"El saber de mis hijos hará mi grandeza"

UNIVERSIDAD DE SONORA

División de Ciencias Exactas y Naturales

Programa de Posgrado en Matemáticas

Optimalidad en índice promedio para procesos de control markovianos y juegos semi-markovianos con probabilidades de transición débilmente continuas

T E S I S

Que para obtener el grado académico de:

Maestro en Ciencias
(Matemáticas)

Presenta:

Mauricio Castro Enríquez

Directores de Tesis: Dr. Fernando Luque Vásquez y
Dr. Oscar Vega Amaya

Hermosillo, Sonora, México,

SINODALES

Dr. Fernando Luque Vásquez
Universidad de Sonora

Dr. Oscar Vega Amaya
Universidad de Sonora

Dr. Jesús Adolfo Minjárez Sosa
Universidad de Sonora

Dr. Alejandra Fonseca Morales
Centro de Investigación y de Estudios Avanzados
del Instituto Politécnico Nacional

UNIVERSIDAD DE SONORA

"El Saber de Mis Hijos
Hará Mi Grandeza"

El saber de mis hijos
hará mi grandeza

MAURICIO CASTRO ENRIQUEZ
216230034

Acta: 49
Foja: 49
Libro: 1

ACTA DE EXAMEN DE GRADO

En la ciudad de Hermosillo, Sonora, siendo las 12:00 horas del día 14 de septiembre de 2018, se reunieron en el Auditorio del Departamento de Matemáticas de la Universidad de Sonora, los integrantes del jurado:

DR. JESÚS ADOLFO MINJÁREZ SOSA
DRA. ALEJANDRA FONSECA MORALES
DR. OSCAR VEGA AMAYA
DR. FERNANDO LUQUE VÁSQUEZ

bajo la presidencia del primero y fungiendo como secretario el último, para realizar el examen de grado del programa de Maestro en Ciencias Matemáticas, a:

MAURICIO CASTRO ENRIQUEZ

quien de acuerdo a la opción de examen de grado presentó un trabajo de investigación titulado

"Optimalidad en índice promedio para procesos de control markovianos y juegos semi-markovianos con probabilidades de transición débilmente continuas"

El jurado, después de debatir entre sí reservada y libremente, emitió el siguiente dictamen:

APROBADO POR UNANIMIDAD

y para constancia se levantó la presente acta.

DR. JESÚS ADOLFO MINJÁREZ SOSA
Presidente

DR. FERNANDO LUQUE VÁSQUEZ
Secretario

DRA. ALEJANDRA FONSECA MORALES
Sinodal externo

DR. OSCAR VEGA AMAYA
Sinodal

DR. JESÚS ADOLFO MINJÁREZ SOSA, Coordinador del Programa de Maestro en Ciencias Matemáticas de la Universidad de Sonora, hace constar que las firmas que anteceden corresponden al jurado que intervino en el examen de grado.

Hermosillo, Sonora, a 14 de septiembre de 2018

DR. JESÚS ADOLFO MINJÁREZ SOSA
Coordinador de programa

DIVISIÓN DE CIENCIAS EXACTAS Y NATURALES, COORDINACIÓN POSGRADO EN MATEMÁTICAS

Folio: 7560

Optimalidad en índice promedio para procesos de control markovianos y juegos semi-markovianos con probabilidades de transición débilmente continuas

Mauricio Castro Enríquez

Septiembre de 2018

II

Índice general

Agradecimientos

Deseo agradecer a los Doctores Fernando Luque Vásquez y Oscar Vega Amaya por haber dirigido esta tesis y sus valiosos consejos,

a los Doctores Jesús Adolfo Minjárez Sosa y Alejandra Fonseca Morales por sus importantes revisiones de la tesis,

a la Doctora Alicia López Betancourt por su apoyo,

a la Doctora Martha Dolores Guzmán Partida por su amistad,

al resto de mis amigos por su apoyo, en especial a la Maestra Mayra Rosalía Tocto Erazo por su hospitalidad al inicio de la maestría,

y a mis padres María Angélica Enríquez Sánchez y José Ángel Castro Domínguez por su apoyo incondicional.

Introducción

En esta tesis se estudian problemas de control óptimo asociados a procesos de control markovianos (PCM) y juegos semi-markovianos (JSM) de suma cero con índice promedio, kérneles de transición débilmente continuos en espacios de Borel y funciones de costo o pago no acotadas.

Un PCM es un sistema dinámico cuyo comportamiento es controlado o modificado mediante la selección de una variable o parámetro de control. A las "reglas" con las que se elige dicha variable de control se les conoce como políticas o estrategias de control. La respuesta del sistema a las políticas de control se mide por medio de un índice de funcionamiento. Así, el problema de control óptimo consiste en encontrar una política de control que optimice el índice de funcionamiento.

En un juego de suma cero el sistema es controlado por dos jugadores quienes eligen sus variables de control independiente y simultáneamente. En este caso el índice de funcionamiento mide la respuesta del sistema a las políticas o estrategias de control seleccionadas por los jugadores y representa el pago (o ganancia) que recibirá uno de ellos del otro. El problema de optimización consiste en encontrar estrategias de control para ambos jugadores que garanticen un equilibrio entre las ganancias de uno y las pérdidas del otro.

Los problemas descritos anteriormente pueden abordarse con distintos enfoques [2, 11]. En el presente trabajo usaremos un enfoque de "punto-fijo" para obtener soluciones de la ecuación de optimalidad para el problema de control óptimo, y soluciones de la ecuación de Shapley para el caso de juegos semi-markovianos de suma cero. Esto se hace bajo tres tipos de condiciones sobre los modelos: una condición de Lyapunov, condiciones de continuidad y compacidad, y una condición de crecimiento sobre la función de costo o de pago.

La condición de Lyapunov permite obtener, por medio de "argumentos de punto fijo", resultados relacionados con la ergodicidad y estabilidad de los sistemas controlados. Por otra parte, las condiciones de continuidad y compacidad, en términos generales, nos permiten encontrar minimizadores medibles para ciertos problemas de optimización "estáticos" relacionados con el problema de control y el de juego. La condición de crecimiento permite considerar funciones de costo y de pago no-acotadas.

En el primer capítulo se estudia el problema de control óptimo. El análisis desarrollado en este capítulo está basado en las referencias [22, 24]. En el segundo capítulo se extiende el enfoque de las referencias anteriores a juegos semi-markovianos de suma cero y se ilustran los resultados con un problema de control mini-max para sistemas de inventario. Se incluyen además tres apéndices que muestran algunos resultados relevantes para este trabajo acerca de funciones semicontinuas, correspondencias y cadenas de Markov.

Notación

- $\mathbb{N}$ representa el conjunto de los **números naturales** y $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$.
- $\mathbb{R}$ es el conjunto de los **números reales** y $\mathbb{R}_+ := [0, \infty)$.
- Para cada $c \in \mathbb{R}$, $c^+ := \max\{c, 0\}$.
- Sea $(X, \mathcal{T})$ un espacio topológico, entonces
  - $\mathcal{B}(X)$ es la $\sigma-$**álgebra de Borel** de X, es decir, la mínima $\sigma-$álgebra que contiene a $\mathcal{T}$. A los elementos de $\mathcal{B}(X)$ los llamaremos **conjuntos de Borel**.
  - $\mathbb{P}(X)$ es el espacio de medidas de probabilidad en $(X, \mathcal{B}(X))$
  - X es un **espacio de Borel** si es un conjunto de Borel de un espacio métrico, separable y completo.
  - $C_b(X)$ es la clase de las funciones $u : X \to \mathbb{R}$ continuas y acotadas.
  - Sea $B \subset X$. La **función indicadora** del conjunto B es la función $I_B : X \to \mathbb{R}$ definida como

$$I_B(x) := \begin{cases} 1 & si\ x \in B \\ 0 & si\ x \notin B \end{cases} .$$

  - Una medida μ sobre X se dice que es **regular** si para cada $B \in \mathcal{B}(X)$ se satisface

$$\begin{aligned} \mu(B) &= \sup\{\mu(F) : F \subset B,\ F\ es\ compacto\} \\ &= \inf\{\mu(G) : B \subset G,\ G\ es\ abierto\} \end{aligned}$$

- Sean $(X, \mathcal{F})$ y $(Y, \mathcal{G})$ espacios medibles.

- Diremos que una función $Q(\cdot|\cdot)$ es un **kérnel** sobre X dado Y si se satisface lo siguiente:
  - $Q(B|\cdot)$ es una función medible para cada $B \in \mathcal{F}$.
  - $Q(\cdot|y)$ es una medida para cada $y \in Y$. Si $Q(\cdot|y)$ es una medida de probabilidad para toda $y \in Y$ diremos que Q es un **kérnel estocástico** sobre Y dado X**.**
- Sean μ_1 y μ_2 dos medidas sobre $(X, \mathcal{F})$. Diremos que μ_2 es absolutamente continua respecto a μ_1 (denotado por $\mu_2 \ll \mu_1$) si $\mu_2(E) = 0$ siempre que $\mu_1(E) = 0$.
- $M(X)$ es la clase de las funciones (Borel-) medibles $u : X \to \mathbb{R}$.
- $M_+(X) := \{u \in M(X) : u \geq 0\}$.

Capítulo 1

Problema de control con costo promedio

1.1. Introducción

En este capítulo se estudia el problema de control óptimo markoviano (PCOM) en costo promedio (esperado) y demostraremos la existencia de políticas de control óptimas bajo condiciones apropiadas.

En un PCOM se tiene un sistema dinámico cuyo comportamiento es controlado por un agente mediante acciones sobre el sistema, las cuales se eligen por medio de políticas de control. El comportamiento del sistema es medido con un índice de funcionamiento cuyo valor depende de la política de control que elige aplicar el agente sobre el sistema. El objetivo del PCOM es encontrar políticas de control que minimicen o maximicen el valor del índice de funcionamiento. Para definir un PCOM en costo promedio se requiere un modelo de control markoviano, un conjunto de políticas de control admisibles y un índice de funcionamiento que mida el comportamiento del sistema; en este trabajo estudiaremos el PCOM con el índice en costo promedio. En las secciones 1.2, 1.3 y 1.4 se introducen estos conceptos de manera precisa.

En la sección 1.4 se demuestra la existencia de políticas de control óptimas estableciendo la existencia de soluciones de la ecuación de optimalidad en costo promedio. Para demostrar la validez de la ecuación de optimalidad, es decir, la existencia de soluciones, requeriremos que el modelo de control satisfaga algunas condiciones. En este trabajo vamos a considerar tres clases de condiciones: condiciones de Lyapunov, de continuidad y compacidad, y de crecimiento. La condición de Lyapunov garantiza que los procesos controlados con políticas estacionarias tienen propiedades de ergodicidad y estabilidad. Las condiciones de continuidad y compacidad garantizan la existencia de "minimizadores" medibles para ciertos problemas de optimización estáticos. La existencia de tales "minimizadores" medibles garantiza la existencia de políticas estacionarias óptimas. Estos condiciones y algunos resultados importantes se presentan en la sección 1.5

En la sección 1.6 demostraremos, bajo las condiciones descritas en el párrafo anterior, la existencia una solución de la ecuación de optimalidad en costo promedio usando "argumentos de punto fijo". El análisis desarrollado en esta sección se basa en las referencias [22, 24].

1.2. Modelo de control markoviano

Para plantear un problema de control óptimo es necesario especificar un modelo de control, un conjunto de políticas de control admisibles y el índice de funcionamiento que se desea optimizar.

Definición 1.2.1 *Un* ***modelo de control markoviano*** *(MCM) consta de los objetos*

$$\mathcal{M} = (X,\ A,\ \{A(x) : x \in X\},\ Q,\ C)$$

donde:

(a) X *es el* ***espacio de estados posibles*** *del sistema.*

(b) A *es el* ***espacio de acciones (controles) disponibles*** *para modificar el comportamiento del sistema.*

Supondremos que los conjuntos X y A son espacios de Borel, i.e., son subconjuntos de Borel de espacios métricos separables y completos.

(c) *Para cada $x \in X$, $A(x)$ es un subconjunto no-vacío de A y representa el conjunto de acciones o controles admisibles para el estado x. Al conjunto*

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

*lo llamaremos **conjunto de pares admisibles**. Supondremos que $\mathbb{K}$ pertenece a la σ-álgebra de Borel del producto cartesiano $X \times A$. Nótese que $A(x)$ es la $x-$sección de $\mathbb{K}$ y por lo tanto $A(x) \in \mathcal{B}(A)$ para cada $x \in X$.*

(d) *la evolución del sistema está dada por el kérnel estocástico $Q(\cdot|\cdot,\cdot)$ sobre X dado $\mathbb{K}$.*

(c) *$C : \mathbb{K} \to \mathbb{R}$ es una función medible y representa el **costo por etapa**.*

De manera intuitiva podemos pensar en el modelo $\mathcal{M}$ como la representación de un sistema que evoluciona en el tiempo de la siguiente manera: en el tiempo $n = 0$, se observa el estado del sistema $x_0 = x \in X$ y elegimos una acción $a_0 = a \in A(x)$ con un costo de operación $C(x,a)$. Luego, el sistema transitará a un nuevo estado $x_1 = y \in X$ elegido de acuerdo a la medida de probabilidad $Q(\cdot|x_0,a_0)$, es decir,

$$Q(B|x,a) = \Pr[x_1 \in B|x_0 = x, a_0 = a].$$

Una vez que ha ocurrido la transición se elige una nueva acción $a_1 = b \in A(y)$ con un costo $C(y,b)$, y el proceso se repite indefinidamente generando el proceso controlado

$$(x_0, a_0, x_1, a_1, \ldots)$$

donde x_n y a_n representan el estado del sistema y la acción elegida en el tiempo $n \in \mathbb{N}_0$, respectivamente. Una **historia admisible** del proceso controlado hasta el tiempo $n \in \mathbb{N}_0$ es un vector de la forma

$$j_n := (x_0, a_0, \ldots, x_{n-1}, a_{n-1}, x_n)$$

donde $(x_i, a_i) \in \mathbb{K}$ para $i = 0, \ldots, n-1$, y $x_n \in X$. Denotaremos por H_n al conjunto de tales historias. Observe que $H_n = \mathbb{K}^n \times X$ para $n \in \mathbb{N}$ y $H_0 = X$.

1.3. Políticas de control

Claramente la evolución del proceso controlado depende del kérnel estocástico $Q(\cdot|\cdot,\cdot)$ y de las "reglas" o "procedimientos" usados para elegir las variables de control $a_n, n \in \mathbb{N}_0$. A la sucesión de tales reglas se les conoce como políticas de control y se definen a continuación.

Definición 1.3.1 *Una* ***política de control (admisible)*** $\pi = \{\pi^n\}_{n\in\mathbb{N}_0}$ *es una sucesión de kérneles estocásticos* π^n *sobre* A *dado* H_n *tales que*

$$\pi^n(A(x_n)|j_n) = 1 \quad \forall j_n \in H_n, n \in \mathbb{N}_0.$$

Denotaremos por Π *a la clase de todas las políticas de control.*

En general, particularmente en este trabajo, son de especial interés las clases de las políticas estacionarias deterministas y la clase de las políticas estacionarias aleatorizadas.

Para definir a las políticas estacionarias deterministas requerimos el concepto de selector. Un **selector** de X a A es una función medible $f : X \to A$ tal que $f(x) \in A(x)$ para todo $x \in X$. Denotaremos por $\mathbb{F}$ a la clase de todos los selectores de X a A.

Por otra parte, para cada $x \in X$, denotaremos por $\mathbb{A}(x)$ al conjunto $\mathbb{P}(A(x))$ y por Φ al conjunto de todos los kérneles estocásticos φ sobre A dado X tales que $\varphi(\cdot|x) \in \mathbb{A}(x)$ para cada $x \in X$.

Note que cada selector $f \in \mathbb{F}$ define un elemento φ_f de Φ de la siguiente manera:

$$\varphi_f(B|x) := I_B(f(x)), \quad x \in X.$$

Entonces, la correspondencia

$$f \mapsto \varphi_f \tag{1.1}$$

permite pensar en $\mathbb{F}$ como un subconjunto de Φ.

Definición 1.3.2 *Sea* $\pi = \{\pi^n\}_{n\in\mathbb{N}_0} \in \Pi$ *una política.*

(a) *Diremos que* π *es* ***estacionaria aleatorizada*** *si existe* $\varphi \in \Phi$ *tal que*

$$\pi^n(\cdot|j_n) = \varphi(\cdot|x_n) \quad \forall j_n \in H_n, n \in \mathbb{N}_0. \tag{1.2}$$

(b) *Diremos que* π *es* ***estacionaria determinista*** *si existe* $f \in \mathbb{F}$ *tal que*

$$\pi(\{f(x_n)\}|j_n) = 1 \quad \forall j_n \in H_n, n \in \mathbb{N}_0. \tag{1.3}$$

Con el propósito de simplificar notación identificaremos a una política estacionaria aleatorizada π con el kérnel estocástico $\varphi \in \Phi$ que la define y en consecuencia a la clase formada por dichas políticas también la denotaremos como Φ. De igual manera, identificaremos a una estacionaria determinista π con el selector $f \in \mathbb{F}$ que la define y representaremos a la familia formada por tales políticas como $\mathbb{F}$. Entonces, usando la correspondencia (1.1), tenemos las siguientes contenciones:

$$\mathbb{F} \subset \Phi \subset \Pi.$$

Notación 1.3.3

(a) *Para cada función medible* $u : \mathbb{K} \to \mathbb{R}$ *y* $\varphi \in \Phi$ *defina*

$$u_\varphi(x) := \int_{A(x)} u(x,a)\varphi(da|x), \quad x \in X,$$

siempre que la integral esté bien definida. En particular, con esta notación tendremos

$$Q_\varphi(\cdot|x) = \int_{A(x)} Q(\cdot|x,a)\varphi(da|x) \quad \forall x \in X.$$

Además, si $\varphi = f \in \mathbb{F}$, *entonces*

$$u_f(x) = u(x, f(x)),$$

$$Q_f(\cdot|x) = Q(\cdot|x, f(x))$$

para todo $x \in X$.

