_{\mathbb{E}} w_0(y)\, K(dy/\cdot)$ está en $F(\mathbb{G})$, i.e., existe una constante k tal que

$$\int_{\mathbb{E}} w_0(y)\, K(dy/e, a) \leqslant k w(e, a) \qquad \forall (e, a) \in \mathbb{G}.$$

d) Existe una política π tal que para toda distribución inicial ν el costo promedio $I_p(\pi,\nu)$ es finito, o equivalentemente, por (3.4),

$$\limsup_{n\to\infty}\frac{1}{n}E_\nu^\pi\left(\sum_{t=0}^{n-1}w(e_t,a_t)\right)<\infty.$$

Hipótesis 3.9

a) $\mathbb{E}$ y $\mathbb{G}$ son espacios métricos separables y localmente compactos.

b) El costo en una etapa c es semicontinuo inferiormente.

c) K es débilmente continua (Definición 2.17) y, además, para todo subconjunto compacto C de $\mathbb{E}$ la función $K(C/\cdot)$ se anula en el infinito, es decir, para todo $\varepsilon>0$ existe un conjunto compacto $C'=C'(\varepsilon,C)$ en $\mathbb{G}$ tal que $K(C/e,a)\leqslant\varepsilon \quad \forall(e,a)\notin C'$.

d) Igual a la Hipótesis 3.8 c).

Observación 3.10

a) Es obvio que las Hipótesis 3.8 y 3.9 no son comparables. Por ejemplo en la Hipótesis 3.8 está implícito que $\mathbb{E}$ y $\mathbb{G}$ son espacios de Borel[2], que es una condición más débil que la Hipótesis 3.9 a). Pero por otro lado, la Hipótesis 3.8 a) es más fuerte (i.e., implica) la Hipótesis 3.9 b). Similarmente, 3.9 c) implica 3.8 b), pero 3.8 d) no es requerido en 3.9.

b) En la mayoría de las aplicaciones de los procesos de decisión de Markov, los espacios $\mathbb{E}$ y $\mathbb{G}$ y también el espacio de controles $\mathbb{A}$, son subconjuntos de espacios Euclideanos, así que la Hipótesis 3.9 a) realmente no es muy restrictiva. Una condición suficiente para la Hipótesis 3.9 a) es que $\mathbb{E}$ y $\mathbb{A}$ sean ambos espacios métricos localmente compactos y separables (Que es una condición suficiente para que $\mathbb{E}\times\mathbb{A}$ sea un espacio métrico separable y localmente compacto; véase, por ejemplo, [15] página 75), y que $\mathbb{G}$ sea abierto o cerrado en $\mathbb{E}\times\mathbb{A}$ ([15], página 66). La razón principal para pedir que $\mathbb{E}$ y $\mathbb{G}$ sean como en la Hipótesis 3.9 a) es explicada en la Observación 3.12.

c) La Hipótesis 3.9 c) sobre K, así como (3.21) abajo (véase también la Observación 3.14), está relacionada con una condición dada por Benes [7] para asegurar la existencia de distribuciones invariantes para cadenas de Markov con la propiedad débil de Feller. (Observe que la continuidad de K

[2] Véase la definición en el párrafo de terminología y notación del capítulo 1.

es una propiedad como la de Feller.) En nuestro contexto, la Hipótesis 3.9 c) implica en particular que, como se puede demostrar fácilmente, la función

$$u \to \int u(y)\, K(dy/\cdot) \quad \text{está en } \; C_0(\mathbb{G}) \; \text{ si } u \text{ está en } C_0(\mathbb{E})\,. \tag{3.10}$$

Por ejemplo, considérese un sistema de control Markoviano en $\mathbb{R}^d$, $e_{t+1} = F(e_t, a_t, w_t)$, con $t = 0, 1, 2, ...$ y donde las w_t son variables aleatorias independientes e idénticamente distribuídas e independientes del estado inicial e_0. Entonces la Hipótesis 3.9 c) se cumple, sí, para toda s, la función $F(e, a, s)$ es continua en (e, a) y más aún, $F(e, a, s) \to \infty$ cuando $(e, a) \to \infty$. Un caso particular es el sistema Markoviano discreto de ruido aditivo en el inciso d) abajo que incluye muchos modelos encontrados en las aplicaciones ([2], [14], [18] y [26]).

d) Como un ejemplo, consideremos el sistema con ruido aditivo $e_{t+1} = G(e_t, a_t) + w_t$ con espacios de estados y de controles iguales a $\mathbb{R}$. Además, consideremos que el costo en una etapa c es una función cuadrática, por decir $c(e, a) = \alpha e^2 + \beta a^2$ con α y β positivos y $G(e, a)$ una función continua dada. En este caso las hipótesis 3.8 a), b) y 3.9 a), b) son obviamente satisfechas, y condiciones suficientes para las Hipótesis 3.8 c) y d) son encontradas fácilmente (por ejemplo[24]). Finalmente la segunda parte de la Hipótesis 3.9 c) se cumple también, si, por ejemplo $|G(e, a)| \to \infty$ cuando $|(e, a)| \to \infty$. En otras palabras, bajo condiciones razonables, las Hipótesis 3.8 y 3.9 se satisfacen ambas. Para el caso vectorial se cumple una afirmación análoga.