(b) *Para cada función* $w : X \to \mathbb{R}$ *medible defina*

$$Qw(x,a) := \int_X w(y)Q(dy|x,a), \quad (x,a) \in \mathbb{K},$$

siempre que la integral esté bien definida. Combinando esta notación con la del inciso (a), para una política estacionaria aleatorizada $\varphi \in \Phi$ defina

$$Q_\varphi w(x) := \int_{A(x)} Qw(x,a)\varphi(da|x), \quad x \in X.$$

En particular, si $\varphi = f \in \mathbb{F}$ entonces

$$Q_f w(x) = Qw(x, f(x))$$

$$= \int_X w(y) Q(dy|x, f(x))$$

$$= \int_X w(y) Q_f(dy|x)$$

para todo $x \in X$.

(c) *Sea $\nu \in \mathbb{P}(X)$. Para cada función $w : X \to \mathbb{R}$ medible defina*

$$\nu(w) := \int_X w(y)\nu(dy)$$

siempre que la integral esté bien definida.

El siguiente teorema garantiza la existencia del proceso $\{(x_n, a_n)\}$ con las distribuciones marginales determinadas por la distribución del estado inicial del sistema ν, la ley de transición $Q(\cdot|\cdot,\cdot)$ y la política $\pi = \{\pi^n\}_{n\in\mathbb{N}_0}$ con la que se eligen controles. Este resultado es una consecuencia del Teorema de Ionescu-Tulcea [3, Theorem 2.7.2].

Teorema 1.3.4 *Sea $\Omega := (X \times A)^\infty$ y $\mathcal{F}$ la σ-álgebra producto correspondiente. Para cada $\nu \in \mathbb{P}(X)$ y cada política $\pi \in \Pi$ existe una medida de probabilidad P_ν^π y un proceso estocástico $\{(x_n, a_n)\}_{n\in\mathbb{N}_0}$ definidos ambos sobre el espacio medible $(\Omega, \mathcal{F})$ que satisfacen las siguientes propiedades para cada $n \in \mathbb{N}_0$:*

(a) $P_\nu^\pi[x_0 \in B] = \nu(B) \quad \forall B \in \mathcal{B}(X);$

(b) $P_\nu^\pi\,[a_n \in D|j_n] = \pi^n(D|j_n) \;\; \forall D \in \mathcal{B}(A)$;

(c) $P_\nu^\pi\,[x_{n+1} \in B|j_n, a_n] = Q(B|x_n, a_n) \;\; \forall B \in \mathcal{B}(X)$.

La esperanza con respecto a la medida de probabilidad P_ν^π será denotada por E_ν^π. La medida ν en el Teorema 1.3.4(a) es la **distribución** del estado inicial. Si ν está concentrada en un estado $x \in X$, es decir,

$$\nu(B) = I_B(x) \;\; \forall B \in \mathcal{B}(X),$$

entonces escribiremos P_x^π en lugar de P_ν^π, y nos referiremos a x como el **estado inicial**. Nos vamos a referir a x_n y a_n como las variables de estado y de control en la etapa $n \in \mathbb{N}_0$, respectivamente.

Observación 1.3.5 *Por el Teorema 1.3.4(c), para cada política $\varphi \in \Phi$, el proceso de los estados $\{x_n\}_{n\in\mathbb{N}_0}$ es una cadena de Markov con probabilidad de transición en un paso $Q_\varphi(\cdot|\cdot)$. En este caso, $Q_\varphi^n(\cdot|\cdot)$ representa el kérnel de transición en n-pasos, para $n \in \mathbb{N}$. Entonces, para cada $u : X \to \mathbb{R}$ se satisface*

$$Q_\varphi^n u(x) := \int_X u(y) Q_\varphi^n(dy|x) = E_x^\varphi u(x_n)$$

para cada $x \in X$, $n \in \mathbb{N}$ y toda función u para la cual la integral está bien definida.

1.4. Problema de control óptimo en costo promedio

El **costo promedio esperado** para una política $\pi \in \Pi$ dado el estado inicial $x_0 = x \in X$ se define como

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

A la función

$$J^*(x) := \inf_{\pi\in\Pi} J(x,\pi), \;\; x \in X,$$

se le llama **función de valor óptimo**. Diremos que una política $\pi^* \in \Pi$ es **óptima** (en costo promedio) si

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

Para espacios de estados y controles finitos se ha probado la existencia de políticas estacionarias (deterministas) óptimas [2]. Sin embargo, si el espacio de estados no es finito, no es posible en general–es decir, sin condiciones adecuadas sobre el modelo de control–garantizar la existencia de políticas óptimas como lo muestra el ejemplo dado en [21, Pag. 89]. Un enfoque para garantizar la existencia de políticas estacionarias óptimas en costo promedio consiste en probar la existencia de soluciones de la ecuación de optimalidad (en costo promedio).

Diremos que la **ecuación de optimalidad** se cumple si existe una constante ρ y una función medible $h : X \to \mathbb{R}$ tales que

$$\rho + h(x) = \inf_{a \in A(x)} [C(x,a) + Qh(x,a)] \quad \forall x \in X. \tag{1.4}$$

Si este es el caso, diremos que (ρ, h) es una solución de la ecuación de optimalidad.

Sea (ρ, h) una solución de la ecuación de optimalidad. Diremos que un selector $f \in \mathbb{F}$ es un minimizador de la ecuación de optimalidad si $f(x)$ alcanza el ínfimo en (1.4) para todo $x \in X$, es decir,

$$\begin{aligned} \rho + h(x) &= \inf_{a \in A(x)} [C(x,a) + Qh(x,a)] \\ &= C_f(x) + Q_f h(x). \end{aligned}$$

A la terna (ρ, h, f) se le llama **terna canónica.**

La relación de la ecuación de optimalidad y de sus minimizadores con el problema de control óptimo se muestra en el siguiente teorema.

Teorema 1.4.1 *Suponga que (ρ, h) es una solución de la ecuación de optimalidad. Si*

$$\lim_{n \to \infty} \frac{1}{n} E_x^{\pi} h(x_n) = 0 \quad \forall x \in X, \pi \in \Pi, \tag{1.5}$$

entonces

$$\rho \leq J^*(x) \quad \forall x \in X. \tag{1.6}$$

Si adicionalmente existe un minimizador $f \in \mathbb{F}$ de la ecuación de optimalidad, entonces f es una política óptima y ρ es el costo óptimo, es decir,

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

Demostración. Sean $\pi \in \Pi$ y $x_0 = x$ arbitrarios fijos. Puesto que (ρ, h) es una solución de la ecuación de optimalidad, se sigue que

$$\begin{aligned} h(x_k) &\leq C(x_k, a_k) - \rho + Qh(x_k, a_k) \\ &= C(x_k, a_k) - \rho + E_x^\pi[h(x_{k+1})|x_k, a_k] \end{aligned}$$

para todo $k \in \mathbb{N}_0$. Aplicando esperanza se obtiene la desigualdad

$$E_x^\pi h(x_k) \leq E_x^\pi C(x_k, a_k) - \rho + E_x^\pi h(x_{k+1}). \tag{1.7}$$

Sumando de $k = 0$ hasta $n-1$ resulta que

$$\sum_{k=0}^{n-1} E_x^\pi h(x_k) \leq \sum_{k=0}^{n-1} E_x^\pi C(x_k, a_k) - n\rho + \sum_{k=0}^{n-1} E_x^\pi h(x_{k+1}),$$

la cual se simplifica a

$$h(x) \leq E_x^\pi \sum_{k=0}^{n-1} C(x_k, a_k) - n\rho + E_x^\pi h(x_n). \tag{1.8}$$

Dividiendo por n en (1.8) y tomando límite cuando $n \to \infty$ se obtiene

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

Puesto que $\pi \in \Pi$ y $x_0 = x$ son arbitrarios, concluimos que

$$\rho \leq J^*(x) \quad \forall x \in X.$$

Por otra parte, puesto que $f \in \mathbb{F}$ es un minimizador de la ecuación de optimalidad se cumple la ecuación

$$\rho + h(x) = C_f(x) + Q_f h(x) \quad \forall x \in X.$$

Entonces, con argumentos similares a los dados previamente, se obtiene la igualdad

$$h(x) = E_x^f \sum_{k=0}^{n-1} C(x_k, a_k) - n\rho + E_x^f h(x_n) \quad \forall x \in X, n \in \mathbb{N},$$

de donde se sigue que

$$J(f, x) = \rho \quad \forall x \in X.$$

Por lo tanto,

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

■

En las siguientes secciones se probará la existencia de políticas estacionarias óptimas demostrando la existencia de una solución de la ecuación de optimalidad que satisface la condición (1.5) así como la existencia de minimizadores de la misma. Esto se hará suponiendo que el modelo de control satisface tres tipos de condiciones: un conjunto de condiciones de continuidad y compacidad; una condición de crecimiento sobre la función de costo, y una condición sobre la ley de transición del sistema conocida como condición de Lyapunov.

1.5. Condiciones de optimalidad

Condición 1.5.1 (C1) *Existe una función medible $W : X \to [1, \infty)$ y una constante positiva c tales que*

$$|C(x, a)| \leq cW(x) \quad \forall (x, a) \in \mathbb{K}.$$

Para cada función medible $u : X \to \mathbb{R}$ definimos la W-norma como

$$\|u\|_W := \sup_{x \in X} \frac{|u(x)|}{W(x)}.$$

Denotaremos por $B_W(X)$ al espacio de las funciones con W-norma finita. Puede demostrarse que $\|\cdot\|_W$ es una norma sobre el espacio vectorial $B_W(X)$ y que dicho espacio es completo con respecto a esta norma. Denotaremos por $L_W(X)$ y $C_W(X)$ a los subconjuntos de $B_W(X)$ formados por las funciones semicontinuas inferiormente y las funciones continuas, respectivamente. Estos espacios son completos con respecto a la métrica inducida por $\|\cdot\|_W$.

Condición 1.5.2 (C2) *Existe una función medible $\phi : \mathbb{K} \to \mathbb{R}_+$, una medida no trivial ν sobre X y una constante $\lambda \in (0,1)$ tal que para cada $(x,a) \in \mathbb{K}$, $B \in \mathcal{B}(X)$, $f \in \mathbb{F}$ se cumple lo siguiente:*

(a) $\nu(W) < \infty$,

(b) $Q(B|x,a) \geq \nu(B)\phi(x,a)$,

(c) $QW(x,a) \leq \lambda W(x) + \nu(W)\phi(x,a)$,

(d) $\nu(\phi_f) > 0$.

Lema 1.5.3 *Si se cumple C2, entonces*

$$1 \leq E_x^\pi W(x_n) \leq \lambda^n W(x) + \frac{\nu(W)}{(1-\lambda)\nu(X)} \tag{1.9}$$

para cada $x \in X$, $\pi \in \Pi$ y $n = 0, 1, \ldots$. Por lo tanto,

$$\lim_{n\to\infty} \frac{1}{n} E_x^\pi \left|u(x_n)\right| = 0 \quad \forall x \in X, \pi \in \Pi, u \in B_W(X). \tag{1.10}$$

Demostración. Por C2(c) y el Teorema 1.3.4 tenemos que

$$\begin{aligned} E_x^\pi \left[W(x_{n+1})|h_n, a_n\right] &= E_x^\pi \left[W(x_{n+1})|x_n, a_n\right] \\ &= \int_X W(y)Q(dy|x_n, a_n) \\ &\leq \lambda W(x_n) + \phi(x_n, a_n)\nu(W). \end{aligned}$$

Entonces

$$\begin{aligned} E_x^\pi W(x_{n+1}) &\leq \lambda E_x^\pi W(x_n) + \nu(W) E_x^\pi \phi(x_n, a_n) \\ &\leq \lambda \left[\lambda E_x^\pi W(x_{n-1}) + \nu(W) E_x^\pi \phi(x_{n-1}, a_{n-1})\right] \\ &\qquad +\nu(W) E_x^\pi \phi(x_n, a_n) \end{aligned}$$

Por otra parte, por C2(b) se cumple

$$1 = Q(X|x,a) \geq \phi(x,a)\nu(X) \quad \forall (x,a) \in \mathbb{K}.$$

Luego, por inducción se obtiene (1.9).

Para la demostración de (1.10) note que

$$\begin{aligned} \frac{1}{n}E_x^\pi \left|u(x_n)\right| &\leq ||u||_W \frac{1}{n}E_x^\pi W(x_n) \qquad (1.11)\\ &\leq \frac{||u||_W}{n}\left(\lambda^n W(x) + \frac{\nu(W)}{(1-\lambda)\nu(X)}\right), \end{aligned}$$

de donde se sigue (1.10). ■

Observación 1.5.4 *Para cada $B \in \mathcal{B}(X)$ y $(x,a) \in \mathbb{K}$, defina*

$$\hat{Q}(B|x,a) := Q(B|x,a) - \nu(B)\phi(x,a).$$

Supongamos que se cumple C2. Entonces, $\hat{Q}(\cdot|\cdot,\cdot)$ es un kérnel sobre X dado $\mathbb{K}$ que satisface la propiedad

$$|\hat{Q}u(x,a) - \hat{Q}v(x,a)| \leq \lambda \left\|u - v\right\|_W W(x) \qquad (1.12)$$

para cada $(x,a) \in \mathbb{K}$, $u, v \in B_W(X)$.

Demostración. Las propiedades de kérnel se siguen directamente de la definición de $\hat{Q}(\cdot|\cdot,\cdot)$ y de la propiedad C2(b). Para verificar que (1.12) se satisface fijemos funciones arbitrarias $u, v \in B_W(X)$. Entonces

$$\begin{aligned} |\hat{Q}u(x,a) - \hat{Q}v(x,a)| &\leq \int_X |u(y) - v(y)|\hat{Q}(dy|x,a)\\ &\leq \left\|u - v\right\|_W \hat{Q}W(x,a)\\ &\leq \lambda \left\|u - v\right\|_W W(x) \quad \forall (x,a) \in \mathbb{K}, \end{aligned}$$

donde la tercer desigualdad se sigue de la propiedad C2(c). ■

Definición 1.5.5 *Para cada $v \in B_W(W)$ y $f \in \mathbb{F}$, defina el operador*

$$\begin{aligned} L_f^v u(x) \quad &: \quad = v(x) + \int_X u(y)\hat{Q}_f(dy|x)\\ &= \quad v(x) + Q_f u(x) - \nu(u)\phi_f(x). \end{aligned}$$

Lema 1.5.6 *Suponga que C2 se satisface. Entonces, para cada $\upsilon \in B_W(X)$ y cada $f \in \mathbb{F}$, el operador L_f^υ es de contracción sobre $B_W(X)$ con módulo λ. Así, por el Teorema de punto fijo de Banach, existe una única función $h_f^\upsilon(x) \in B_W(X)$ tal que*

$$h_f^\upsilon(x) = \upsilon(x) + \int_X h_f^\upsilon(y) Q_f(dy|x) - \nu(h_f^\upsilon)\phi_f(x) \quad \forall x \in X. \tag{1.13}$$

Demostración. Sean $\upsilon \in B_W(X)$ y $f \in \mathbb{F}$ arbitrarios fijos. Observemos primero que para todo $u \in B_W(X)$ se satisface

$$\begin{aligned} |L_f^\upsilon u(x)| &\leq |\upsilon(x)| + \|u\|_W \int_X W(y)\hat{Q}_f(dy|x) \\ &\leq |\upsilon(x)| + \lambda ||u||_W W(x) \\ &\leq (\|\upsilon\|_W + \lambda \|u\|_W) W(x) \quad \forall x \in X, \end{aligned}$$

lo cual prueba que $L_f^\upsilon u \in B_W(X)$. Por otra parte, por la Observación 1.5.4, tenemos

$$\left| L_f^\upsilon u(x) - L_f^\upsilon w(x) \right| \leq \lambda ||u - w||_W W(x) \quad \forall x \in X.$$

Se sigue entonces que $||L_f^\upsilon u - L_f^\upsilon w||_W \leq \lambda ||u - w||_W$ para cada $u, w \in B_W(X)$. Por lo tanto, L_f^υ es un operador de contracción sobre $B_W(X)$. ■

Lema 1.5.7 *Suponga que C2 se satisface. Entonces, para cada $f \in \mathbb{F}$ existe una constante $k_f > 0$ tal que*

$$\lim_{n\to\infty} \frac{1}{n} E_x^f \sum_{k=0}^{n-1} \phi_f(x_k) = k_f \quad \forall x \in X. \tag{1.14}$$

Demostración. Elijamos $f \in \mathbb{F}$ arbitrario y tomemos $\upsilon \equiv 1$ en el Lema 1.5.6. Entonces existe una única función $h_f^1 \in B_W(X)$ tal que

$$h_f^1(x) = 1 + \int_X h_f^1(y) Q_f(dy|x) - \nu(h_f^1)\phi_f(x) \quad \forall x \in X.$$

Iterando esta ecuación obtenemos

$$h_f^1(x) = n + E_x^f h_f^1(x_n) - \nu(h_f^1) E_x^f \sum_{k=0}^{n-1} \phi_f(x_k) \quad \forall x \in X, n \in \mathbb{N}.$$

Entonces

$$\frac{1}{n}h_f^1(x) = 1 + \frac{1}{n}E_x^f h_f^1(x_n) - \nu(h_f^1)\left[\frac{1}{n}E_x^f \sum_{k=0}^{n-1} \phi_f(x_k)\right] \quad \forall n \in \mathbb{N}.$$

Puesto que

$$\left|\frac{1}{n}E_x^f h_f^1(x_n)\right| \leq \frac{1}{n}E_x^f \left|h_f^1(x_n)\right| \quad \forall x \in X,$$

tomando límite cuando $n \to \infty$ y usando el Lema 1.5.3 se obtiene la igualdad

$$\nu(h_f^1) \lim_{n\to\infty} \frac{1}{n}E_x^f \sum_{k=0}^{n-1} \phi_f(x_k) = 1 \tag{1.15}$$

lo cual implica que $\nu(h_f^1) > 0$; por lo tanto, (1.14) se cumple tomando $k_f = 1/\nu(h_f^1)$. ■

Teorema 1.5.8 *Suponga que se cumple C2. Entonces para cada $f \in \mathbb{F}$:*

(a) *la cadena de Markov con probabilidad de transición $Q_f(\cdot|\cdot)$ es ν-irreducible y Harris recurrente positiva; por lo tanto, tiene una única medida de probabilidad invariante $\mu_f(\cdot)$;*

(b) $\mu_f(W) < \infty$;

(c) $\mu_f(\phi_f) \geq \dfrac{1-\lambda}{\nu(W)}\mu_f(W) \geq \theta := \dfrac{1-\lambda}{\nu(W)} > 0.$

Si además se cumple C1, entonces:

(d) $\rho_f := \mu_f(C_f)$ *y* $\rho^* := \inf_{f\in\mathbb{F}} \rho_f$ *son constantes finitas;*

(e) *existe una única función $h_f^0 \in B_W(X)$ que satisface la ecuación de Poisson*

$$h_f(x) = C_f(x) - \rho_f + \int_X h_f(y)Q_f(dy|x) \quad \forall x \in X, \tag{1.16}$$

tal que $\nu(h_f) = 0$;

(f) *además*

$$J(x,f) = \lim_{n\to\infty} \frac{1}{n}E_x^f \sum_{k=0}^{n-1} C_f(x_n) = \rho_f \quad \forall x \in X.$$

Demostración. Sea $f \in \mathbb{F}$ una política estacionaria arbitraria.

(a) Por el Lema 1.5.6, para cada $u \in B_W(X)$ existe una única función $h_f^u \in B_W(X)$ que satisface la ecuación

$$h_f^u(x) = u(x) + \int_X h_f^u(y) Q_f(dy|x) - \nu(h_f^u)\phi_f(x) \quad \forall x \in X.$$

Así, iterando como en la demostración del Lema 1.5.6, tenemos

$$h_f^u(x) = E_x^f \sum_{k=0}^{n-1} u(x_k) + E_x^f h_f^u(x_n) - \nu(h_f^u) E_x^f \sum_{k=0}^{n-1} \phi_f(x_k) \quad \forall n \in \mathbb{N}.$$

Entonces, por el Lema 1.5.3 y el Lema 1.5.7, obtenemos

$$\lim_{n\to\infty} \frac{1}{n} E_x^f \sum_{k=0}^{n-1} u(x_k) = \nu(h_f^u) \lim_{n\to\infty} \frac{1}{n} E_x^f \sum_{k=0}^{n-1} \phi(x_k) = \nu(h_f^u) k_f < \infty. \tag{1.17}$$

Luego, por el Teorema 1.1 en [10] lo siguiente se satisface:

(i) La cadena de Markov correspondiente a la probabilidad de transición $Q_f(\cdot|\cdot)$ es Harris recurrente positiva; por lo tanto existe una única medida de probabilidad invariante $\mu_f(\cdot)$. Esto prueba la primera afirmación en (a).

(ii) Para cualquier función medible y acotada u sobre X

$$\lim_{n\to\infty} \frac{1}{n} E_x^f \sum_{k=0}^{n-1} u(x_k) = \mu_f(u) \quad \forall x \in X. \tag{1.18}$$

A continuación probaremos que ν es una medida de irreducibilidad para la cadena con probabilidad de transición $Q_f(\cdot|\cdot)$. Por la condición C2(b) se tiene que

$$Q_f(D|x) \geq \nu(D)\phi_f(x) \quad \forall x \in X, D \in \mathcal{B}(X).$$

Iterando la desigualdad anterior se obtiene

$$Q_f^k(D|x) \geq \nu(D) Q_f^{k-1}\phi_f(x) \quad \forall x \in X, D \in \mathcal{B}(X).$$

Luego, de la Observación 1.3.5 y sumando desde $k = 0$ hasta $n-1$ resulta la desigualdad

$$\frac{1}{n} \sum_{k=0}^{n-1} P_x^f[x_{k+1} \in D] \geq \nu(D) \frac{1}{n} E_x^f \sum_{k=0}^{n-1} \phi_f(x_k) \quad \forall x \in X, n \in \mathbb{N}.$$

Tomando límite cuando $n \to \infty$, por (1.18) y el Lema 1.5.7 resulta que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P_x^f[x_{k+1} \in D] \geq \nu(D) \lim_{n\to\infty} \frac{1}{n} E_x^f \sum_{k=0}^{n-1} \phi_f(x_k) = \nu(D) k_f.$$

Entonces, si $\nu(D) > 0$ para cada $x \in X$ existe n_0 tal que $P_x^f[x_{n_0} \in D] > 0$. Por lo tanto, ν es una medida de irreducibilidad.

(b) Tomando $u \equiv W$ en el (1.17), obtenemos una función $h_f^W \in B_W(X)$ tal que

$$\lim_{n\to\infty} \frac{1}{n} E_x^f \sum_{k=0}^{n-1} W(x_k) = \nu(h_f^W) k_f < \infty \quad \forall x \in X. \tag{1.19}$$

Ahora considere una sucesión $\{w_n\}$ de funciones medibles acotadas no-negativas que convergen crecientemente a W. Entonces

$$\frac{1}{N} E_x^f \sum_{k=0}^{N-1} w_n(x_k) \leq \frac{1}{N} E_x^f \sum_{k=0}^{N-1} W(x_k) \quad \forall N \in \mathbb{N}.$$

Entonces, (1.18) y (1.19) implican que

$$\mu_f(w_n) \leq \nu(h_f^W) k_f < \infty.$$

Finalmente, tomando límite cuando $n \to \infty$ vemos que

$$\mu_f(W) \leq \nu(h_f^W) k_f < \infty.$$

(c) Por la propiedad C2(c), para cada $f \in \mathbb{F}$ se cumple la desigualdad

$$\int_X W(y) Q_f(dy|x) \leq \lambda W(x) + \nu(W)\phi_f(x) \quad \forall x \in X.$$

Integrando ambos lados de la desigualdad respecto a $\mu_f(\cdot)$ resulta la desigualdad

$$\mu_f(W) \leq \lambda \mu_f(W) + \nu(W)\mu_f(\phi_f)$$

la cual implica que

$$\mu_f(W)(1-\lambda) \leq \nu(W)\mu_f(\phi_f)$$

Puesto que W está acotada inferiormente por 1, vemos que

$$0 < 1 - \lambda \leq (1-\lambda)\mu_f(W) \leq \nu(W)\mu_f(\phi_f), \tag{1.20}$$

de donde se sigue que

$$\mu_f(\phi_f) \geq \frac{(1-\lambda)\mu_f(W)}{\nu(W)} \geq \theta = \frac{1-\lambda}{\nu(W)} > 0.$$

(**d**) Por la condición C1 tenemos que

$$-cW \leq C_f \leq cW \quad \forall f \in \mathbb{F}.$$

Entonces, por la parte (b), concluimos que $\rho_f \in \mathbb{R}$ para cada $f \in \mathbb{F}$, y también que $\rho^* = \inf_{f\in\mathbb{F}} \rho_f \in \mathbb{R}$.

(e) Para cada $f \in \mathbb{F}$ definamos el operador $\hat{T}_f$ en $B_W(X)$ como

$$\hat{T}_f u(x) := C_f(x) - \rho_f + \int_X u(y)\hat{Q}_f(dy|x).$$

Los mismos argumentos usados en la demostración del Lema 1.5.6 prueban que $\hat{T}_f$ es un operador de contracción de $B_W(X)$ en sí mismo con módulo λ. Entonces existe una única función $h_f \in B_W(X)$ que satisface la ecuación

$$h_f(x) = C_f(x) - \rho_f + \int_X h_f(y)Q_f(dy|x) - \nu(h_f)\phi_f(x) \quad \forall x \in X.$$

Luego, integrando ambos lados de la ecuación con respecto a $\mu_f(\cdot)$ obtenemos

$$\mu_f(h_f) = \mu_f(C_f) - \rho_f + \mu_f(h_f) - \nu(h_f)\mu_f(\phi_f),$$

lo cual implica que

$$\nu(h_f) = 0,$$

puesto que $\mu_f(\phi_f) > 0$. Por lo tanto, h_f es la única función en $B_W(X)$ que satisface la condición $\nu(h) = 0$ y la ecuación

$$h(x) = C_f(x) - \rho_f + \int_X h(y)Q_f(dy|x) \quad \forall x \in X. \tag{1.21}$$

(**f**) Iterando la ecuación (1.16)–o bien la ecuación (1.21) tomando $h \equiv h_f$– se obtiene la igualdad

$$h_f(x) = E_x^f \sum_{k=0}^{n-1} C_f(x_k) - n\rho_f + E_x^f h_f(x_n) \quad \forall n \in \mathbb{N},$$

la cual implica que

$$J(x,f) = \lim_{n\to\infty} \frac{1}{n} E_x^f \sum_{k=0}^{n-1} C_f(x_n) = \rho_f \quad \forall x \in X.$$

■

Observación 1.5.9 *Para cada* $u \in B_W(X)$ *y* $\varphi \in \Phi$ *se cumple que*

$$\lim_{n\to\infty} \frac{1}{n} E_x^{\varphi} \sum_{k=0}^{\infty} u(x_n) = \mu_{\varphi}(u) \quad \forall x \in X.$$

Esta afirmación se prueba sustituyendo C_f *por la función* u *en la demostración del Teorema 1.5.8 (c) y (d).*

Observación 1.5.10 *El Teorema 1.5.8 también se cumple para las políticas estacionarias aleatorizadas* $\varphi \in \Phi$ *y su demostración es exactamente la misma. Solamente falta verificar que se satisface la condición* $\nu(\phi_{\varphi}) > 0$ *para* $\varphi \in \Phi$*. Esta condición es consecuencia directa de la proposición [8, Proposition D.8, p. 184], la cual garantiza que para cada* $\varphi \in \Phi$ *existe un selector* $f \in \mathbb{F}$ *tal que*

$$\begin{aligned} \phi_{\varphi}(x) &= \int_A \phi(x,a)\varphi(da|x) \\ &\geq \phi(x, f(x)) = \phi_f(x) \end{aligned}$$

para todo $x \in X$*.*

Además de la condiciones C2 y C3, también supondremos que el modelo de control satisface el siguiente conjunto de condiciones de continuidad y compacidad.

Condición 1.5.11 (C3) *:*

(a) C *es semicontinua inferiormente en* $\mathbb{K}$*.*

(b) *El mapeo* $x \mapsto A(x)$ *es semicontinuo superiormente y toma valores compactos no-vacíos.*

(c) *La ley de transición* Q *es débilmente continua, esto es,* Qu *es continua en* $\mathbb{K}$ *para cada* $u \in C_b(X)$*.*

(d) *Las funciones* W *y* QW *son continuas en* X *y* $\mathbb{K}$*, respectivamente.*

Como se muestra en el Teorema de Selección que se presenta en la Observación 1.5.12, la condición C3 tiene un papel doble: por una parte garantiza que ciertas funciones son semicontinuas inferiormente y, por otra, la existencia de selectores que son "minimizadores".

Observación 1.5.12 (Teorema de Selección) *Sea $v : \mathbb{K} \to \mathbb{R}$ una función acotada inferiormente y defina*

$$v^*(x) := \inf_{a \in A(x)} v(x, a), \quad x \in X, \tag{1.22}$$

(a) *Si v es semicontinua inferiormente y el mapeo $x \mapsto A(x)$ es semicontinuo superiormente, entonces la función v^* es semicontinua inferiormente y acotada inferiormente. Además, existe $g^* \in \mathbb{F}$ tal que*

$$v^*(x) = v_{g^*}(x) \quad \forall x \in X.$$

(a) *Si v es continua y el mapeo $x \mapsto A(x)$ es continuo, entonces v^* es continua y existe $g_* \in \mathbb{F}$ tal que*

$$v^*(x) = v_{g^*}(x) \quad \forall x \in X.$$

La parte (a) es una consecuencia de [5, Theorem 1.2, Lemma 2.1, Theorem 3.3] y la parte (b) se sigue de [5, Theorem 4.1].

Proposición 1.5.13 *Suponga que se cumplen C3(c) y (d).*

(a) *Si $u \in L_W(X)$, entonces Qu es semicontinua inferiormente en $\mathbb{K}$.*

(a) *Si $u \in C_W(X)$, entonces Qu es continua en $\mathbb{K}$.*

Demostración. (a) Sea $u \in L_W(X)$ arbitraria. Note que

$$-\|u\|_W W \leq u \leq \|u\|_W W,$$

lo cual implica que la función

$$\hat{u} := u + \|u\|_W W$$

es no-negativa y semicontinua inferiormente. Entonces, existe una sucesión $(u_n)_{n\in\mathbb{N}}$ de funciones en $C_b(X)$ tal que $u_n \uparrow \hat{u}$ (Teorema A.0.6). Entonces, por C3(c), $(Qu_n)_{n\in\mathbb{N}}$ es una sucesión creciente de funciones continuas y acotadas en $\mathbb{K}$. Luego, por el teorema de convergencia monótona,

$$Qu_n(x, a) \uparrow Q\hat{u}(x, a) \quad \forall (x, a) \in \mathbb{K},$$

lo cual implica, por el Teorema A.0.6, que la función

$$Q\hat{u} = Qu + \|u\|_W QW$$

es semicontinua inferiormente en $\mathbb{K}$. Puesto que QW es continua, esto último implica que Qu es semicontinua inferiormente en $\mathbb{K}$.

(b) Ahora supongamos que $u \in C_W(X)$. Por (a) sabemos que Qu es semicontinua inferiormente. A continuación probaremos que Qu es semicontinua superiormente. Primero observe que la función

$$\widetilde{u} := -u + ||u||_W W$$

es no-negativa y semicontinua inferiormente–de hecho, esta función es continua. Argumentando como en la parte (a), vemos que

$$Q\widetilde{u} = -Qu + ||u||_W QW$$

es semicontinua inferiormente $\mathbb{K}$, lo cual implica que Qu es semicontinua superiormente en $\mathbb{K}$. Por lo tanto, Qu es continua. ■

1.6. Existencia de políticas estacionarias óptimas

El Teorema 1.6.1, el cual se enuncia más abajo, es el resultado principal de este capítulo. Dicho teorema afirma que si el modelo de control satisface las condiciones C1, C2 y C3, entonces existe una terna canónica (ρ^*, h^*, f^*) que satisface la condición (1.5). Entonces, por el Teorema 1.4.1, f^* es una política óptima y ρ^* es el costo promedio óptimo.

Bajo las condiciones C1 y C2, por el Teorema 1.6.1, la constantes

$$\begin{aligned} \rho_f &= \mu_f(C_f), \quad f \in \mathbb{F}, \\ \rho^* &= \inf_{f \in \mathbb{F}} \rho_f, \end{aligned}$$

son finitas y para cada $f \in \mathbb{F}$ existe una única función $h_f \in B_W(X)$ que satisface la ecuación de Poisson

$$\rho_f + h_f = C_f + Q_f h_f$$

y la condición $\nu(h_f) = 0$.

Teorema 1.6.1 *Suponga que C1, C2, y C3 se satisfacen. Entonces:*

(a) *Existe $h^* \in L_W(X)$ y $f^* \in \mathbb{F}$ tal que (ρ^*, h^*, f^*) es una terna canónica, es decir,*

$$\begin{aligned} \rho^* + h^*(x) &= \inf_{a \in A(x)} [C(x,a) + Qh^*(x,a)] \\ &= C_{f^*}(x) + Q_{f^*}h^*(x) \quad \forall x \in X. \end{aligned}$$

(b) *La constante ρ^* es el valor óptimo y f^* es una política óptima, esto es,*

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

(c) *La función h^* satisface la condición*

$$h^*(x) = \inf_{f \in \mathbb{F}^*} h_f(x) = h_{f^*}(x) \quad \forall x \in X,$$

donde

$$\mathbb{F}^* := \{f \in \mathbb{F} : \rho^* = \rho_f\}.$$

Corolario 1.6.2 *Suponga que C1, C2, y C3 se satisfacen. Si la función de costo C y la multifunción $x \mapsto A(x)$ son continuas, entonces h^* es continua.*

La demostración del Teorema 1.6.1 requiere de algunos resultados preliminares, los cuales se presentan en los Lemas 1.6.3 y 1.6.4. En estos lemas usaremos el concepto de envolvente semicontinua inferiormente. La envolvente semicontinua inferiormente de una función v sobre un espacio métrico (Y, d) se define como

$$v^l(x) := \sup_{r>0} \inf_{y \in B_r(x)} u(y), \quad x \in Y,$$

donde $B_r(x)$ denota la bola con centro en $x \in Y$ y radio $r > 0$ (véase Apéndice A).

Lema 1.6.3 *Suponga que las condiciones C1, C2, y C3(d) se satisfacen y defina $S := \phi^l$. Entonces:*

(a) Para todo $B \in \mathcal{B}(X), (x,a) \in \mathbb{K}$ se cumplen las desigualdades

$$Q(B|x,a) \geq \nu(B)S(x,a), \tag{1.23}$$

$$QW(x,a) \leq \lambda W(x) + \nu(W)S(x,a). \tag{1.24}$$

(b) además para cada $f \in \mathbb{F}$ se cumplen las desigualdades

$$\mu_f(S_f) \geq \frac{1-\lambda}{\nu(W)}\mu_f(W) \geq \theta > 0 \tag{1.25}$$

donde $\theta = (1-\lambda)/\nu(W)$.