e) Sean $(\mathbb{X}, \mathbb{Y})$ y $(\mathbb{Z}, \mathbb{W})$ dos pares duales de espacios vectoriales, y $G : \mathbb{X} \to \mathbb{Z}$ un operador lineal con operador adjunto G^*. Entonces el operador G es débilmente continuo si y sólo si G^* mapea $\mathbb{W}$ en $\mathbb{Y}$ ([1] página 37 o [13] página 984). Por ejemplo la Hipótesis 3.8 c)[3], da como resultado que para toda función g en $F(\mathbb{E})$

$$\begin{aligned} \left| \int g(x)\, K(dx/e, a) \right| &\leqslant \|g\|_{w_0} \int w_0(x)\, K(dx/e, a) \leqslant \|g\|_{w_0}\, k w(e, a) \\ &= k' w(e, a) \end{aligned}$$

para alguna constante k'. De aquí que L_1^* mapea $F(\mathbb{E})$ en $F(\mathbb{G})$, y por lo tanto L_1 es débilmente continuo. De lo anterior se sigue que L es débilmente

[3]Que es igual a la Hipótesis 3.9 d).

continuo y así obtenemos el inciso a) del siguiente lema; el inciso b) se sigue de la continuidad débil de K.

Lema 3.11

Bajo las Hipótesis 3.8 o 3.9:
a) los operadores L_1 y L son débilmente continuos y
b) L_1^* mapea $C_a(\mathbb{E})$ en $C_a(\mathbb{G})$.

3.5 CONSISTENCIA

En esta sección estudiamos la consistencia de los programas lineales P y P^*. Comenzamos observando que el programa dual P^* es consistente y, en el Teorema 3.13, usando las Hipótesis 3.9 y el Teorema Generalizado de Farkas (Teorema 3.4), damos condiciones necesarias y suficientes para que el programa lineal P sea consistente. Además, la Observación 3.12 contiene algunos detalles técnicos que serán usados en las demostraciones de los Teoremas 3.13 y 3.15.

El programa dual P^* es trivialmente consistente. Por ejemplo, para cualquier constante k la pareja $(\rho, h) := (0, k)$ satisface la restricción (3.11). En cambio para obtener una condición necesaria y suficiente para que el programa lineal P sea consistente, a partir del Teorema Generalizado de Farkas (Teorema 3.4), necesitamos mostrar que el conjunto $L(M_+(\mathbb{G})) \subseteq \mathbb{R} \times M(\mathbb{E})$ es débilmente cerrado. Como no hemos podido demostrar esto directamente, "pertubaremos" el programa lineal P de la siguiente manera:

Sea v_0 una función estrictamente positiva en $C_0(\mathbb{E})$ y considérese el operador lineal $T : M(\mathbb{G}) \times \mathbb{R}^2 \to M(\mathbb{E}) \times \mathbb{R}^2$ definido por

$$T(\mu, r_1, r_2) := (L_1\mu, \langle \mu, 1\rangle + r_1, \langle \mu, v_0\rangle - r_2). \tag{3.11}$$

El operador adjunto $T^* : F(\mathbb{E}) \times \mathbb{R}^2 \to F(\mathbb{G}) \times \mathbb{R}^2$ es

$$T^*(h, \rho_1, \rho_2) := (L_1^* h + \rho_1 + \rho_2 v_0, \rho_1, -\rho_2). \tag{3.12}$$

Ahora nótese que P es consistente[4] si y sólo si la ecuación lineal

$$T(\mu, r_1, r_2) := (0, 1, \varepsilon) \tag{3.13}$$

[4] Es decir, existe una medida $\mu \in M_+(\mathbb{G})$ que satisface $L\mu = (1, 0)$; véase (3.8).

tiene una solución (μ, r_1, r_2) en $M_+(\mathbb{G}) \times \mathbb{R}_+^2$ para algún $\varepsilon > 0$. En efecto, es obvio que si μ satisface (3.8), entonces $(\mu, 0, 0)$ satisface (3.13) con $\varepsilon := \langle \mu, v_0 \rangle$; recíprocamente, si (μ, r_1, r_2) satisface (3.13), entonces $\langle \mu, v_0 \rangle \geqslant \varepsilon$ implica que $\langle \mu, 1 \rangle > 0$ y, por lo tanto, $\mu' = \frac{\mu}{\langle \mu, 1 \rangle}$ satisface (3.8). Usaremos (3.13) y el Teorema Generalizado de Farkas (Teorema 3.4) en la demostración del Teorema 3.13 a).

Observación 3.12

a) Consideremos a $M(\mathbb{G})$ (Definición 3.6) como el espacio de Banach de las medidas finitas con signo en $\mathbb{G}$ con la norma de variación total $\|\cdot\|_{VT}$. Observe que por las definiciones de las funciones de ponderación w y w_0 ((3.4), (3.5) y (3.6)) tenemos que

$$\|\mu\|_w = \int w(x) \, |\mu| \, (dx) \geqslant \|\mu\|_{VT}$$

así que $(M(\mathbb{G}), \|\cdot\|_w)$ es un subespacio de $(M(\mathbb{G}), \|\cdot\|_{VT})$. Por otro lado con $\mathbb{G}$ como en la Hipótesis 3.9 a),el espacio $(M(\mathbb{G}), \|\cdot\|_{VT})$ es el dual del espacio separable de Banach $C_0(\mathbb{G})$.

b) Consideraremos tres topologías en $M(\mathbb{G})$: la topología débil $\sigma(M(\mathbb{G}), F(\mathbb{G}))$, la topología débil $\sigma(M(\mathbb{G}), C_a(\mathbb{G}))$ y la topología débil* $\sigma(M(\mathbb{G}), C_0(\mathbb{G}))$. Sin embargo, como lo establecimos en el párrafo que sigue a la Definición 3.6, "topología débil" significará siempre $\sigma(M(\mathbb{G}), F(\mathbb{G}))$ a menos que establezcamos explícitamente otro significado.

c) Bajo la Hipótesis 3.9 a), si $\{\mu^j\}$ es una sucesión acotada de medidas en $\mathbb{G}$ que converge en la topología débil* $\sigma(M(\mathbb{G}), C_0(\mathbb{G}))$, entonces la sucesión de medidas marginales $\hat{\mu}^j$ en $\mathbb{E}$ también converge en la topología débil*. Esto es, si

$$\langle \mu^j, v \rangle \underset{j \to \infty}{\longrightarrow} \langle \mu, v \rangle \qquad \forall v \in C_0(\mathbb{G}) \tag{3.14}$$

entonces

$$\langle \hat{\mu}^j, u \rangle \underset{j \to \infty}{\longrightarrow} \langle \hat{\mu}, u \rangle \qquad \forall u \in C_0(\mathbb{E}) \tag{3.15}$$

Para probar (3.15), primero nótese que, bajo la Hipótesis 3.9 a), $\mathbb{G}$ es un conjunto σ-compacto y, por lo tanto, existe una sucesión creciente de conjuntos compactos $\{C_n\}$ tal que $C_n \uparrow \mathbb{G}$. Más aún, por el Lema de

Urysohn, para cada $\varepsilon > 0$ y para todo $n = 1, 2, 3, ...$, existe una función α_n en $C_0(\mathbb{G})$ tal que $0 \leqslant \alpha_n \leqslant 1$, con $\alpha_n(e,a) = 1$ si $(e,a) \in C_n$ y $\alpha_n(e,a) = 0$ si la distancia de (e,a) a C_n es mayor o igual que ε. Ahora dada $u \in C_0(\mathbb{E})$, defina $u_n(e,a) := u(e)\,\alpha_n(e,a)$ en $\mathbb{G}$, entonces u_n está en $C_0(\mathbb{G})$ para todo n y, además, para todo (e,a) en $\mathbb{G}$

$$|u_n(e,a)| \leqslant |u(e)| \leqslant \|u\| < \infty \quad \text{y} \quad u_n(e,a) \underset{n\to\infty}{\longrightarrow} u(e).$$

De aquí que, por el Teorema de Convergencia Dominada, para cada j

$$\langle \mu^j, u_n \rangle \underset{n\to\infty}{\longrightarrow} \langle \mu^j, u \rangle \text{ y } \langle \mu, u_n \rangle \underset{n\to\infty}{\longrightarrow} \langle \mu, u \rangle \tag{3.16}$$

Por otro lado, para cada n, (3.14) conduce a

$$\langle \mu^j, u_n \rangle \underset{j\to\infty}{\longrightarrow} \langle \mu, u_n \rangle \tag{3.17}$$

Finalmente la conclusión (3.15) se sigue de (3.16) y (3.17) y de la desigualdad

$$\begin{aligned}
\left|\langle \hat{\mu}^j, u \rangle - \langle \hat{\mu}, u \rangle\right| &\leqslant \left|\langle \hat{\mu}^j, u \rangle - \langle \mu^j, u_n \rangle\right| \\
&\quad + \left|\langle \mu^j, u_n \rangle - \langle \mu, u_n \rangle\right| + \left|\langle \mu, u_n \rangle - \langle \hat{\mu}, u \rangle\right| \\
&= \left|\langle \mu^j, u \rangle - \langle \mu^j, u_n \rangle\right| \\
&\quad + \left|\langle \mu^j, u_n \rangle - \langle \mu, u_n \rangle\right| + \left|\langle \mu, u_n \rangle - \langle \mu, u \rangle\right|.
\end{aligned}$$

En esta última igualdad hemos usado la Observación 3.7 c) viendo a u como una función en $\mathbb{G}$, constante en las acciones para cada $e \in \mathbb{E}$.

Teorema 3.13

Suponga que se satisfacen las Hipótesis 3.9. Entonces
a) P es consistente si y sólo si

$$L_1^* h + \rho_1 + \rho_2 v_0 \geqslant 0 \text{ con } h \in F(\mathbb{E}),\ \rho_1 \geqslant 0 \text{ y } \rho_2 \leqslant 0 \tag{3.18}$$

implica

$$\rho_1 + \varepsilon\rho_2 \geqslant 0 \text{ para algún } \varepsilon > 0 \tag{3.19}$$

b) Sea v_0 una función estrictamente positiva en $C_0(\mathbb{E})$ y w_0 la función de peso (Definición 3.6), y supóngase que existe $\varepsilon > 0$ y una política estacionaria $\pi = \phi^{\infty}$, con $\phi \in \Phi$, tal que $\forall e \in \mathbb{E}$

$$\liminf_{n\to\infty} \frac{E_e^{\pi}\left[w_0\left(e_n\right)\right]}{n} = 0 \tag{3.20}$$

y

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} E_e^{\pi}\left(v_0\left(e_t\right)\right) \geqslant \varepsilon. \tag{3.21}$$

Entonces P es consistente.