Demostración. (a) La primera desigualdad es inmediata puesto que $\phi \geq S$. Para probar la segunda desigualdad, note que C3(d) implica que la función $QW - \lambda W$ es continua sobre $\mathbb{K}$. Entonces, combinando C2(c), (A.2), y el Lema A.0.9, vemos que

$$\begin{aligned}
\nu(W)S(x,a) &= (\nu(W)\phi)^l(x,a) \\
&\geq (QW(x,a) - \lambda W(x))^l \\
&= QW(x,a) - \lambda W(x),
\end{aligned}$$

para todo $(x,a) \in \mathbb{K}$, lo cual prueba (1.24).

(b) Para demostrar (1.25) observemos que para cada $f \in \mathbb{F}$ se cumple la desigualdad

$$\nu(W)S_f(x) \geq Q_f W(x) - \lambda W(x) \quad \forall x \in X,$$

la cual implica que

$$\nu(W)\mu_f(S_f) \geq (1-\lambda)\mu_f(W).$$

Entonces

$$\begin{aligned}
\mu_f(S_f) &\geq \frac{1-\lambda}{\nu(W)}\mu_f(W) \\
&\geq \frac{1-\lambda}{\nu(W)} = \theta > 0,
\end{aligned}$$

puesto que $\mu_f(W) \geq 1$. ■

Defina sobre el espacio $B_W(X)$ los siguientes operadores:

$$\begin{aligned}
Pu(x,a) &:= Qu(x,a) - \nu(u)S(x,a), \\
Lu(x,a) &:= C(x,a) - \rho^* + Pu(x,a), \\
L^l u(x,a) &:= (Lu)^l(x,a),
\end{aligned}$$

donde $(x,a) \in \mathbb{K}$. Además defina el operador T en $B_W(X)$ como

$$Tu(x) := \inf_{a \in A(x)} L^l u(x,a), \quad \forall x \in X. \tag{1.26}$$

Lema 1.6.4 *Suponga que C1, C2, y C3 se satisfacen. Entonces el operador T es un operador de contracción de $L_W(X)$ en sí mismo con módulo λ.*

Demostración. Primero probaremos que $Tu \in L_W(X)$ para cada $u \in B_W(X)$. Note que $P(\cdot|\cdot,\cdot)$ es un kérnel sobre X dado $\mathbb{K}$ y que $PW \leq \lambda W$. Entonces

$$\begin{aligned}
|Lu(x,a)| &= |C(x,a) - \rho^* + Pu(x,a)| \\
&\leq |C(x,a)| + |\rho^*| + \|u\|_W \int_X W(y)P(dy|x,a) \\
&\leq cW(x) + |\rho^*| W(x) + \lambda \|u\|_W W(x) \\
&\leq \kappa W(x)
\end{aligned}$$

para todo $(x,a) \in \mathbb{K}$, donde $\kappa := c + |\rho^*| + \lambda \|u\|_W$. Entonces

$$\left| L^l u(x,a) \right| \leq \kappa W(x) \ \ \forall (x,a) \in \mathbb{K},$$

que a su vez implica que

$$||Tu||_W \leq \kappa.$$

Por otra parte, puesto que $L^l u$ es semicontinua inferiormente, la desigualdad anterior y el Teorema de Selección (Observación 1.5.12) garantizan que $Tu \in L_W(X)$ para cada $u \in B_W(X)$.

Ahora probaremos que T es un operador de contracción. Comencemos observando que

$$|Pu(x,a) - Pv(x,a)| \leq \lambda \|u - v\|_W W(x)$$

para todo $(x,a) \in \mathbb{K}, u, v \in B_W(X)$. De aquí se sigue la desigualdad

$$\begin{aligned} |Lu(x,a) - Lv(x,a)| &= |Pu(x,a) - Pv(x,a)| \\ &\leq \lambda \|u - v\|_W W(x). \end{aligned}$$

Entonces

$$\begin{aligned} |Tu(x) - Tw(x)| &= \left| \inf_{a \in A(x)} L^l u(x,a) - \inf_{a \in A(x)} L^l w(x,a) \right| \\ &\leq \sup_{a \in A(x)} \left| L^l u(x,a) - L^l w(x,a) \right| \\ &\leq \sup_{a \in A(x)} |Lu(x,a) - Lw(x,a)| \\ &\leq \lambda \|u - w\|_W W(x), \quad \forall x \in X. \end{aligned}$$

La última desigualdad implica que

$$||Tu - Tv||_W \leq \lambda \|u - w\|_W \quad \forall u, v \in B_W(X),$$

lo cual prueba que T es un operador de contracción sobre $L_W(X)$. ■

A continuación presentamos la demostración del Teorema 1.6.1.

Demostración del Teorema 1.6.1. En el lema 1.6.4 se probó que $T : L_W(X) \to L_W(X)$ es un operador de contracción. Puesto que $L_W(X)$ es completo con respecto a la métrica inducida por $\|\cdot\|_W$, el Teorema de punto-fijo de Banach garantiza la existencia de una función $h^* \in L_W(X)$ tal que

$$\begin{aligned} h^*(x) &= Th^*(x) \\ &= \inf_{a \in A(x)} (C - \rho^* + Qh^* - \nu(h^*)S)^l(x,a), \end{aligned}$$

para todo $x \in X$.

Probaremos que $\nu(h^*) = 0$. Dado que $L^l h^* \leq Lh^*$, se sigue que

$$h^*(x) \leq C(x,a) - \rho^* + Qh^*(x,a) - \nu(h^*)S(x,a) \quad \forall (x,a) \in \mathbb{K}.$$

Entonces para cada $f \in \mathbb{F}$ se tiene la desigualdad

$$h^*(x) \leq C_f(x) - \rho^* + Q_f h^*(x) - \nu(h^*)S_f(x) \; \forall x \in X.$$

Integrando respecto a la medida invariante μ_f se obtiene

$$\mu_f\left(h^*\right) \leq \mu_f\left(C_f\right) - \rho^* + \mu_f\left(h^*\right) - \nu(h^*)\mu_f\left(S_f\right),$$

y así vemos que

$$\nu(h^*)\mu_f\left(S_f\right) \leq \rho_f - \rho^* \quad \forall f \in \mathbb{F}. \tag{1.27}$$

Tomando ínfimo en ambos lados de (1.27) se obtiene

$$\inf_{f \in \mathbb{F}} \left[\nu(h^*)\mu_f\left(S_f\right)\right] \leq 0,$$

lo cual implica que $\nu(h^*) \leq 0$ puesto que $\inf\limits_{f \in \mathbb{F}} \mu_f\left(S_f\right) \geq \theta > 0$ por el Lema 1.6.3.

Demostremos ahora que $\nu(h^*) \geq 0$. Por el Lema 1.5.13 Qh^* es semicontinua inferiormente en $\mathbb{K}$; entonces

$$Lh^* = C - \rho^* + Qh^* - \nu(h^*)S$$

también es semicontinua inferiormente puesto que las funciones $-\nu(h^*)S$ y C lo son. Por lo tanto

$$L^l h^* = Lh^* = C - \rho^* + Qh^* - \nu(h^*)S.$$

Por el Teorema de selección 1.5.12 existe $f^* \in \mathbb{F}$ tal que

$$\begin{aligned} h^*(x) &= Th^*(x) = \inf_{a \in A(x)} Lh^*(x,a) \qquad (1.28) \\ &= C_{f^*}(x) - \rho^* + Q_{f^*}h^*(x) - \nu(h^*)S_{f^*}(x). \end{aligned}$$

para cada $x \in X$. Integrando (1.28) con respecto a μ_{f^*} tenemos la igualdad

$$\mu_{f^*}(h^*) = \mu_{f^*}(C_{f^*}) - \rho^* + \mu_{f^*}(h^*) - \nu(h^*)\mu_{f^*}(S_{f^*}),$$

por lo que

$$\nu(h^*)\mu_{f^*}(S_{f^*}) = \rho_{f^*} - \rho^* \geq 0.$$

Puesto que $\inf_{f\in\mathbb{F}} \mu_f(S_f) > 0$ tenemos que $\mu_{f^*}(S_{f^*}) > 0$. Por lo tanto, $\nu(h^*) \geq 0$.

Con esto hemos demostrado que $\nu(h^*) = 0$. Entonces la terna (ρ^*, h^*, f^*) satisface las ecuaciones

$$h^*(x) = \inf_{a\in A(x)} [C(x,a) - \rho^* + Qh^*(x,a)] \tag{1.29}$$

$$(1.30)$$

$$= C_{f^*}(x) - \rho^* + Q_{f^*}h^*(x) \quad \forall x \in X,$$

lo cual demuestra la parte **(a)** del teorema.

La parte (b) se sigue directamente del Teorema 1.4.1 y el Lema 1.5.3.

Para la demostración de la parte **(c)** recordemos que $\mathbb{F}^*$ es la clase de las políticas estacionarias óptimas. Sea $f \in \mathbb{F}^*$ y consideremos la solución de la ecuación de Poisson h_f que satisface la condición $\nu(h_f) = 0$. Entonces

$$\begin{aligned} h_f(x) &= L_f h_f(x) = Lh_f(x, f(x)) \\ &\geq \inf_{a\in A(x)} Lh_f(x,a) = Th_f(x) \quad \forall x \in X. \end{aligned}$$

De la desigualdad anterior se sigue

$$h_f \geq T^n h_f \quad \forall n \in \mathbb{N}.$$

Puesto que h^* es el único punto fijo de T en $L_W(X)$ y además $T^n h_f \to h^*$, vemos que

$$h_f \geq h^* \quad \forall n \in \mathbb{N}.$$

Por otra parte, sabemos que $\rho^* = \rho_{f^*}$ y $h^* = h_{f^*}$. Por lo tanto

$$h^* = \inf_{f\in\mathbb{F}^*} h_f = h_{f^*}.$$

■

Capítulo 2

Juego semi-markoviano de suma cero

2.1. Introducción

En este capítulo se estudia el juego semi-markoviano (JSM) de suma cero en costo promedio y demostraremos la existencia de estrategias óptimas bajo condiciones apropiadas.

En los JSM de suma cero se tiene un sistema dinámico cuyo comportamiento es controlado por dos jugadores mediante estrategias. El sistema cuenta una función de pago que permite medir su comportamiento y sus valores dependen de las estrategias elegidas por los jugadores, donde una ganancia para uno de los jugadores es una pérdida para el otro. El objetivo del juego es encontrar estrategias de "equilibrio". Para definir el sistema del JSM de suma cero se necesita definir un modelo de juego semi-markoviano, conjuntos de estrategias para cada jugador, y un índice de pago (en costo promedio). En las secciones 2.2, 2.3 y 2.4 definiremos estos conceptos.

El enfoque que usaremos para demostrar la existencia de estrategias óptimas consiste en garantizar la existencia de soluciones de la ecuación de Shapley. Demostraremos que si existen soluciones a la ecuación de Shapley con ciertas propiedades, entonces existen estrategias de equilibrio para el JSM de suma cero en costo promedio. La ecuación de Shapley y su relación con la existencia de equilibrios se presentan en la sección 2.4.

Para demostrar la existencia de soluciones para la ecuación de Shapley se necesita que juego satisfaga algunas condiciones. En este trabajo se va a considerar cuatro clases de condiciones: condiciones de Lyapunov, de continuidad y compacidad, de crecimiento, y una condición sobre el tiempo medio de permanencia. Las condiciones de Lyapunov nos permitirán obtener propiedades de ergodicidad y estabilidad de los procesos asociados a las clases de las estrategias estacionarias. Las condiciones de continuidad y compacidad nos permitirán encontrar equilibrios para ciertos "juegos estáticos" asociados al JSM. Estas clases de condiciones y algunos resultados derivados de ellas se verán en la sección 2.5, mientras que en la sección 2.6 demostraremos la existencia de estrategias de equilibrio para el JSM por medio de "argumentos de punto fijo".

En la mayoría de los trabajos relacionados con juegos markovianos y semi-markovianos se ha utilizado kérneles de transición fuertemente continuos (véase [15] y sus referencias). Pero en años recientes se ha mostrado en los trabajos [13, 16] que es posible considerar kérneles de transición débilmente continuos. En este trabajo también utilizaremos la condición de continuidad débil del kérneles de transición para el JSM de suma cero en pago promedio utilizando en el enfoque de "argumentos de punto fijo" usado en [23].

En la sección 2.7 se presenta un problema de control mini-max para un sistema de inventario, el cual satisface las condiciones de optimalidad de la sección 2.5. Finalizaremos el estudio del JSM de suma cero en costo promedio en la sección 2.8 con una breve comparación de nuestro trabajo con los resultados de la referencia [13].

2.2. El modelo de juego semi-markoviano

Para plantear un problema de optimalidad en un juego semi-markoviano es necesario definir un modelo de juego semi-markoviano, los conjuntos de estrategias admisibles para los jugadores y el índice de funcionamiento bajo el cual se planteara el equilibrio del juego.

Definición 2.2.1 *Un **juego semi-markoviano (JSM) de suma cero de dos jugadores** consta de los objetos*

$$\mathcal{J} = (X, A, B, \mathbb{K}_A, \mathbb{K}_B, F, r)$$

donde:

(a) X *es el* ***espacio de estados posibles*** *del juego.*

(b) A *y* B *son los* ***espacios de acciones (controles) disponibles*** *del primer y segundo jugador para influir en el comportamiento del sistema, respectivamente.*

Supondremos que los conjuntos X, A *y* B *son espacios de Borel, i.e., son subconjuntos de Borel de espacios métricos separables y completos.*

(c) *Los conjuntos de restricciones* $\mathbb{K}_A$ *y* $\mathbb{K}_B$ *son subconjuntos de Borel de* $X \times A$ *y* $X \times B$*, respectivamente. Para cada* $x \in X$*, las* x*-secciones*

$$\begin{aligned} A(x) &= \{a \in A : (x,a) \in \mathbb{K}_A\} \\ B(x) &= \{b \in B : (x,b) \in \mathbb{K}_B\} \end{aligned}$$

representan los conjuntos de acciones admisibles para el primer y el segundo jugador en el estado x*, respectivamente. El conjunto*

$$\mathbb{K} := \{(x,a,b) : x \in X, (x,a) \in \mathbb{K}_A, (x,b) \in \mathbb{K}_B\},$$

pertenece a la $\sigma-$*álgebra de Borel del producto cartesiano* $X \times A \times B$ *(véase [18]).*

(d) F *es un kérnel estocástico sobre* $\mathbb{R}_+ \times X$ *dado* $\mathbb{K}$*. Definamos*

$$\begin{aligned} F(t,D|x,a,b) :=& \quad F([0,t] \times D|x,a,b) \\ G(t|x,a,b) :=& \quad F(t,X|x,a,b) \\ Q(D|x,a,b) :=& \quad F(\mathbb{R}_+,X|x,a,b) \end{aligned}$$

para cada $t \in \mathbb{R}_+$, $D \in \mathcal{B}(X)$ *y* $(x,a,b) \in \mathbb{K}$*. Suponemos que el kérnel* F *satisface lo siguiente:*

(d.1) $F(\cdot|x,a,b)$ *es una medida regular sobre* $\mathbb{R}_+ \times X$ *para cada* $(x,a,b) \in \mathbb{K}$,

(d.2) *para cada* $t \in \mathbb{R}_+$, $D \in \mathcal{B}(X)$ *y* $(x,a,b) \in \mathbb{K}$ *se cumple que*

$$F(t,D|x,a,b) = G(t|x,a,b)Q(D|x,a,b).$$

(e) $r : \mathbb{K} \times \mathbb{R}_+ \to \mathbb{R}$ *es una función medible y representa la función de* ***pago por etapa****.*

Intuitivamente nuestro modelo $\mathcal{J}$ representa un sistema dinámico que evoluciona en el tiempo de la siguiente manera: el juego inicia en el estado $x_0 = x \in X$ y los jugadores seleccionan las acciones $a_0 = a \in A(x)$ (jugador 1) y $b_0 = b \in B(x)$ (jugador 2). Entonces el juego permanece en el estado x_0 un tiempo aleatorio $\delta_1 = l \in \mathbb{R}_+$ determinado por la función de distribución $G(\cdot|x_0, a_0, b_0)$, es decir,

$$G(t|x, a, b) = \Pr[\delta_1 \leq t|x_0 = x, a_0 = a, b_0 = b].$$

El primer jugador recibe un pago $r(x, a, b, l)$ del segundo jugador y el juego pasa al estado $x_1 = y \in X$ de acuerdo a la medida de probabilidad $Q(\cdot|x_0, a_0, b_0)$, es decir,

$$Q(D|x, a, b) = \Pr[x_1 \in D|x_0 = x, a_0 = a, b_0 = b].$$

Después de la transición al estado y ambos jugadores seleccionan las acciones $a_1 = c \in A(y)$ y $b_1 = d \in B(y)$. El juego permanece en el estado y un tiempo aleatorio $\delta_2 = k \in \mathbb{R}_+$ determinado por $G(\cdot|y, c, d)$ y así el primer jugador recibe un pago $r(y, c, d, k)$ del segundo jugador. Este proceso se repite indefinidamente generando el proceso

$$(x_0, a_0, b_0, \delta_1, x_1, a_1, b_1, \delta_2, ...),$$

donde x_n y a_n, b_n son el estado y las acciones elegidas por los jugadores en la etapa n de juego, y δ_n es el tiempo que el juego permanece en el estado x_n antes de pasar al siguiente estado x_{n+1}. Ahora, para cada $n \in \mathbb{N}$, el **conjunto de historias admisibles** hasta la etapa n es el conjunto $H_n := \mathbb{K} \times \mathbb{R}_+ \times H_{n-1}$, donde $H_0 := X$. Una **historia admisible** hasta la etapa n es un elemento j_n de H_n y es denotado por

$$j_n := (x_0, a_0, b_0, \delta_1, \ldots, x_{n-1}, a_{n-1}, b_{n-1}, \delta_n, x_n).$$

2.3. Estrategias de juego

Las "reglas" por las cuales los jugadores eligen sus acciones en cada etapa del juego se llaman estrategias. La forma en cómo se definen estas sucesiones de reglas es análoga a la forma en que se definen las políticas de control del primer capítulo, solo que ahora se consideran dos controladores (jugadores).