Demostración:

a) En esta demostración usamos el Teorema Generalizado de Farkas (Teorema 3.4) con las siguientes identificaciones [véase (3.11), (3.12) y (3.13)]:

$$\begin{aligned} (\mathbb{X},\mathbb{Y}) &: = \left(M(\mathbb{G})\times\mathbb{R}^2, F(\mathbb{G})\times\mathbb{R}^2\right) \quad (\mathbb{Z},\mathbb{W}) := \left(M(\mathbb{E})\times\mathbb{R}^2, F(\mathbb{E})\times\mathbb{R}^2\right), \\ G &: = T, \qquad \mathbb{C} := M_+(\mathbb{G})\times\mathbb{R}^2_+, \qquad b := (0,1,\varepsilon) \end{aligned}$$

Entonces la condición a) en el Teorema 3.4 es la misma que (3.13), y la condición b) del mismo teorema se traduce en

$$T^*\left(h,\rho_1,\rho_2\right) \geqslant 0 \Rightarrow \left\langle (0,1,\varepsilon), (h,\rho_1,\rho_2)\right\rangle \geqslant 0$$

que es precisamente la condición "(3.18) implica (3.19)". De aquí se sigue que para probar el Teorema 3.13 a) sólo necesitamos verificar las hipótesis del Teorema de Farkas (Teorema 3.4), a saber, i) T es débilmente continua y ii) $T\left(M_+(\mathbb{G})\times\mathbb{R}^2_+\right)$ es débilmente cerrado. La parte i) se sigue del Lema 3.11 a) y de la definición de T dada en (3.11). Por lo tanto, para completar la demostración del Teorema 3.13 a) sólo nos resta demostrar ii).

Demostración de ii): Para probar que $T\left(M_+(\mathbb{G})\times\mathbb{R}^2_+\right)$ es débilmente cerrado, sea $(D,\leqslant)$ un conjunto dirigido y sea $\left\{\left(\mu^{\alpha},\sigma_1^{\alpha},\sigma_2^{\alpha}\right) : \alpha\in D\right\}$ una red en el cono $M_+(\mathbb{G})\times\mathbb{R}^2_+$, tal que $T\left(\mu^{\alpha},\sigma_1^{\alpha},\sigma_2^{\alpha}\right)$ converge débilmente a $(\nu,\rho_1,\rho_2)\in M(\mathbb{E})\times\mathbb{R}^2$, i.e,[5]

$$\left\langle L_1\mu^{\alpha},u\right\rangle \to \langle \nu,u\rangle \qquad \forall u\in\mathbf{F}(\mathbb{E}) \tag{3.22}$$

[5]Por la definición (3.11) de T.

$$\langle \mu^\alpha, 1\rangle + \sigma_1^\alpha \to \rho_1, \text{ y} \tag{3.23}$$

$$\langle \mu^\alpha, v_0\rangle - \sigma_2^\alpha \to \rho_2. \tag{3.24}$$

Deseamos mostrar que (ν, ρ_1, ρ_2) está en la imagen del cono $M_+(\mathbb{G}) \times \mathbb{R}^2_+$, i.e., existe $(\mu^0, \sigma_1^0, \sigma_2^0) \in M_+(\mathbb{E}) \times \mathbb{R}^2_+$ tal que

$$a)\ L_1\mu^0 = \nu, \tag{3.25}$$
$$b)\ \langle \mu^0, 1\rangle + \sigma_1^0 = \rho_1, \text{ y} \tag{3.26}$$
$$c)\ \langle \mu^0, v_0\rangle - \sigma_2^0 = \rho_2 \tag{3.27}$$

Si $\rho_1 = 0$ en (3.23), entonces $\sigma_1^\alpha \to 0$ y $\langle \mu^\alpha, 1\rangle \to 0$ y es fácilmente verificable que necesariamente $\nu = 0$ y que (3.25) se satisface con $(\mu^0, \sigma_1^0, \sigma_2^0) = (0, 0, -\rho_2)$. Consideremos entonces el caso $\rho_1 > 0$.

Supóngase que $\rho_1 > 0$. Entonces, por (3.23), existe $\alpha_0 \in D$ tal que

$$0 \leqslant \langle \mu^\alpha, 1\rangle = \mu^\alpha(\mathbb{G}) \leqslant 2\rho_1 \qquad \forall \alpha \geqslant \alpha_0$$

De aquí que, por el Teorema de Banach-Alaoglu-Bourbaki (Teorema 3.5) y la Observación 3.12 a) existe una medida μ^0 en $\mathbb{G}$ y una sucesión $\{j\}$ en D tal que $\mu^j \to \mu^0$ en la topología débil* $\sigma(M(\mathbb{G}), C_0(\mathbb{G}))$, i.e.,

$$\langle \mu^j, v\rangle \to \langle \mu^0, v\rangle \qquad \forall v \in C_0(\mathbb{G}) \tag{3.28}$$

Más aún, μ^0 está en $M(\mathbb{G})$, ya que $0 \leqslant \mu^0(\mathbb{G}) \leqslant \liminf_{j\to\infty} \mu^j(\mathbb{G}) \leqslant 2\rho_1$ (por ejemplo, [12] Proposición III.12). Ahora, sea u una función arbitraria en $C_0(\mathbb{E})$. Entonces (3.28) y (3.10) producen

$$\left\langle \mu^j, \int_{\mathbb{E}} u(x)\, K(dx/\cdot)\right\rangle \to \left\langle \mu^0, \int_{\mathbb{E}} u(x)\, K(dx/\cdot)\right\rangle,$$

mientras que (3.28) y la Observación 3.12 c) nos llevan a

$$\langle \hat{\mu}^j, u\rangle \to \langle \hat{\mu}^0, u\rangle .$$

Por lo tanto, por (3.7) $L_1\mu^j$ converge a $L_1\mu^0$ en la topología débil*, i.e.,

$$\langle L_1\mu^j, u\rangle \to \langle L_1\mu^0, u\rangle \qquad \forall u \in C_0(\mathbb{E}).$$

De esta expresión junto con (3.22) obtenemos $L_1\mu^0 = \nu$ que es la condición en (3.25) a). Finalmente (3.25) b) y c) se satisfacen definiendo $\sigma_1^0 := \rho_1 - \langle \mu^0, 1\rangle$ y $\sigma_2^0 := \langle \mu^0, v_0\rangle - \rho_2$. Esto demuestra ii), que, como fue ya señalado, concluye la demostración del inciso a) del Teorema 3.13.

b) Por el inciso a), para demostrar b) es suficiente mostrar que (3.18), junto con (3.20) y (3.21), implican (3.19). Así que tomemos $\pi = \phi^\infty$ como en (3.20) y reescribamos (3.18) como

$$h(e) \geqslant \int h(x)\, K(dx/e, a) - \rho_1 - \rho_2 v_0(e), \quad h \in F(\mathbb{E}), \quad \rho_1 \geqslant 0, \quad \rho_2 \leqslant 0. \tag{3.29}$$

Note que (3.20) implica obviamente que

$$\liminf_{n\to\infty} \frac{E_e^\pi [h(e_n)]}{n} = 0 \qquad \forall h \in F(\mathbb{E}), \quad e \in \mathbb{E}. \tag{3.30}$$

Por otro lado, ya que (3.29) se satisface para todo (e, a) en $\mathbb{G}$, tomando en cuenta la Observación 1.3, la integración con respecto a $\varphi(\cdot/e)$ conduce a

$$h(e) \geqslant \int h(x)\, K(dx/e, \varphi) - \rho_1 - \rho_2 v_0(e) \qquad \forall e \in \mathbb{E}$$

que a su vez, por iteración, conduce a

$$h(e) \geqslant E_e^\pi [h(e_n)] - n\rho_1 - \rho_2 \sum_{t=0}^{n-1} E_e^\pi v_0(e_t),$$

i.e.,

$$h(e) + n\rho_1 + \rho_2 \sum_{t=0}^{n-1} E_e^\pi v_0(e_t) \geqslant E_e^\pi [h(e_n)].$$

Así multiplicando por $\frac{1}{n}$ y tomando lim inf en esta última expresión, (3.30) y (3.21) producen (3.19). Por lo tanto, por el inciso a), P es consistente.▲

Observación 3.14

Las hipótesis del Teorema 3.13 b) implican que el kérnel estocástico $K(\cdot/\cdot, \varphi)$ satisface la condición (v) de Benes en el artículo [7], para la existencia de una medida invariante no trivial. Efectivamente, supóngase que

$\varepsilon > 0$ es tal que $\varepsilon \leqslant \|v_0\|$, y escojamos $0 < \varepsilon_0 < \varepsilon$. Como $v_0 \in C_0(\mathbb{E})$ existe un conjunto compacto C_0 tal que $0 < v_0(e) < \varepsilon_0$ para todo $e \in C_0^c$. Entonces, escribiendo a $\mathbb{E}$ como la unión de C_0 y C_0^c, para todo $t = 0, 1, 2, \ldots$ y todo $e \in \mathbb{E}$, tenemos

$$E_e^\pi\left[v_0(e_t)\right] = \int_{\mathbb{E}} v_0(x)\, K^t(dx/e, \varphi) \leqslant (\|v_0\| - \varepsilon_0)\, K^t(C_0/e, \varphi) + \varepsilon_0$$

Esto implica, por (3.21),

$$\liminf_n \frac{1}{n} \sum_{t=0}^{n-1} K^t(C_0/e, \varphi) \geq \frac{\varepsilon - \varepsilon_0}{\|v_0\| - \varepsilon_0} > 0$$

Finalmente, tome la medida en la condición (v) de Benes [7] como, por ejemplo, la distribución inicial ν en (3.1) para obtener la condición referida.

3.6 EXISTENCIA DE SOLUCIONES ÓPTIMAS

En esta sección discutiremos sobre la existencia de soluciones óptimas para los programas lineales P y P^*, y para las ecuaciones de optimalidad (3.2) y (3.3). Demostraremos en el Teorema 3.15 que: a) El programa lineal P es resoluble y no hay abertura dual, i.e., $\sup P^* = \inf P$. b) Las sucesiones minimizantes para P son débilmente precompactas y c) Si existe una sucesión maximizante para P^*, acotada en la norma con peso w_0, entonces el programa dual P^* es resoluble y se satisface la dualidad fuerte, i.e., $\max P^* = \min P$, más aún, existen una función $h \in F(\mathbb{E})$ y un selector medible $f \in \mathfrak{F}$ que satisfacen las ecuaciones de optimalidad (3.2) y (3.3).

En el Teorema 3.15 usamos la siguiente definición. Una sucesión de medidas $\{\mu_n\}$ en $M(\mathbb{G})$ se dice que es una sucesión minimizante para P si cada μ_n es factible para P y $\langle \mu_n, c\rangle \downarrow \inf P$. Análogamente, (ρ_n, h_n) es una sucesión maximizante para P^*, si cada (ρ_n, h_n) es factible para P^* y $\rho_n \uparrow \sup(P^*)$.