Definición 2.3.1 *Una **estrategia de control** $\pi_1 = \{\pi_1^n\}_{n\in\mathbb{N}_0}$ para el primer jugador es una sucesión de kérneles estocásticos π_1^n sobre A dado H_n tales que*

$$\pi_1^n(A(x_n)|j_n) = 1 \quad \forall j_n \in H_n,\ n \in \mathbb{N}_0.$$

La clase de tales estrategias será denotada por Π_1. Las estrategias para el segundo jugador y la clase Π_2 de estas estrategias se definen de forma similar, remplazando A por B.

En este capítulo las estrategias de mayor interés son las estacionarias aleatorizadas y se definen de la siguiente manera: para cada $x \in X$, denotaremos por $\mathbb{A}(x)$ al conjunto $\mathbb{P}(A(x))$ y por Φ_1 al conjunto de todos los kérneles estocásticos φ_1 sobre A dado X tales que, $\varphi_1(\cdot|x) \in \mathbb{A}(x)$ para cada $x \in X$. Los conjuntos $\mathbb{B}(x)$ y Φ_2 para el segundo jugador se definen de forma similar, remplazando A por B. Diremos que una estrategia $\pi_i \in \Pi_i$ $(i = 1, 2)$ es **estacionaria** si existe $\varphi_i \in \Phi_i$ tal que

$$\pi_i^n(\cdot|j_n) = \varphi_i(\cdot|x_n) \quad \forall j_n \in H_n, n \in \mathbb{N}_0. \tag{2.1}$$

Abusando de la notación, si $\pi_i \in \Pi_i$ $(i = 1, 2)$ es una estrategia estacionaria y $\varphi_i \in \Phi_i$ satisface (2.1), entonces escribiremos φ_i en lugar de π_i. En adelante Φ_i va a representar a la clase de las estrategias estacionarias para el jugador i. De este modo $\Phi_i \subset \Pi_i$, para $i = 1, 2$.

Notación 2.3.2 ***(a)*** *Para cada función $u : \mathbb{K} \to \mathbb{R}$ medible, $(\varphi_1, \varphi_2) \in$*

$\Phi_1 \times \Phi_2$ y $(\gamma_1, \gamma_2) \in \mathbb{A}(x) \times \mathbb{B}(x)$ defina

$$\begin{aligned} u_{\varphi_1,\varphi_2}(x) &:= \int_{B(x)} \int_{A(x)} u(x,a,b)\varphi_1(da|x)\varphi_2(db|x), \\ u_{\gamma_1,\varphi_2}(x) &:= \int_{B(x)} \int_{A(x)} u(x,a,b)\gamma_1(da)\varphi_2(db|x), \\ u_{\varphi_1,\gamma_2}(x) &:= \int_{B(x)} \int_{A(x)} u(x,a,b)\varphi_1(da|x)\gamma_2(db), \\ u_{\gamma_1,\gamma_2}(x) &:= \int_{B(x)} \int_{A(x)} u(x,a,b)\gamma_1(da)\gamma_2(db), \end{aligned}$$

para cada $x \in X$ y aquellas funciones u para las cuales las integrales tengan sentido. En particular, con esta notación tendremos

$$Q_{\varphi_1,\varphi_2}(\cdot|x) := \int_{B(x)} \int_{A(x)} Q(\cdot|x,a,b)\varphi_1(da|x)\varphi_2(db|x),$$

para cada $x \in X$.

***(b)** Para cada función $w : X \to \mathbb{R}$ medible defina*

$$Qw(x,a,b) := \int_X w(y)Q(dy|x,a,b),$$

siempre que la integral esté bien definida. Combinando esta notación con la del inciso (a), para cada $(\varphi_1, \varphi_2) \in \Phi_1 \times \Phi_2$ y $(\gamma_1, \gamma_2) \in \mathbb{A}(x) \times \mathbb{B}(x)$ se tiene

$$\begin{aligned} Q_{\varphi_1,\varphi_2}w(x) &:= \int_{B(x)} \int_{A(x)} Qw(x,a,b)\varphi_1(da|x)\varphi_2(db|x) \\ Q_{\gamma_1,\varphi_2}w(x) &:= \int_{B(x)} \int_{A(x)} Qw(x,a,b)\gamma_1(da)\varphi_2(db|x), \\ Q_{\varphi_1,\gamma_2}w(x) &:= \int_{B(x)} \int_{A(x)} Qw(x,a,b)\varphi_1(da|x)\gamma_2(db), \\ Q_{\gamma_1,\gamma_2}w(x) &:= \int_{B(x)} \int_{A(x)} Qw(x,a,b)\gamma_1(da)\gamma_2(db). \end{aligned}$$

para cada $x \in X$ y aquellas funciones w para las cuales las integrales tengan sentido.

(c) *Sea ν una medida sobre X. Para cada función $u : X \to \mathbb{R}$ medible defina*

$$\nu(u) := \int_X u(y)\nu(dy).$$

siempre que la integral esté bien definida.

El siguiente teorema garantiza la existencia del proceso $\{x_n, a_n, b_n, \delta_{n+1}\}_{n\in\mathbb{N}_0}$ con las distribuciones marginales determinadas por la distribución del estado inicial del sistema ν, las estrategias π_1 y π_2 con las que eligen sus acciones los jugadores, la ley de transición $Q(\cdot|\cdot,\cdot,\cdot)$ y la distribución de los tiempos de permanencia $G(\cdot|\cdot,\cdot,\cdot)$. Por el Teorema de Ionescu-Tulcea [3, Theorem 2.7.2] se tiene el siguiente resultado.

Teorema 2.3.3 *Sea $\Omega := (\mathbb{K} \times \mathbb{R}_+)^\infty$ y $\mathcal{F}$ la σ-álgebra producto correspondiente. Para cada $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ y cada $\nu \in \mathbb{P}(X)$ existe una medida de probabilidad $P_\nu^{\pi_1,\pi_2}$ definida sobre el espacio medible $(\Omega, \mathcal{F})$ y un proceso estocástico $\{x_n, a_n, b_n, \delta_{n+1}\}_{n\in\mathbb{N}_0}$ que satisfacen las siguientes propiedades para cada $n \in \mathbb{N}_0$:*

(a) $P_\nu^{\pi_1,\pi_2}[x_0 \in D] = \nu(D) \quad \forall D \in \mathcal{B}(X)$;

(b) $P_\nu^{\pi_1,\pi_2}[a_n \in C_1|j_n] = \pi_1^n(C_1|j_n) \quad \forall C_1 \in \mathcal{B}(A)$;

(c) $P_\nu^{\pi_1,\pi_2}[b_n \in C_2|j_n] = \pi_2^n(C_2|j_n) \quad \forall C_2 \in \mathcal{B}(B)$;

(d) $P_\nu^{\pi_1,\pi_2}[x_{n+1} \in D|j_n, a_n, b_n, \delta_{n+1}] = Q(D|x_n, a_n, b_n) \quad \forall D \in \mathcal{B}(X)$.

(e) $P_\nu^{\pi_1,\pi_2}[\delta_{n+1} \le t|j_n, a_n, b_n] = G(t|x_n, a_n, b_n) \quad \forall t \ge 0$.

La esperanza con respecto a la medida de probabilidad $P_\nu^{\pi_1,\pi_2}$ será denotada por $E_\nu^{\pi_1,\pi_2}$. La medida ν en el Teorema 2.3.3 representa la **distribución inicial** del juego; si la medida ν está concentrada en un estado $x_0 = x \in X$, es decir,

$$\nu(D) = I_D(x) \quad \forall D \in \mathcal{B}(X),$$

entonces escribiremos $P_x^{\pi_1,\pi_2}$ en lugar de $P_\nu^{\pi_1,\pi_2}$, y a x le llamaremos el **estado inicial** del juego. Nos vamos a referir a x_n, a_n, y b_n como las variables de estado y de control del primer y segundo jugador en la etapa n, respectivamente.

2.4. El índice de pago promedio

El **pago esperado promedio** para el par de estrategias $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ dado el estado inicial $x_0 = x \in X$ se define como

$$J(x, \pi_1, \pi_2) := \liminf_{n\to\infty} \frac{E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} r(x_k, a_k, b_k, \delta_{k+1})}{E_x^{\pi_1,\pi_2} T_n},$$

donde $T_{n+1} := T_n + \delta_{n+1}$, con $T_0 \equiv 0$.

A las funciones

$$N(x) := \sup_{\pi_1\in\Pi_1} \inf_{\pi_2\in\Pi_2} J(x, \pi_1, \pi_2),$$

$$M(x) := \inf_{\pi_2\in\Pi_2} \sup_{\pi_1\in\Pi_1} J(x, \pi_1, \pi_2).$$

se les llama **valor inferior** y **valor superior** del juego, respectivamente. En general se tiene que

$$N(x) \leq M(x) \quad \forall x \in X. \tag{2.2}$$

Si se satisface la igualdad en (2.2), diremos que el juego tiene un valor y a la función $\rho : X \to \mathbb{R}$ definida por

$$\rho(x) := N(x) = M(x).$$

le llamaremos el **valor del juego** semi-markoviano**.**

Supongamos que el juego tiene valor ρ y sea $\varepsilon > 0$. Diremos que una estrategia $\pi_1^* \in \Pi_1$ es ε**-óptima para el primer jugador** si se satisface la desigualdad

$$\inf_{\pi_2\in\Pi_2} J(x, \pi_1^*, \pi_2) \geq \rho(x) - \varepsilon \quad \forall x \in X. \tag{2.3}$$

En caso de que la desigualdad (2.3) se satisfaga con $\varepsilon = 0$, diremos que π_1^* es **óptima para el primer jugador**. De manera similar, diremos que una estrategia $\pi_2^* \in \Pi_2$ es ε-**óptima para el segundo jugado**r si cumple

$$\sup_{\pi_1\in\Pi_1} J(x, \pi_1, \pi_2^*) - \varepsilon \leq \rho(x) \quad \forall x \in X. \tag{2.4}$$

Si la desigualdad (2.4) se satisface con $\varepsilon = 0$, diremos que π_2^* es **óptima para el segundo jugador.**

La función

$$\tau(x,a,b) := \int_0^{+\infty} tG(dt|x,a,b) \quad \forall (x,a,b) \in \mathbb{K},$$

es el **tiempo medio de permanencia** en el estado $x \in X$ cuando se toman las decisiones $a \in A(x)$ y $b \in B(x)$, mientras que

$$R(x,a,b) := \int_0^{+\infty} r(x,a,b,t)G(dt|x,a,b) \quad \forall (x,a,b) \in \mathbb{K},$$

es el **pago esperado** en el estado $x \in X$ cuando se toman las decisiones $a \in A(x)$ y $b \in B(x)$.

Observación 2.4.1 *Para cada* $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$, $x \in X$, *y* $k \in \mathbb{N}_0$ *se satisface*

$$E_x^{\pi_1,\pi_2} r(x_k, a_k, b_k, \delta_{k+1}) = E_x^{\pi_1,\pi_2} R(x_k, a_k, b_k), \tag{2.5}$$

$$E_x^{\pi_1,\pi_2} \delta_{k+1} = E_x^{\pi_1,\pi_2} \tau(x_k, a_k, b_k). \tag{2.6}$$

Para verificar esta observación note que para $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$, $x \in X$, y $k \in \mathbb{N}_0$ se cumple lo siguiente

$$\begin{aligned} E_x^{\pi_1,\pi_2}\left[r(x_k, a_k, b_k, \delta_{k+1})|j_k, a_k, b_k\right] &= E_x^{\pi_1,\pi_2}\left[r(x_k, a_k, b_k, \delta_{k+1})|x_k, a_k, b_k\right] \\ &= \int_0^{+\infty} r(x_k, a_k, b_k, t)G(dt|x_k, a_k, b_k) \\ &= R(x_k, a_k, b_k). \end{aligned}$$

De igual manera tenemos

$$\begin{aligned} E_x^{\pi_1,\pi_2}\left[\delta_{k+1}|j_k, a_k, b_k\right] &= E_x^{\pi_1,\pi_2}\left[\delta_{k+1}|x_k, a_k, b_k\right] \\ &= \int_0^{+\infty} tG(dt|x_k, a_k, b_k) \\ &= \tau(x_k, a_k, b_k) \end{aligned}$$

Tomando esperanza se obtiene (2.5) y (2.6).

Observación 2.4.2 *Para cada* $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ *y* $x \in X$, *la función* $J(x, \pi_1, \pi_2)$ *se puede re-escribir como*

$$J(x, \pi_1, \pi_2) := \liminf_{n\to\infty} \frac{E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} R(x_k, a_k, b_k)}{E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k)}. \tag{2.7}$$

Para ver que se cumple esta observación tome $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$, $x \in X$, y $n \in \mathbb{N}_0$ arbitrarios. Ya que $T_n = \sum_{k=0}^{n-1} \delta_{k+1}$, por la Observación 2.4.1 se tiene que

$$\begin{aligned} E_x^{\pi_1,\pi_2} T_n &= E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} \delta_{k+1} = \sum_{k=0}^{n-1} E_x^{\pi_1,\pi_2} \delta_{k+1} \\ &= \sum_{k=0}^{n-1} E_x^{\pi_1,\pi_2} \tau(x_k, a_k, b_k) \\ &= E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k). \end{aligned}$$

Con un procedimiento análogo se demuestra que

$$E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} r(x_k, a_k, b_k, \delta_{k+1}) = E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} R(x_k, a_k, b_k).$$

El enfoque que utilizaremos para garantizar la existencia de equilibrios en juegos semi-markovianos consiste en establecer condiciones para la existencia de soluciones de la llamada **ecuación de Shapley**. Diremos que (ρ, h), con $\rho \in \mathbb{R}$ y $h : X \to \mathbb{R}$ medible, es una solución de la ecuación de Shapley si

$$\begin{aligned} h(x) &= \sup_{\gamma_1 \in \mathbb{A}(x)} \inf_{\gamma_2 \in \mathbb{B}(x)} \left[R_{\gamma_1,\gamma_2}(x) - \rho \tau_{\gamma_1,\gamma_2}(x) + Q_{\gamma_1,\gamma_2} h(x) \right] \\ &= \inf_{\gamma_2 \in \mathbb{B}(x)} \sup_{\gamma_1 \in \mathbb{A}(x)} \left[R_{\gamma_1,\gamma_2}(x) - \rho \tau_{\gamma_1,\gamma_2}(x) + Q_{\gamma_1,\gamma_2} h(x) \right] \end{aligned}$$

para cada $x \in X$.

Sea $\varepsilon > 0$ y (ρ, h) una solución de la ecuación de Shapley. Nuestro objetivo es encontrar un par de estrategias estacionarias $(\varphi_1^\varepsilon, \varphi_2) \in \Phi_1 \times \Phi_2$ tales que

$$h(x) = \sup_{\gamma_1 \in \mathbb{A}(x)} \left[R_{\gamma_1,\varphi_2}(x) - \rho \tau_{\gamma_1,\varphi_2}(x) + Q_{\gamma_1,\varphi_2} h(x) \right] \tag{2.8}$$

$$\leq \inf_{\gamma_2 \in \mathbb{B}(x)} \left[R_{\varphi_1^\varepsilon,\gamma_2}(x) - \rho \tau_{\varphi_1^\varepsilon,\gamma_2}(x) + Q_{\varphi_1^\varepsilon,\gamma_2} h(x) \right] + \varepsilon. \tag{2.9}$$

El siguiente teorema muestra la relación entre la ecuación de Shapley y la existencia de estrategias óptimas.

Teorema 2.4.3 *Suponga que (ρ, h) es una solución de la ecuación de Shapley y que para cada $\varepsilon > 0$ existe un par de estrategias estacionarias $(\varphi_1^\varepsilon, \varphi_2) \in \Phi_1 \times \Phi_2$ que satisface (2.8) y (2.9). Si para cada $x \in X$ y $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ se satisface para alguna constante σ*

$$\lim_{n\to\infty} E_x^{\pi_1,\pi_2} \frac{1}{n} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k) \geq \sigma > 0, \tag{2.10}$$

y

$$\lim_{n\to\infty} \frac{E_x^{\pi_1,\pi_2} h(x_n)}{E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k)} = 0, \tag{2.11}$$

entonces ρ es el valor del juego, φ_1^ε es ε-óptima para el primer jugador y φ_2 es óptima para el segundo jugador.