En la demostración del Teorema 3.15 usaremos el Lema 2.4 a) y c), y también la siguiente observación: si μ es factible para P y (ρ, h) es factible para P^* entonces

$$\langle L_1\mu, h\rangle = 0 \text{ y } \langle L\mu, (\rho, h)\rangle = \rho \tag{3.31}$$

Teorema 3.15

Supóngase que P es consistente con un valor finito y que se satisfacen:
i) las Hipótesis 3.8, ó
ii) 3.9 y 3.8 a).
Entonces:

a) P es resoluble y no hay abertura dual, i.e., existe una solución factible μ^* para P tal que[6]

$$\langle \mu^*, c\rangle = \min P = \sup P^* = \rho^* \tag{3.32}$$

b) Si $\{\mu^n\}$ es una sucesión minimizante para P, entonces existe una subsucesión $\{\mu^j\}$ de $\{\mu^n\}$ tal que μ^j converge en la topología débil $\sigma(M(\mathbb{G}), C_a(\mathbb{G}))$ (Observación 3.12 b) a una solución óptima para P.

c) Si (ρ_n, h_n) es una sucesión maximizante para P^* con $\{h_n\}$ acotada en la norma con peso w_0 [7], entonces P^* es resoluble, se satisface la dualidad fuerte. Además, si $\hat{\mu}^*$ es óptima para P y $\hat{\mu}^*$ es la marginal en $\mathbb{E}$, entonces la ecuación de optimalidad (3.2) se satisface $\hat{\mu}$-casi donde quiera, de hecho, existe una función $h \in F(\mathbb{E})$ y una política determinista f^∞ con f selector medible tal que para $\hat{\mu}$-casi todo $e \in \mathbb{E}$

$$\begin{aligned} \rho^* + h(e) &= \min_{a\in A(e)} \left\{ c(e,a) + \int_{\mathbb{E}} h(y)\, Q(dy/e,a) \right\} \\ &= c(e,f) + \int_{\mathbb{E}} h(y)\, Q(dy/e,f) \end{aligned} \tag{3.33}$$

Demostración:

a) Esta parte está demostrada en [24] bajo las Hipótesis i), y de hecho la demostración funciona exactamente en la misma forma bajo las Hipótesis ii).

b) Sea $\{\mu_n\}$ una sucesión miminizante para P, esto es, μ^n es factible para P, que, por (3.8), significa que

$$i) \quad \langle \mu^n, 1\rangle = 1, \qquad \text{y } ii) \quad L_1\mu^n = 0 \qquad \forall n \tag{3.34}$$

$$\text{y } \langle \mu^n, c\rangle \downarrow \inf P. \tag{3.35}$$

[6] Véase (3.1) y el Teorema 1.14.

[7] Es decir, existe una constante k tal que $\|h_n\|_{w_0} \leqslant k\ \forall n$.

Por el inciso a) el valor $\inf P$ se alcanza para alguna medida μ^* factible para P, y así podemos escribir $\min P$ en su lugar. En particular, (3.35) implica que para cualquier $\varepsilon > 0$ existe $n(\varepsilon)$ tal que

$$\min P \leqslant \langle \mu^n, c\rangle \leqslant \min P + \varepsilon \qquad \forall n \geqslant n(\varepsilon).$$

El lado derecho de la desigualdad y la Hipótesis 3.8 a) implican la existencia de una medida de probabilidad μ^* en $\mathbb{G}$ y una subsucesión $\{\mu^j\}$ de $\{\mu^n\}$ tal que μ^j converge a μ^* en la topología $\sigma(M(\mathbb{G}), C_a(\mathbb{G}))$ (por ejemplo [5],[6], [7], [24] y [26] y los Teoremas 2.3 y 2.8), i.e.,

$$\langle \mu^j, v\rangle \to \langle \mu^*, v\rangle \qquad \forall v \in C_a(\mathbb{G}). \tag{3.36}$$

De aquí que, como c es semicontinua inferiormente,

$$\langle \mu^*, c\rangle \leqslant \liminf_j \langle \mu^j, c\rangle \leqslant \min P + \varepsilon.$$

Como $\varepsilon > 0$ es arbitrario, la última desigualdad muestra que μ^* satisface (3.32). Así, para completar la demostración de la parte b) sólo resta mostrar que μ^* es factible para P [véase (3.8)]. Sin embargo, esto se sigue directamente de (3.36) y (3.34). En efecto, la Hipótesis 3.8 b), que es una parte de la Hipótesis 3.9 c), implica que L_1^* mapea $C_a(\mathbb{E})$ en $C_a(\mathbb{G})$ [Lema 3.11 b)] y por lo tanto $\forall u \in C_a(\mathbb{E})$

$$\langle L_1\mu^*, u\rangle = \langle \mu^*, L_1^* u\rangle = \lim_j \langle \mu^j, L_1^* u\rangle = \lim_j \langle L_1\mu^j, u\rangle = 0,$$

esto es, $L_1\mu^* = 0$

c) Sea $\{(\rho_n, h_n)\}$ una sucesión maximizante para P^*, i.e., (ρ_n, h_n) es factible para P^*, así que, por (3.9), para todo $n = 1, 2, 3, ...$ y todo $(e, a) \in \mathbb{G}$

$$\rho_n + h_n(e) \leqslant c(e,a) + \int_{\mathbb{E}} h_n(x)\, K(dx/e,a) \tag{3.37}$$

y

$$\rho_n = \langle (1,0), (\rho_n, h_n)\rangle \uparrow \sup P^* = \rho^*.$$

Defina $h(e) := \limsup_n h_n(e)$, $e \in \mathbb{E}$. Dado que por hipótesis $\|h_n\|_{w_0}$ es acotada, h está en $F(\mathbb{E})$ y, más aún, aplicando el Lema de Fatou en (3.37) vemos que

$$\rho^* + h(e) \leqslant c(e,a) + \int_{\mathbb{E}} h(x)\, K(dx/e,a) \qquad \forall (e,a) \in \mathbb{G}; \tag{3.38}$$

esto es, (ρ^*, h) es factible para P^*, por lo tanto óptima, puesto que $\rho^* = \sup P^*$. Este hecho, combinado con (3.32) produce la dualidad fuerte: existe una solución óptima μ^* para P y una solución óptima (ρ^*, h) para P^*, y $\langle \mu^*, c\rangle = \rho^*$. Por otro lado, por (3.31)

$$\langle \mu^*, L^*(\rho^*, h)\rangle = \langle L\mu^*, (\rho^*, h)\rangle = \rho^*$$