Demostración. Sea $\varepsilon > 0$. Demostraremos que $\varphi_2 \in \Phi_2$ es óptima y $\varphi_1^\varepsilon \in \Phi_1$ es ε-óptima. Sean $\pi_1 = \{\pi_1^0, \pi_1^1, \ldots\} \in \Pi_1$, $n \in \mathbb{N}_0$, $j_n \in H_n$ con $x_0 = x \in X$ arbitrarios pero fijos. Observemos primero que por (2.8) se satisface

$$h(x) \geq \left[R_{\gamma_1,\varphi_2}(x) - \rho\tau_{\gamma_1,\varphi_2}(x) + Q_{\gamma_1,\varphi_2} h(x)\right] \quad \forall \gamma_1 \in \mathbb{A}(x). \tag{2.12}$$

Iterando esta desigualdad se obtiene

$$\begin{aligned} h(x) \quad &\geq \quad E_x^{\pi_1,\varphi_2} \sum_{k=0}^{n-1} R(x_k, a_k, b_k) - \rho E_x^{\pi_1,\varphi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k) \\ &\quad + E_x^{\pi_1,\varphi_2} h(x_n) \end{aligned} \tag{2.13}$$

para cada $n \in \mathbb{N}$. Dividiendo a ambos lados de (2.13) por $E_x^{\pi_1,\varphi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k)$ y haciendo $n \to \infty$, por (2.10) y (2.11) obtenemos que

$$\begin{aligned} 0 \quad &\geq \quad \liminf_{n\to\infty} \frac{E_x^{\pi_1,\varphi_2} \sum_{k=0}^{n-1} R(x_k, a_k, b_k)}{E_x^{\pi_1,\varphi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k)} - \rho \\ &= \quad J(x, \pi_1, \varphi_2) - \rho. \end{aligned} \tag{2.14}$$

Entonces

$$\begin{aligned} \rho \quad &\geq \quad \sup_{\pi_1 \in \Pi_1} J(x, \pi_1, \varphi_2) \\ &\geq \quad \inf_{\pi_2 \in \Pi_2} \sup_{\pi_1 \in \Pi_1} J(x, \pi_1, \pi_2) \\ &= \quad M(x). \end{aligned}$$

Ahora sea $\pi_2 \in \Pi_2$. Procediendo como antes obtenemos

$$\begin{aligned} 0 &\leq \liminf_{n\to\infty} \frac{E_x^{\varphi_1^\varepsilon,\pi_2} \sum_{k=0}^{n-1} R(x_k,a_k,b_k)}{E_x^{\varphi_1^\varepsilon,\pi_2} \sum_{k=0}^{n-1} \tau(x_k,a_k,b_k)} - \rho + \liminf_{n\to\infty} \frac{n\varepsilon}{E_x^{\varphi_1^\varepsilon,\pi_2} \sum_{k=0}^{n-1} \tau(x_k,a_k,b_k)} \qquad (2.15)\\ &= J(x,\varphi_1^\varepsilon,\pi_2) - \rho + \frac{\varepsilon}{\sigma}, \end{aligned}$$

lo cual implica que

$$\begin{aligned} \rho &\leq \inf_{\pi_2\in\Pi_2} J(x,\varphi_1^\varepsilon,\pi_2) + \frac{\varepsilon}{\sigma} \qquad (2.16)\\ &\leq \sup_{\pi_1\in\Pi_1} \inf_{\pi_2\in\Pi_2} J(x,\pi_1,\pi_2) + \frac{\varepsilon}{\sigma}\\ &= N(x) + \frac{\varepsilon}{\sigma}. \end{aligned}$$

Pero (2.16) es válido para cada $\varepsilon > 0$, de donde se sigue que

$$M(x) \leq \rho \leq N(x).$$

Por lo tanto $M(x) = N(x) = \rho$ para cada $x \in X$, i.e., ρ es el valor del juego. Así, por (2.14) y (2.15) hemos llegado al resultado deseado. ■

En lo que sigue, probaremos que bajo condiciones adecuadas en el juego, existen $h : X \to \mathbb{R}$, $\rho \in \mathbb{R}$ y $(\varphi_1^\varepsilon, \varphi_2) \in \Phi_1 \times \Phi_2$ ($\varepsilon > 0$) que satisfacen la ecuación de Shapley, las desigualdades (2.8) y (2.9), y los límites (2.10) y (2.11). Por el Teorema 2.4.3 esto garantiza la existencia de estrategias óptimas (ε-óptimas).

Las condiciones que se impondrán al juego son variaciones de las condiciones para el caso de procesos de control markoviano del capítulo anterior. Además se incluye una condición relacionada con la función τ.

Veamos ahora una observación que será de utilidad para las secciones siguientes.

Observación 2.4.4 *Sea $(\varphi_1, \varphi_2) \in \Phi_1 \times \Phi_2$ fijo pero arbitrario. El proceso de los estados $\{x_n\}_{n\in\mathbb{N}_0}$ es una cadena de Markov con kérnel de transición en un paso $Q_{\varphi_1,\varphi_2}(\cdot|\cdot)$. En este caso, $Q^n_{\varphi_1,\varphi_2}(\cdot|\cdot)$ representa el kérnel de transición en la n-ésima etapa, para $n \in \mathbb{N}_0$. Entonces, para cada $w : X \to \mathbb{R}$ se satisface*

$$Q^n_{\varphi_1,\varphi_2} w(x) := \int_X w(y) Q^n_{\varphi_1,\varphi_2}(dy|x) = E_x^{\varphi_1,\varphi_2} w(x_n),$$

para cada $x \in X$ y $n \in \mathbb{N}_0$.

2.5. Condiciones de optimalidad

En esta sección veremos las condiciones que le impondremos al modelo para poder garantizar la existencia de estrategias óptimas (ε-optimas).

Condición 2.5.1 (D1) *Existe una constante positiva σ tal que, $\tau(x,a,b) \geq \sigma > 0$ para cada $(x,a,b) \in \mathbb{K}$.*

Observación 2.5.2 *Suponga que se satisface D1. Para cada $(\pi_1,\pi_2) \in \Pi_1 \times \Pi_2$*

$$\lim_{n\to\infty} \frac{1}{n} E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} \tau(x_k,a_k,b_k) \geq \sigma > 0.$$

Condición 2.5.3 (D2) *Existe una función medible $W : X \to \mathbb{R}$ tal que $W \geq 1$ y $\max\{\tau(x,a,b), |R(x,a,b)|\} \leq cW(x)$ para toda $(x,a,b) \in \mathbb{K}$, donde c es una constante positiva.*

La $W-$norma $\|\cdot\|_W$ y los conjuntos $B_W(X)$, $L_W(X)$ y $C_W(W)$ se definen de la misma forma como se hizo en la sección 1.5 del primer capítulo.

Condición 2.5.4 (D3) *Existe una función medible $\phi : \mathbb{K} \to \mathbb{R}_+$, una medida no trivial ν sobre X y una constante $\lambda \in (0,1)$ tal que, para cada $(x,a,b) \in \mathbb{K}$, $B \in \mathcal{B}(X)$, y $(\varphi_1,\varphi_2) \in \Phi_1 \times \Phi_2$ se cumple lo siguiente:*

(a) $\nu(W) < \infty$,

(b) $Q(B|x,a,b) \geq \nu(B)\phi(x,a,b)$,

(c) $QW(x,a,b) \leq \lambda W(x) + \nu(W)\phi(x,a,b)$,

(d) $\nu(\phi_{\varphi_1,\varphi_2}) > 0$.

A continuación veremos una serie de resultados que son consecuencia de las condiciones D1, D2 y D3.

Lema 2.5.5 *Si se satisface D2 y D3 entonces,*

$$
\begin{aligned}
1 \quad &\leq \quad E_x^{\pi_1,\pi_2} W(x_n) \leq \lambda^n W(x) + \frac{\nu(W)}{(1-\lambda)\nu(X)} &(2.17)\\
&\leq \quad \lambda W(x) + \frac{\nu(W)}{(1-\lambda)\nu(X)} &(2.18)
\end{aligned}
$$

para cada $x \in X$, $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ *y* $n \in \mathbb{N}_0$. *Por lo tanto, para cada* $u \in B_W(X)$

$$
\lim_{n\to\infty} \frac{1}{n} E_x^{\pi_1,\pi_2} |u(x_n)| = 0, \tag{2.19}
$$

Si adicionalmente se satisface D1, también se cumple que

$$
\lim_{n\to\infty} \frac{E_x^{\pi_1,\pi_2} u(x_n)}{E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k)} = 0. \tag{2.20}
$$

Demostración. La demostración de (2.17) y (2.19) es análoga a la demostración del Lema 1.5.3, por lo que solo demostraremos (2.20). Sean $x \in X$, $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$, y $u \in B_W(X)$ arbitrarios pero fijos, entonces

$$
\begin{aligned}
\left| \frac{E_x^{\pi_1,\pi_2} u(x_n)}{E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k)} \right| \quad &\leq \quad \frac{E_x^{\pi_1,\pi_2} |u(x_n)|}{E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k)} \\
&\leq \quad \frac{\|u\|_W E_x^{\pi_1,\pi_2} W(x_n)}{E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k)} \\
&\leq \quad \left(\lambda W(x) + \frac{\nu(W)}{(1-\lambda)\nu(X)} \right) \frac{\|u\|_W}{E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k)}.
\end{aligned}
$$

De la Observación 2.5.2, se sigue (2.20), con lo que se concluye la demostración.
■

Observación 2.5.6 *Para cada* $B \in \mathcal{B}(X)$ *y* $(x, a, b) \in \mathbb{K}$, *defina*

$$
\hat{Q}(D|x, a, b) := Q(D|x, a, b) - \nu(D)\phi(x, a, b).
$$

Supongamos que se cumple D2. Entonces $\hat{Q}(\cdot|\cdot,\cdot,\cdot)$ *es un kérnel sobre* X *dado* $\mathbb{K}$ *que satisface la desigualdad*

$$
|\hat{Q}u(x, a, b) - \hat{Q}\upsilon(x, a, b)| \leq \lambda \|u - \upsilon\|_W
$$

para cada $(x, a, b) \in \mathbb{K}$ *y* $u, \upsilon \in B_W(X)$.

Demostración. Este resultado se prueba como la Observación 1.5.4. ■

Definición 2.5.7 *Para* $u, \upsilon \in B_W(W)$ *y* $(\varphi_1, \varphi_2) \in \Phi_1 \times \Phi_2$*, definimos* $L^{\upsilon}_{\varphi_1,\varphi_2} u : X \to \mathbb{R}$ *como*

$$\begin{aligned} L^{\upsilon}_{\varphi_1,\varphi_2} u(x) := & \quad \upsilon(x) + \int_X u(y) \hat{Q}_{\varphi_1,\varphi_2}(dy|x) \\ = & \quad \upsilon(x) + \int_X u(y) Q_{\varphi_1,\varphi_2}(dy|x) - \nu(u)\phi_{\varphi_1,\varphi_2}(x). \end{aligned}$$

Lema 2.5.8 *Suponga que D3 se satisface. Entonces, para cada* $\upsilon \in B_W(X)$ *y cada* $(\varphi_1, \varphi_2) \in \Phi_1 \times \Phi_2$*,* $L^{\upsilon}_{\varphi_1,\varphi_2}$ *es un operador de contracción de* $B_W(X)$ *en sí mismo con módulo* λ*. Así, por el teorema de punto fijo de Banach, existe una única función* $h^{\upsilon}_{\varphi_1,\varphi_2}(x) \in B_W(X)$ *tal que*

$$h^{\upsilon}_{\varphi_1,\varphi_2}(x) = \upsilon(x) + \int_X h^{\upsilon}_{\varphi_1,\varphi_2}(y) Q_f(dy|x) - \nu(h^{\upsilon}_{\varphi_1,\varphi_2})\phi_{\varphi_1,\varphi_2}(x) \quad \forall x \in X. \tag{2.21}$$

Demostración. La demostración de este lema es análoga a la del Lema 1.5.6 al remplazar $f \in \mathbb{F}$ por $(\varphi_1, \varphi_2) \in \Phi_1 \times \Phi_2$. ■

Lema 2.5.9 *Suponga que D3 se satisface. Entonces, para cada* $(\varphi_1, \varphi_2) \in \Phi_1 \times \Phi_2$ *existe una constante* $k_{\varphi_1,\varphi_2} > 0$ *tal que*

$$\lim_{n\to\infty} \frac{1}{n} E^{\varphi_1,\varphi_2}_x \sum_{k=0}^{n-1} \phi_{\varphi_1,\varphi_2}(x_n) = k_{\varphi_1,\varphi_2} \quad \forall x \in X.$$

Demostración. La demostración de este lema es análoga a la del Lema 1.5.7 al remplazar $f \in \mathbb{F}$ por $(\varphi_1, \varphi_2) \in \Phi_1 \times \Phi_2$. ■

En el siguiente teorema veremos una serie de resultados importantes, entre los que se encuentra la existencia de una medida de probabilidad invariante $\mu_{\varphi_1,\varphi_2}(\cdot)$ para la cadena de Markov con probabilidad de transición $Q_{\varphi_1,\varphi_2}(\cdot|\cdot)$.

Teorema 2.5.10 *Bajo D3 se satisface lo siguiente. Para cada* $(\varphi_1, \varphi_2) \in \Phi_1 \times \Phi_2$.

(a) *La cadena Markov con probabilidad de transición* $Q_{\varphi_1,\varphi_2}(\cdot|\cdot)$ *es* ν*-irreducible y Harris recurrente positiva con una única medida de probabilidad invariante* $\mu_{\varphi_1,\varphi_2}(\cdot)$*;*

(b) $\mu_{\varphi_1,\varphi_2}(W) < \infty$;

(c) $\mu_{\varphi_1,\varphi_2}(\phi_{\varphi_1,\varphi_2}) \geq \dfrac{1-\lambda}{\nu(W)}\mu_{\varphi_1,\varphi_2}(W) \geq \theta := \dfrac{1-\lambda}{\nu(W)} > 0$;

(d) *si* $u \in B_W(X)$, *entonces*

$$\lim_{n\to\infty} \frac{1}{n} E_x^{\varphi_1,\varphi_2} \sum_{k=1}^{n-1} u(x_k) = \mu_{\varphi_1,\varphi_2}(u) \quad \forall x \in X.$$

Si además D1 y D2 se satisfacen, entonces:

(e) *la constante*

$$\rho_{\varphi_1,\varphi_2} := \frac{\mu_{\varphi_1,\varphi_2}(R_{\varphi_1,\varphi_2})}{\mu_{\varphi_1,\varphi_2}(\tau_{\varphi_1,\varphi_2})} \tag{2.22}$$

es finita;

(f) *existe una única función* $h_{\varphi_1,\varphi_2} \in B_W(X)$ *que satisface la ecuación de Poisson*

$$\begin{aligned} h_{\varphi_1,\varphi_2}(x) &= R_{\varphi_1,\varphi_2}(x) - \rho_{\varphi_1,\varphi_2}\tau_{\varphi_1,\varphi_2}(x) \\ &+ \int_X h_{\varphi_1,\varphi_2}(y) Q_{\varphi_1,\varphi_2}(dy|x) \quad \forall x \in X, \end{aligned} \tag{2.23}$$

y $\nu(h_{\varphi_1,\varphi_2}) = 0$*;*

(g)

$$J(x,\varphi_1,\varphi_2) = \lim_{n\to\infty} \frac{E_x^{\varphi_1,\varphi_2} \sum_{k=0}^{n-1} R(x_k,a_k,b_k)}{E_x^{\varphi_1,\varphi_2} \sum_{k=0}^{n-1} \tau(x_k,a_k,b_k)} = \rho_{\varphi_1,\varphi_2},$$

para cada $x \in X$.

Demostración. Sea $(\varphi_1,\varphi_2) \in \Phi_1 \times \Phi_2$ arbitrario pero fijo.

Las demostraciones de las partes **(a)**, **(b)** y **(c)** son análogas a las demostraciones de las partes (a), (b) y (c) del Teorema 1.5.8, remplazando solamente $f \in \mathbb{F}$ por $(\varphi_1,\varphi_2) \in \Phi_1 \times \Phi_2$.

(**d**) Fijemos $u \in B_W(X)$. Para cada $w \in B_W(X)$ definamos la función $T^u_{\varphi_1,\varphi_2} w : X \to \mathbb{R}$ como

$$T^u_{\varphi_1,\varphi_2} w(x) = u(x) - \mu_{\varphi_1,\varphi_2}(u) + \hat{Q}_{\varphi_1,\varphi_2} w(x).$$

Por D3(c), $T^u_{\varphi_1,\varphi_2}$ mapea $B_W(X)$ en si mismo. Veamos que $T^u_{\varphi_1,\varphi_2}$ es un operador de contracción.

Sean $w, \upsilon \in B_W(X)$, entonces

$$\begin{aligned} \left| T^u_{\varphi_1,\varphi_2} w(x) - T^u_{\varphi_1,\varphi_2} \upsilon(x) \right| &\leq \int_X |w(y) - \upsilon(y)| \, \hat{Q}_{\varphi_1,\varphi_2}(dy|x) \\ &\leq \|w - \upsilon\|_W \int_X W(y) \hat{Q}_{\varphi_1,\varphi_2}(dy|x) \\ &\leq \|w - \upsilon\|_W \, \lambda W(x) \ \ \forall x \in X. \end{aligned}$$

Se sigue que $\left\| T^u_{\varphi_1,\varphi_2} w - T^u_{\varphi_1,\varphi_2} \upsilon \right\|_W \leq \lambda \|w - \upsilon\|_W$, i.e., $T^u_{\varphi_1,\varphi_2}$ es un operador de contracción de $B_W(X)$ en si mismo con módulo λ. Por el teorema de punto fijo de Banach, existe una función $h_u \in B_W(X)$ que satisface

$$\begin{aligned} h_u(x) &= u(x) - \mu_{\varphi_1,\varphi_2}(u) \\ &+ Q_{\varphi_1,\varphi_2} h_u(x) - \nu(h_u) \phi_{\varphi_1,\varphi_2}(x) \ \ \forall x \in X. \end{aligned}$$

Integrando respecto a $\mu_{\varphi_1,\varphi_2}(\cdot)$ obtenemos

$$\nu(h_u) \mu_{\varphi_1,\varphi_2}(\phi_{\varphi_1,\varphi_2}) = 0,$$

y por la parte (c) obtenemos $\nu(h_u) = 0$. Por lo tanto, h_u satisface la ecuación

$$h_u(x) = u(x) - \mu_{\varphi_1,\varphi_2}(u) + Q_{\varphi_1,\varphi_2} h_u(x) \ \forall x \in X. \tag{2.24}$$

Iterando (2.24), obtenemos para cada $n = 1, 2, \ldots$

$$\begin{aligned} h_u(x) &= E^{\varphi_1,\varphi_2}_x \sum_{k=0}^{n-1} u(x_k) \\ &\quad - n\mu_{\varphi_1,\varphi_2}(u) + E^{\varphi_1,\varphi_2}_x h_u(x_n), \end{aligned}$$

para cada $x \in X$. Dividiendo por n, haciendo $n \to \infty$, y aplicando el Lema 2.5.5 obtenemos

$$\mu_{\varphi_1,\varphi_2}(u) = \lim_{n\to\infty} \frac{1}{n} E^{\varphi_1,\varphi_2}_x \sum_{k=0}^{n-1} u(x_k) \ \ \forall x \in X, \tag{2.25}$$

(e) Las condiciones D1 y D2 nos dicen que $R_{\varphi_1,\varphi_2}, \tau_{\varphi_1,\varphi_2} \in B_W(X)$ y $\tau_{\varphi_1,\varphi_2}(\cdot) \geq \sigma$, obtenemos (2.22).

(**f**) Para cada $u \in B_W(X)$ definamos la función $\hat{T}_{\varphi_1,\varphi_2}u : X \to \mathbb{R}$ como

$$\hat{T}_{\varphi_1,\varphi_2}u(x) := \bar{R}_{\varphi_1,\varphi_2}(x) + \hat{Q}_{\varphi_1,\varphi_2}u(x),$$

donde $\bar{R}_{\varphi_1,\varphi_2}(\cdot) := R_{\varphi_1,\varphi_2}(\cdot) - \rho_{\varphi_1,\varphi_2}\tau_{\varphi_1,\varphi_2}(\cdot)$ y $\hat{Q}_{\varphi_1,\varphi_2}u(\cdot) := Q_{\varphi_1,\varphi_2}u(\cdot) - \nu(u)\phi_{\varphi_1,\varphi_2}(\cdot)$. Se puede observar que $\hat{T}_{\varphi_1,\varphi_2}$ es un operador de contracción de $B_W(X)$ en si mismo, con módulo λ. Entonces existe una única función $h_{\varphi_1,\varphi_2} \in B_W(X)$ tal que

$$h_{\varphi_1,\varphi_2}(x) = \bar{R}_{\varphi_1,\varphi_2}(x) + \hat{Q}_{\varphi_1,\varphi_2}h_{\varphi_1,\varphi_2}(x) \quad \forall x \in X.$$

Luego, integrando ambos lados de la ecuación respecto a $\mu_{\varphi_1,\varphi_2}(\cdot)$ se obtiene

$$\begin{aligned} \mu_{\varphi_1,\varphi_2}(h_{\varphi_1,\varphi_2}) &= \mu_{\varphi_1,\varphi_2}(R_{\varphi_1,\varphi_2}) - \rho_{\varphi_1,\varphi_2}\mu_{\varphi_1,\varphi_2}(\tau_{\varphi_1,\varphi_2}) \\ &\quad + \mu_{\varphi_1,\varphi_2}(h_{\varphi_1,\varphi_2}) - \nu(h_{\varphi_1,\varphi_2})\mu_{\varphi_1,\varphi_2}(\phi_{\varphi_1,\varphi_2}). \end{aligned}$$

Entonces, por la definición de $\rho_{\varphi_1,\varphi_2}$,

$$\nu(h_{\varphi_1,\varphi_2})\mu_{\varphi_1,\varphi_2}(\phi_{\varphi_1,\varphi_2}) = 0.$$

Pero $\mu_{\varphi_1,\varphi_2}(\phi_{\varphi_1,\varphi_2}) > 0$ por (c). Esto implica que

$$\nu(h_{\varphi_1,\varphi_2}) = 0$$

y así h_{φ_1,φ_2} satisface (2.23).

(**g**) La parte (d) implica que

$$\mu_{\varphi_1,\varphi_2}(\tau_{\varphi_1,\varphi_2}) = \lim_{n\to\infty} \frac{1}{n} E_x^{\varphi_1,\varphi_2} \sum_{k=0}^{n-1} \tau_{\varphi_1,\varphi_2}(x_k) > 0 \quad \forall x \in X,$$

ya que por D2 se tiene que $\tau_{\varphi_1,\varphi_2}(\cdot) \in B_W(X)$ y por D1 se cumple que $\tau_{\varphi_1,\varphi_2}(\cdot) > 0$.

Por otra parte, iterando (2.23) y aplicando inducción, obtenemos para cada $n = 1, 2, \ldots$

$$\begin{aligned} h_{\varphi_1,\varphi_2}(x) &= E_x^{\varphi_1,\varphi_2}\left[\sum_{k=0}^{n-1} R_{\varphi_1,\varphi_2}(x_k) - \rho_{\varphi_1,\varphi_2}\sum_{k=0}^{n-1}\tau_{\varphi_1,\varphi_2}(x_k)\right] \\ &\quad + E_x^{\varphi_1,\varphi_2}h_{\varphi_1,\varphi_2}(x_n) \quad \forall x \in X. \end{aligned}$$

Dividiendo por n, haciendo $n \to \infty$ y aplicando el Lema 2.5.5 obtenemos

$$0 = \lim_{n\to\infty} \frac{1}{n} E_x^{\varphi_1,\varphi_2} \sum_{k=0}^{n-1} R_{\varphi_1,\varphi_2}(x_k) - \rho_{\varphi_1,\varphi_2} \lim_{n\to\infty} \frac{1}{n} E_x^{\varphi_1,\varphi_2} \sum_{k=0}^{n-1} \tau_{\varphi_1,\varphi_2}(x_k).$$

Por (2.25), tenemos que

$$\begin{aligned} \rho_{\varphi_1,\varphi_2} &= \lim_{n\to\infty} \frac{E_x^{\varphi_1,\varphi_2} \sum_{k=0}^{n-1} R_{\varphi_1,\varphi_2}(x_k)}{E_x^{\varphi_1,\varphi_2} \sum_{k=0}^{n-1} \tau_{\varphi_1,\varphi_2}(x_k)} \\ &= J(x, \varphi_1, \varphi_2), \end{aligned}$$

para cada $x \in X$. ■

A continuación se muestra la última condición que le impondremos al modelo

Condición 2.5.11 (D4) *Para cada* $(x, a, b) \in \mathbb{K}$,

(a) R *es semicontinua inferiormente sobre* $\mathbb{K}$.

(b) *El mapeo* $x \mapsto A(x)$ *es semicontinuo superiormente y* $A(x)$ *es completo para cada* $x \in X$*; el mapeo* $x \mapsto B(x)$ *es semicontinuo inferiormente con* $B(x)$ *compacto y no-vacío para cada* $x \in X$.

(c) τ *es continua sobre* $\mathbb{K}$.

(d) *La ley de transición* Q *es débilmente continua, esto es,* Qu *es continua en* $\mathbb{K}$ *para cada* $u \in C_b(X)$.

(e) W *y* QW *son funciones continuas sobre* X *y* $\mathbb{K}$, *respectivamente.*

Asignamos a los conjuntos $\mathbb{A} := \mathbb{P}(A)$, $\mathbb{B} := \mathbb{P}(B)$, $\mathbb{A}(x)$ y $\mathbb{B}(x)$ $(x \in X)$ la topología de la convergencia débil, es decir, una sucesión $\{\gamma_n\}_{n\in\mathbb{N}}$ en $\mathbb{A}(x)$ converge a $\gamma \in \mathbb{A}(x)$ si y sólo si γ_n converge débilmente a γ, i.e.,

$$\int_{A(x)} u(a)\gamma_n(da) \to \int_{A(x)} u(a)\gamma(da),$$

para cada $u \in C_b(A(x))$. Esta topología es metrizable, es decir, existe una métrica en $\mathbb{A}(x)$ que induce la topología. Además $\mathbb{A}$, $\mathbb{B}$, $\mathbb{A}(x)$ y $\mathbb{B}(x)$ son espacios de Borel por ser A, B, $A(x)$ y $B(x)$ espacios de Borel (véase [4, Corollary 7.25.1]).

Ahora defina

$$\bar{\mathbb{K}} := \{(x, \gamma_1, \gamma_2) : x \in X,\ \gamma_1 \in \mathbb{A}(x),\ \gamma_2 \in \mathbb{B}(x)\}.$$

Por [18, Corollary 4.1 and Lemma 1.1], $\bar{\mathbb{K}}$ es un subconjunto de Borel del espacio $X \times \mathbb{A} \times \mathbb{B}$.

Lema 2.5.12 *Suponga que D2 y D4 se satisfacen. Para cada $u : \mathbb{K} \to \mathbb{R}$ s.c.i. en $\mathbb{K}$, el mapeo $(x, \gamma_1, \gamma_2) \mapsto u_{\gamma_1,\gamma_2}(x)$, donde*

$$u_{\gamma_1,\gamma_2}(x) = \int_{A(x)} \int_{B(x)} u(x, a, b)\gamma_2(db)\gamma_1(da),$$

es s.c.i. sobre $\bar{\mathbb{K}}$. Además, si existe una constante positiva k tal que $|u(x, a, b)| \leq kW(x)$ para cada $(x, a, b) \in \mathbb{K}$, entonces

$$\left|u_{\gamma_1,\gamma_2}(x)\right| \leq kW(x) \quad \forall\,(x, \gamma_1, \gamma_2) \in \mathbb{K}.$$

Demostración. Por el Teorema A.0.6 (en el Apéndice), existe una sucesión $\{u_n\}$ en $C_b(\mathbb{K})$ tales que $u_n \uparrow u$. El teorema de convergencia monótona implica que

$$\begin{aligned} \lim_{n\to\infty} (u_n)_{\gamma_1,\gamma_2}(x) &= \lim_{n\to\infty} \int_{A(x)} \int_{B(x)} u_n(x, a, b)\gamma_2(db)\gamma_1(da) \\ &= u_{\gamma_1,\gamma_2}(x) \quad \forall\,(x, \gamma_1, \gamma_2) \in \mathbb{K}. \end{aligned}$$

Por [19, Lemma 5.4], el mapeo $(x, \gamma_1, \gamma_2) \mapsto (u_n)_{\gamma_1,\gamma_2}(x)$ es continuo sobre $\bar{\mathbb{K}}$. Entonces, por el Teorema A.0.7 del apéndice, el mapeo $(x, \gamma_1, \gamma_2) \mapsto u_{\gamma_1,\gamma_2}(x)$ es semicontinuo sobre $\bar{\mathbb{K}}$. La segunda parte de este lema resulta evidente ya que $W(\cdot)$ no tiene argumentos en A y B. ■

Observemos que si se cumple la condición D4(b), por [12, Theorem 3] y [4, Propositions 7.22 and 7.23], el mapeo $x \mapsto \mathbb{A}(x)$ es semicontinuo superiormente y $\mathbb{A}(x)$ es completo para cada $x \in X$, mientras que el mapeo $x \mapsto \mathbb{B}(x)$ es semicontinuo inferiormente con $\mathbb{B}(x)$ compacto y no-vacío para cada $x \in X$.

Teorema 2.5.13 (de selección minimax) *Suponga que D4(b) se satisface. Sea* $U : \mathbb{K} \to \mathbb{R}$ *una función medible y defina*

$$\hat{U}(x) := \inf_{\gamma_2 \in \mathbb{B}(x)} \sup_{\gamma_1 \in \mathbb{A}(x)} U_{\gamma_1,\gamma_2}(x), \quad \forall x \in X.$$

Si el mapeo $(x, \gamma_1, \gamma_2) \mapsto U_{\gamma_1,\gamma_2}(x)$ *es s.c.i. sobre* $\bar{\mathbb{K}}$ *y* $U_{\gamma_1,\gamma_2}(\cdot) \in B_W(X)$ *para cada* $(x, \gamma_1, \gamma_2) \in \bar{\mathbb{K}}$, *entonces:*

(a) $\hat{U}$ *es s.c.i. en* X *y además*

$$\begin{aligned} \hat{U}(x) &= \inf_{\gamma_2 \in \mathbb{B}(x)} \sup_{\gamma_1 \in \mathbb{A}(x)} U_{\gamma_1,\gamma_2}(x) \\ &= \sup_{\gamma_1 \in \mathbb{A}(x)} \inf_{\gamma_2 \in \mathbb{B}(x)} U_{\gamma_1,\gamma_2}(x), \end{aligned}$$

(b) *existe* $\varphi_2 \in \Phi_2$ *tal que*

$$\hat{U}(x) = \sup_{\gamma_1 \in \mathbb{A}(x)} U_{\gamma_1,\varphi_2}(x),$$

(c) *para cada* $\varepsilon > 0$ *existe* $\varphi_1^\varepsilon \in \Phi_1$ *tal que*

$$\hat{U}(x) \leq \inf_{\gamma_2 \in \mathbb{B}(x)} U_{\gamma_1,\varphi_1^\varepsilon}(x) + \varepsilon.$$

Demostración. Consulte Lemma 3.3 en [13]. ■

Los teoremas de selección garantizan la existencia de selectores medibles, como se muestra en la Observación 1.5.12. Pero en el caso de juegos con dos o más jugadores ya no podemos garantizar la existencia de estrategias deterministas, por lo que los selectores deben de tomar valores en espacios de probabilidad. De este modo, los selectores en el Teorema 2.5.13 son los kérneles φ_2 y φ_1^ε.

Con lo discutido en esta sección ya estamos preparados para demostrar la existencia de estrategias estacionarias óptimas.

2.6. Existencia de estrategias estacionarias óptimas

El resultado más importante de este capítulo es el Teorema 2.6.1, el cual se enuncia líneas más abajo. En este teorema se afirma que si se satisfacen las condiciones D1-D4, entonces existe una solución a la ecuación de Shapley (ρ^*, h^*), una estrategia $\varphi_2^* \in \Phi_2$ óptima y para cada $\varepsilon > 0$ una estrategia $\varphi_1^{*,\varepsilon} \in \Phi_1$ tal que $\varphi_1^{*,\varepsilon}$ es ε-óptima.

Para la demostración del Teorema 2.6.1 se va a recurrir a una serie de resultados que, entre otras cosas, nos dicen que el valor del juego está dado por la constante finita

$$\rho^* = \sup_{\varphi_1 \in \Phi_1} \inf_{\varphi_2 \in \Phi_2} \rho_{\varphi_1,\varphi_2} = \inf_{\varphi_2 \in \Phi_2} \sup_{\varphi_1 \in \Phi_1} \rho_{\varphi_1,\varphi_2}.$$

Teorema 2.6.1 *Suponga que D1-D4 se satisfacen. Entonces existe $h^* \in L_W(X)$ y una constante ρ^* tal que*

$$\begin{aligned} h^*(x) &= \sup_{\gamma_1 \in \mathbb{A}(x)} \inf_{\gamma_2 \in \mathbb{B}(x)} \left[R_{\gamma_1,\gamma_2}(x) - \rho^* \tau_{\gamma_1,\gamma_2}(x) + Q_{\gamma_1,\gamma_2} h^*(x) \right] \\ &= \inf_{\gamma_2 \in \mathbb{B}(x)} \sup_{\gamma_1 \in \mathbb{A}(x)} \left[R_{\gamma_1,\gamma_2}(x) - \rho^* \tau_{\gamma_1,\gamma_2}(x) + Q_{\gamma_1,\gamma_2} h^*(x) \right] \quad \forall x \in X. \end{aligned}$$

Además, existe $\varphi_2^* \in \Phi_2$ tal que

$$h^*(x) = \sup_{\gamma_1 \in \mathbb{A}(x)} \left[R_{\gamma_1,\varphi_2^*}(x) - \rho^* \tau_{\gamma_1,\varphi_2^*}(x) + Q_{\gamma_1,\varphi_2^*} h^*(x) \right] \tag{2.26}$$

y para cada $\varepsilon > 0$, existe $\varphi_1^{*,\varepsilon} \in \Phi_1$ tal que

$$h^*(x) \leq \inf_{\gamma_2 \in \mathbb{B}(x)} \left[R_{\varphi_1^{*,\varepsilon},\gamma_2}(x) - \rho^* \tau_{\varphi_1^{*,\varepsilon},\gamma_2}(x) + Q_{\varphi_1^{*,\varepsilon},\gamma_2} h^*(x) \right] + \varepsilon. \tag{2.27}$$

Como consecuencia, ρ^* es el valor del juego, $\varphi_1^{*,\varepsilon}$ es una política ε-óptima para el primer jugador, mientras que φ_2^* es una política óptima para el segundo jugador.

Lema 2.6.2 *Suponga que D1-D3, y D4(e) se satisfacen, y sea $S := \phi^l$, donde ϕ^l es la envolvente inferior de ϕ (véase la Definición A.0.8 y el Lema A.0.9 en el Apéndice con $Y = \mathbb{K}$). Entonces*

$$QW(x,a,b) \leq \lambda W(x) + \nu(W)S(x,a,b) \quad \forall (x,a,b) \in \mathbb{K}. \tag{2.28}$$

Además

$$\mu_{\varphi_1,\varphi_2}(S_{\varphi_1,\varphi_2}) \geq \frac{1-\lambda}{\nu(W)}\mu_{\varphi_1,\varphi_2}(W) \geq \theta > 0 \quad \forall\, (\varphi_1,\varphi_2) \in \Phi_1 \times \Phi_2. \tag{2.29}$$

Demostración. La demostración de este lema es análoga a la del Lema 1.6.3 al remplazar $f \in \mathbb{F}$ por $(\varphi_1,\varphi_2) \in \Phi_1 \times \Phi_2$. ■

Corolario 2.6.3 *Si se satisfacen D1-D3, y D4(e), entonces*

$$\mu_{\varphi_1,\varphi_2}(W) \leq \frac{\nu(W)}{(1-\lambda)\nu(X)}, \tag{2.30}$$

para cada $(\varphi_1,\varphi_2) \in \Phi_1 \times \Phi_2$.