Esto nos lleva a

$$\langle \mu^*, c - L^*(\rho^*, h)\rangle = 0$$

o, equivalentemente, escribiendo $\mu^*(d(e,a)) = \varphi^*(da/e)\,\hat{\mu}^*(de)$ como en el Lema 2.4 a),

$$\begin{aligned} 0 &= \int_{\mathbf{G}} [c(e,a) - L^*(\rho^*, h)(e,a)]\,\mu^*(d(e,a)) \\ &= \int_{\mathbb{E}} \left\{ c(e,\varphi^*) - \rho^* - h(e) + \int_{\mathbb{E}} h(x)\,K(dx/e,\varphi^*) \right\} \hat{\mu}^*(de). \end{aligned}$$

Esta ecuación y (3.38) implican que para $\hat{\mu}^*-$ casi todo $e \in \mathbb{E}$ (Véanse (3.2) y (3.3))

$$\begin{aligned} \rho^* + h(e) &= \min_{a \in A(\mathbb{E})} \left\{ c(e,a) + \int_{\mathbb{E}} h(x)\,K(dx/e,a) \right\} \\ &= c(e,\varphi^*) + \int_{\mathbb{E}} h(x)\,K(dx/e,\varphi^*). \end{aligned}$$

Finalmente, por el Lema 2.4 c), existe una política estacionaria determinista f^∞ tal que

$$\rho^* + h(e) \geqslant c(e, f^\infty) + \int_{\mathbb{E}} h(x)\,K(dx/e, f^\infty)$$

para $\hat{\mu}^*-$ casi todo $e \in \mathbf{E}$, que combinado con (3.38) produce (3.33). Esto completa la demostración del teorema.▲

Observación 3.16

Bajo las Hipótesis 3.8 o 3.9, el conjunto de soluciones óptimas para P es un conjunto convexo, débilmente cerrado, y es un subconjunto extremal del conjunto de soluciones factibles, donde "extremal" significa que si μ es una medida óptima para P y es combinación convexa de dos medidas factibles, estas tienen que ser óptimas [10].

COMENTARIO FINAL AL CAPÍTULO 3

En este capítulo vimos otro enfoque al problema de control de Markov cuando el índice de funcionamiento es el costo promedio, para el caso unicadena, aplicando la programación lineal infinita. Dimos condiciones necesarias y suficientes para que el programa lineal P, asociado a dicho problema de control sea consistente, y también encontramos condiciones suficientes para que el programa lineal P sea resoluble y para que existan políticas óptimas deterministas.

CONCLUSIONES Y PROBLEMAS ABIERTOS

En la tesis se estudiaron dos tipos de índices de funcionamiento de políticas, el de costo esperado descontado y el de costo promedio. Con el índice descontado se abordaron tanto el problema de control sin restricciones como el problema restringido. Con el índice promedio solo se abordó el problema sin restricciones.

La idea que subyace en la tesis es la de ver al índice de funcionamiento como una integral lo cual permite desarrollar teorías muy generales, porque se pueden desarrollar bajo hipótesis relativamente sencillas -no muy restrictivas- que permiten obtener resultados importantes, como por ejemplo la existencia de controles óptimos. Así en el capítulo 2, si el mapeo $\mu \rightarrow \int c d\mu = I_\alpha(\pi, \nu)$ es semicontinuo inferiormente y la familia de medidas μ cumple alguna condición de "compacidad" (tensión), entonces obtenemos condiciones suficientes para que el problema de control tenga solución. En el capítulo 3, el escribir el índice de funcionamiento como una integral permite plantearlo como un problema de programación lineal y obtener condiciones necesarias y suficientes para la existencia de políticas óptimas.

Esto deja como problemas abiertos el como calcular, al menos aproximadamente, las funciones de costo óptimo (las funciones de valor) y el de como calcular los controles óptimos. También queda abierto atacar el problema de control con restricciones en espacios generales vía el enfoque de la programación lineal. En una etapa posterior trataremos de abordar alguno de estos problemas.

Bibliografía

[1] E.J. Anderson and P. Nash, *Linear Programming in Infinite-Dimensional Spaces*, Wiley, Chichester, 1987.

[2] A. Arapostathis, V.S. Borkar, E. Fernández-Gaucherand, M.K. Ghosh and S.I. Marcus, *Discrete-time controlled Markov processes with average cost criterion: a survey*, SIAM J. Control Optim., 31 (1993), pp. 282-344.

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

[4] E.J. Balder, *On compactness of the space of policies in stochastic dynamic programming*, Stoch. Proc. Appl., 32 (1989), pp. 141-150.

[5] E.J. Balder, *On equivalence of strong and weak convergence in L_1-spaces under extreme point conditions*, Israel J. Math., 75 (1991), pp. 21-47.

[6] E.J. Balder, *Lectures on Young Measures*, Cahiers de Mathématiques de la Décision, CEREMADE, Université Paris IX-Dauphine, 1995.

[7] V.E. Benes, *Finite regular invariant measures for Feller processes*, J. Appl. Probab.,5 (1967), pp. 203-209.

[8] A.G. Bhatt and V.S. Borkar, *Occupation measures for controlled markov processes: characterization and optimality.* Ann. Probab., 24 (1996), pp. 1531-1562.

[9] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

[10] J.M. Borwein, *On the existence of Pareto efficient points*, Math. Oper. Res., 8 (1983), pp. 64-73.

[11] N. Bourbaki, *Intégration*, Cap. IX, Hermann, Paris, 1969.

[12] H. Brezis, *Analyse Fonctionnelle: Théorie et Applications*, 4^e tirage, Masson, Paris, 1993.

[13] B.D. Craven and J.J. Koliha, *Generalizations of Farkas' theorem*, SIAM J. Math. Anal. Appl, 8 (1997), pp. 983-997.

[14] G. Deodhare and M. Vidyasagar, *Infinite control synthesis and infinite linear programming*, in Mathematical Control Theory, M.C. Joshi and A.V. Balakrishnan, eds., Lecture Notes in Pure and Appl. Math., 142 (1993), Marcel Dekker, New York, pp. 87-107.

[15] J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York, 1969.

[16] J.L. Doob, *Measure Theory.* Springer-Verlag, New York, 1994.

[17] N. Dunford and J.T. Schwartz, *Linear Operators, Part I: General Theory*, Interscience-Wiley, New York, 1957.

[18] E.B. Dynkin and A.A. Yushkevich, *Controlled Markov Processes*, Springer-Verlag, Berlín, 1979.

[19] J. González-Hernández and O. Hernández-Lerma, *Envelopes of sets of measures, tightness, and Markov control processes*, Appl. Math. Optim.

[20] J. González-Hernández and R. Pérez-Hernández, *An example of a Bounded-in-probability, but non-tight Markov chain*, Stat. Prob. Lett., submitted.

[21] D. Hernández-Hernández, O. Hernández-Lerma and M. Taksar Taksar, *The linear programming approach to deterministic optimal control problems.* Appl. Math. (Warsaw), 24 (1996), pp.17-33.

[22] O. Hernández-Lerma and D. Hernández-Hernández, *Discounted cost markov decision processes on Borel spaces: The linear programming formulation.* J. of Math. Anal. and App., 183, #2, (1994), pp. 335-351.

[23] O. Hernández-Lerma and J. González-Hernández, *Infinite linear programming and multichain Markov control processes in uncountable spaces*, SIAM J. Control Optim., 36, #1, (1998), pp. 313-335.

[24] O. Hernández-Lerma and J.B. Lasserre, *Linear programming and average optimality of Markov control processes on Borel spaces–unbounded costs*, SIAM J. Control Optim., 32 (1994), pp. 480-500.

[25] O. Hernández-Lerma and J.B. Lasserre, *Average optimality in Markov control processes via discounted cost problems and linear programming*, SIAM J. Control Optim., 34 (1996), pp.295-310.

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

[27] O. Hernández-Lerma and J.B. Lasserre, *An extension of the Vitali-Hahn-Saks theorem*. Proc. Amer. Math. Soc., 124 (1996), pp. 3673-3676; correction, ibid. 126 (1998), p. 949.

[28] K. Hinderer, *Foundations of Non-stationary Dynamic programming with Discrete Time Parameter*, Lecture Notes Oper. Res. Math. Syst. 33. Springer-Verlag, Berlin (1970).

[29] A. Hordijk and J.B. Lasserre, *Linear programming formulation of MDPs in countable state space: the multichain case*, Zeit. Oper. Res., 40 (1994), pp. 91-108.

[30] L.C.M. Kallenberg, *Linear Programming and Finite Markovian Control Problems*, Mathematical Centre Tracts vol. 148, Amsterdam, 1983.

[31] J. Kawabe, *A criterion for weak compactness of measures on product spaces with applications,* Yokohama Math. J., 42 (1994), pp. 159-169.

[32] M. Kurano and M. Kawai, *Existence of optimal stationary policies in discounted decision processes: approaches by occupation measures.* Computers Math. Appl., 27 (1994), pp. 95-101.

[33] J.B. Lasserre, *Average optimal stationary policies and linear programming in countable space MDPs*, J. Math. Anal. Appl., 183 (1994), pp. 233-249.

[34] E. Mascolo and L. Migliaccio, *Relaxation methods in control theory*. Appl. Math. Optim., 20 (1989), pp. 97-103.

[35] C.A. Rogers, *Probability measures on compact sets.* Proc. London Math. Soc., 52 (1986), pp. 328-348.

[36] H.L. Royden, *Real Analysis*, 2nd ed. Macmillan, New York, 1968.

[37] J. Rubio, *Control and Optimization: The Linear Treatment of Nonlinear Problems*, Manchester University Press, Manchester, and John Wiley, New York and London, 1986.

[38] J. Rubio, The global control of nonlinear diffusion equations, SIAM J. Control Optim.,33 (1995), pp. 308-322.

[39] M. Schäl, *On dynamic programming: compactness of the space of policies*, Stoch. Proc. Appl., 3 (1975), pp. 345-364.

[40] A.A. Yushkevich, *Compactness of a measure space in dynamic programming revisited.* Technical Rpt. #95-2, Department of Mathematics, University of North Carolina, Charlotte, NC 28223, 1995.