Demostración. Sean $x \in X$ y $(\varphi_1,\varphi_2) \in \Phi_1 \times \Phi_2$ arbitrarios fijos. Como $S \leq \phi$, entonces por D3(b), $Q_{\varphi_1,\varphi_2}(\cdot|x) \geq \nu(\cdot)S_{\varphi_1,\varphi_2}(x)$ de donde sigue que $1 \geq \nu(X)S_{\varphi_1,\varphi_2}(x)$. Entonces $1 \geq \nu(X)\mu_{\varphi_1,\varphi_2}(S_{\varphi_1,\varphi_2})$. Luego, por (2.29)

$$\frac{1}{\nu(X)} \geq \frac{1-\lambda}{\nu(W)}\mu_{\varphi_1,\varphi_2}(W),$$

y así obtenemos (2.30). ■

Lema 2.6.4 *Suponga que D1-D3, y D4(e) se satisfacen y defina*

$$\rho^l := \sup_{\varphi_1 \in \Phi_1} \inf_{\varphi_2 \in \Phi_2} \rho_{\varphi_1,\varphi_2} \quad y \quad \rho^u := \inf_{\varphi_2 \in \Phi_2} \sup_{\varphi_1 \in \Phi_1} \rho_{\varphi_1,\varphi_2}.$$

Entonces

$$|\rho^l| < \infty \quad y \quad |\rho^u| < \infty.$$

Demostración. Sea $(\varphi_1, \varphi_2) \in \Phi_1 \times \Phi_2$. Dado que $\mu_{\varphi_1,\varphi_2}(\tau_{\varphi_1,\varphi_2}) \geq \sigma > 0$ por D1, aplicando el Corolario 2.6.3 obtenemos

$$\begin{aligned} \left|\rho_{\varphi_1,\varphi_2}\right| &\leq \frac{\mu_{\varphi_1,\varphi_2}(\left|R_{\varphi_1,\varphi_2}\right|)}{\mu_{\varphi_1,\varphi_2}(\tau_{\varphi_1,\varphi_2})} \leq \frac{c\mu_{\varphi_1,\varphi_2}(W_{\varphi_1,\varphi_2})}{\mu_{\varphi_1,\varphi_2}(\tau_{\varphi_1,\varphi_2})} \qquad (2.31) \\ & \qquad (2.32) \\ &\leq \frac{c\nu(W)}{\sigma(1-\lambda)\nu(X)} = \frac{ck}{\sigma}, \end{aligned}$$

donde $k := \frac{\nu(W)}{(1-\lambda)\nu(X)}$. Entonces,

$$-\infty < -\frac{ck}{\sigma} \leq \rho_{\varphi_1,\varphi_2} \leq \frac{ck}{\sigma} < +\infty,$$

lo cual implica las desigualdades

$$-\infty < -\frac{ck}{\sigma} \leq \inf_{\psi_2 \in \Phi_2} \rho_{\varphi_1,\psi_2} \quad \forall \varphi_1 \in \Phi_1,$$

y

$$\sup_{\psi_1 \in \Phi_1} \rho_{\psi_1,\varphi_2} \leq \frac{ck}{\sigma} < +\infty \quad \forall \varphi_2 \in \Phi_2.$$

Por lo tanto,

$$\begin{aligned} -\infty &< -\frac{ck}{\sigma} \leq \rho^l = \sup_{\varphi_1 \in \Phi_1} \inf_{\varphi_2 \in \Phi_2} \rho_{\varphi_1,\varphi_2} \\ &\leq \inf_{\varphi_2 \in \Phi_2} \sup_{\varphi_1 \in \Phi_1} \rho_{\varphi_1,\varphi_2} = \rho^u \leq \frac{ck}{\sigma} < +\infty \end{aligned}$$

lo cual demuestra el resultado deseado. ■

Definición 2.6.5 *Para cada $u \in B_W(X)$ definimos las siguientes funciones sobre $\mathbb{K}$*

$$\begin{aligned} Pu(x,a,b) :=& \quad Qu(x,a,b) - \nu(u)S(x,a,b), \\ Lu(x,a,b) :=& \quad R(x,a,b) - \rho^l \tau(x,a,b) + Pu(x,a,b), \\ L^l u(x,a,b) :=& \quad (Lu)^l\,(x,a,b). \end{aligned}$$

Para cada $u \in B_W(X)$ definimos

$$Tu(x) := \sup_{\gamma_1 \in \mathbb{A}(x)} \inf_{\gamma_2 \in \mathbb{B}(x)} L^l_{\gamma_1,\gamma_2} u(x).$$

Lema 2.6.6 *Suponga que D1-D4 se satisfacen. Entonces T es un operador de contracción de $L_W(X)$ en si mismo con módulo λ y existe $h^* \in L_W(X)$ tal que $h^* = Th^*$.*

Demostración. Demostraremos primero que $Tu \in L_W(X)$ para cada $u \in B_W(X)$. Observemos que $P(\cdot|\cdot,\cdot,\cdot)$ es un kérnel sobre X dado $\mathbb{K}$ y que $PW \leq \lambda W$. Entonces

$$\begin{aligned}
|Lu(x,a,b)| &= \left|R(x,a,b) - \rho^l \tau(x,a,b) + Pu(x,a,b)\right| \\
&\leq |R(x,a,b)| + \left|\rho^l \tau(x,a,b)\right| + |Pu(x,a,b)| \\
&\leq cW(x) + c\left|\rho^l\right| W(x) + \lambda \|u\|_W W(x) \\
&\leq \kappa W(x)
\end{aligned}$$

para todo $(x,a,b) \in \mathbb{K}$, donde $\kappa := c + \left|\rho^l\right| c + \lambda \|u\|_W$. Se sigue que

$$\begin{aligned}
|Tu(x)| &= \left|\sup_{\gamma_1 \in \mathbb{A}(x)} \inf_{\gamma_2 \in \mathbb{B}(x)} L^l_{\gamma_1,\gamma_2} u(x)\right| \\
&\leq \sup_{\gamma_1 \in \mathbb{A}(x)} \sup_{\gamma_2 \in \mathbb{B}(x)} \left|L^l_{\gamma_1,\gamma_2} u(x)\right| \\
&\leq \kappa W(x) \quad \forall x \in X.
\end{aligned}$$

Por lo tanto

$$\|Tu\|_W \leq \kappa,$$

i.e., $Tu \in B_W(X)$ para cada $u \in B_W(X)$.

Por otra parte, el mapeo $(x,\gamma_1,\gamma_2) \mapsto L^l{}_{\gamma_1,\gamma_2} u(x)$ es s.c.i. en $\bar{\mathbb{K}}$ y $\left|u_{\gamma_1,\gamma_2}(x)\right| \leq \|u\|_W W(x)$ por el Lema 2.5.12. El Teorema de selección minimax (Teorema 2.5.13 (b)) y el Teorema A.0.7 (b) en el apéndice garantizan que $Tu \in L_W(X)$.

Ahora demostraremos que T es un operador de contracción. Primero observemos que

$$\begin{aligned}\left|L^l u(x,a,b) - L^l w(x,a,b)\right| &\leq \sup_{a\in A(x),\ b\in B(x)} |Lu(x,a,b) - Lw(x,a,b)| \\ &= \sup_{a\in A(x),\ b\in B(x)} |Pu(x,a,b) - Pw(x,a,b)| \\ &\leq \lambda \|u-w\|_W W(x) \quad \forall (x,a,b)\in \mathbb{K}.\end{aligned}$$

Entonces, por el Lema A.0.11 y la condición D3(c) tenemos que

$$\begin{aligned}|Tu(x) - Tw(x)| &= \left|\sup_{\gamma_1\in\mathbb{A}(x)} \inf_{\gamma_2\in\mathbb{B}(x)} L^l_{\gamma_1,\gamma_2}u(x) - \sup_{\gamma_1\in\mathbb{A}(x)} \inf_{\gamma_2\in\mathbb{B}(x)} L^l_{\gamma_1,\gamma_2}w(x)\right| \\ &\leq \sup_{\gamma_1\in\mathbb{A}(x)} \left|\inf_{\gamma_2\in\mathbb{B}(x)} L^l_{\gamma_1,\gamma_2}u(x) - \inf_{\gamma_2\in\mathbb{B}(x)} L^l_{\gamma_1,\gamma_2}w(x)\right| \\ &\leq \sup_{\gamma_1\in\mathbb{A}(x)} \sup_{\gamma_2\in\mathbb{B}(x)} \left|L^l_{\gamma_1,\gamma_2}u(x) - L^l_{\gamma_1,\gamma_2}w(x)\right| \\ &= \sup_{\gamma_1\in\mathbb{A}(x)} \sup_{\gamma_2\in\mathbb{B}(x)} \left|\int_{B(x)}\int_{A(x)} \left[L^l u(x,a,b) - L^l w(x,a,b)\right]\gamma_1(da)\gamma_2(db)\right| \\ &\leq \lambda \|u-w\|_W W(x)\end{aligned}$$

para cada $x \in X$. La ultima desigualdad implica que

$$\|Tu - Tw\| \leq \lambda \|u-w\|_W .$$

Por lo tanto, T es operador de contracción de $L_W(X)$ en sí mismo con módulo λ. ■

Lema 2.6.7 *Suponga que D1-D4 se satisfacen. Sea $h^* \in L_W(X)$ como en el Lema 2.6.6. Entonces $\nu(h^*) \leq 0$.*

Demostración. Sea $\varepsilon > 0$. Para cada $x \in X$, por el Teorema 2.5.13 (d), existe $\varphi_1^{*,\varepsilon} \in \Phi_1$ tal que

$$h^*(x) = Th^*(x) \leq \inf_{\gamma_2\in\mathbb{B}(x)} L^l_{\varphi_1^{*,\varepsilon},\gamma_2} h^*(x) + \varepsilon.$$

Pero $L^l h^* \leq Lh^*$, por lo cual, para cada $\gamma_2 \in \mathbb{B}(x)$

$$\begin{aligned} h^*(x) &\leq R_{\varphi_1^{*,\varepsilon},\gamma_2}(x) - \rho^l \tau_{\varphi_1^{*,\varepsilon},\gamma_2}(x) \\ &\quad + Q_{\varphi_1^{*,\varepsilon},\gamma_2} h^*(x) - \nu(h^*) S_{\varphi_1^{*,\varepsilon},\gamma_2}(x) + \varepsilon. \end{aligned}$$

Se sigue, para cada $\varphi_2 \in \Phi_2$ que

$$h^*(x) \leq R_{\varphi_1^{*,\varepsilon},\varphi_2}(x) - \rho^l \tau_{\varphi_1^{*,\varepsilon},\varphi_2}(x) \tag{2.33}$$

$$\qquad + Q_{\varphi_1^{*,\varepsilon},\varphi_2} h^*(x) - \nu(h^*) S_{\varphi_1^{*,\varepsilon},\varphi_2}(x) + \varepsilon. \tag{2.34}$$

Integrando (2.33) respecto a $\mu_{\varphi_1^{*,\varepsilon},\varphi_2}(\cdot)$ se tiene que

$$\begin{aligned} \mu_{\varphi_1^{*,\varepsilon},\varphi_2}(h^*) &\leq \mu_{\varphi_1^{*,\varepsilon},\varphi_2}(R_{\varphi_1^{*,\varepsilon},\varphi_2}) - \rho^l \mu_{\varphi_1^{*,\varepsilon},\varphi_2}(\tau_{\varphi_1^{*,\varepsilon},\varphi_2}) \\ &\quad + \mu_{\varphi_1^{*,\varepsilon},\varphi_2}(h^*) - \nu(h^*) \mu_{\varphi_1^{*,\varepsilon},\varphi_2}(S_{\varphi_1^{*,\varepsilon},\varphi_2}) + \varepsilon. \end{aligned}$$

Entonces

$$\rho^l \leq \rho_{\varphi_1^{*,\varepsilon},\varphi_2} + \frac{\varepsilon - \nu(h^*) \mu_{\varphi_1^{*,\varepsilon},\varphi_2}(S_{\varphi_1^{*,\varepsilon},\varphi_2})}{\mu_{\varphi_1^{*,\varepsilon},\varphi_2}(\tau_{\varphi_1^{*,\varepsilon},\varphi_2})}.$$

Esta desigualdad es válida para cada $\varepsilon > 0$.

Ahora demostraremos que $\nu(h^*) \leq 0$. Supongamos que $\nu(h^*) > 0$ y sea ε tal que

$$0 < \varepsilon < \nu(h^*)\theta.$$

Por el Corolario 2.6.3 tenemos que

$$\mu_{\varphi_1^{*,\varepsilon},\varphi_2}(W) \leq \frac{\nu(W)}{(1-\lambda)\nu(X)}.$$

Por D2, resulta

$$\mu_{\varphi_1^{*,\varepsilon},\varphi_2}(\tau_{\varphi_1^{*,\varepsilon},\varphi_2}) \leq c \mu_{\varphi_1^{*,\varepsilon},\varphi_2}(W),$$

y por el Lema 2.6.2

$$\mu_{\varphi_1^{*,\varepsilon},\varphi_2}(S_{\varphi_1^{*,\varepsilon},\varphi_2}) \geq \theta > 0.$$

Entonces se tiene que

$$\rho^l \leq \rho_{\varphi_1^{*,\varepsilon},\varphi_2} + \frac{\left[\varepsilon - \nu(h^*)\mu_{\varphi_1^{*,\varepsilon},\varphi_2}(S_{\varphi_1^{*,\varepsilon},\varphi_2})\right](1-\lambda)\nu(X)}{\nu(W)c} \quad (2.35)$$

$$\leq \rho_{\varphi_1^{*,\varepsilon},\varphi_2} + \frac{\left[\varepsilon - \nu(h^*)\theta\right](1-\lambda)\nu(X)}{\nu(W)c}. \quad (2.36)$$

Luego, aplicando ínfimo sobre φ_2 en ambos lados de (2.35) resulta la desigualdad

$$\begin{aligned}\rho^l &\leq \inf_{\varphi_2\in\Phi_2} \rho_{\varphi_1^{*,\varepsilon},\varphi_2} + \frac{\left[\varepsilon - \nu(h^*)\theta\right](1-\lambda)\nu(X)}{\nu(W)} \\ &\leq \rho^l + \frac{\left[\varepsilon - \nu(h^*)\theta\right](1-\lambda)\nu(X)}{\nu(W)c}.\end{aligned}$$

Esto es una contradicción puesto que $\varepsilon - \nu(h^*)\theta < 0$, por lo tanto $\nu(h^*) \leq 0$. ■

Lema 2.6.8 *Suponga que D1-D4 se satisfacen. Entonces* $\rho^* := \rho^l = \rho^u$.

Demostración. Sea $u \in L_W(X)$. Las funciones Qu, $-\nu(h^*)S$, R y τ son s.c.i. en $\mathbb{K}$ por la definición de S, la condición D4 y el Lema 1.5.13. Entonces $Lh^* = R - \rho^l\tau + Qh^* - \nu(h^*)S$ es s.c.i. en $\mathbb{K}$. Por lo cual

$$L^l h^* = Lh^* = R - \rho^l\tau + Qh^* - \nu(h^*)S.$$

Por el Teorema 2.5.13 existe $\varphi_2^* \in \Phi_2$ tal que

$$\begin{aligned}h^*(x) &= Th^*(x) = \sup_{\gamma_1\in\mathbb{A}(x)} \inf_{\gamma_2\in\mathbb{B}(x)} L_{\gamma_1,\gamma_2}h^*(x) \\ &= \sup_{\gamma_1\in\mathbb{A}(x)} L_{\gamma_1,\varphi_2^*}h^*(x) \quad \forall x \in X.\end{aligned}$$

Entonces, para cada $\varphi_1 \in \Phi_1$

$$\begin{aligned}h^*(x) &\geq R_{\varphi_1,\varphi_2^*}(x) - \rho^l\tau_{\varphi_1,\varphi_2^*}(x) + Q_{\varphi_1,\varphi_2^*}h^*(x) \\ &\quad -\nu(h^*)S_{\varphi_1,\varphi_2^*}(x) \quad \forall x \in X.\end{aligned}$$

Integrando respecto a $\mu_{\varphi_1,\varphi_2^*}(\cdot)$ obtenemos

$$\begin{aligned}\rho^l &\geq \frac{\mu_{\varphi_1,\varphi_2^*}(R_{\varphi_1,\varphi_2^*})}{\mu_{\varphi_1,\varphi_2^*}(\tau_{\varphi_1,\varphi_2^*})} - \nu(h^*)\frac{\mu_{\varphi_1,\varphi_2^*}(S_{\varphi_1,\varphi_2^*})}{\mu_{\varphi_1,\varphi_2^*}(\tau_{\varphi_1,\varphi_2^*})} \\ &= \rho_{\varphi_1,\varphi_2^*} - \nu(h^*)\frac{\mu_{\varphi_1,\varphi_2^*}(S_{\varphi_1,\varphi_2^*})}{\mu_{\varphi_1,\varphi_2^*}(\tau_{\varphi_1,\varphi_2^*})} \\ &\geq \rho_{\varphi_1,\varphi_2^*}.\end{aligned}$$

Calculando supremo sobre Φ_1 a ambos lados de la desigualdad, tenemos que

$$\rho^l \geq \sup_{\varphi_1\in\Phi_1} \rho_{\varphi_1,\varphi_2^*} \geq \inf_{\varphi_2\in\Phi_2} \sup_{\varphi_1\in\Phi_1} \rho_{\varphi_1,\varphi_2} = \rho^u$$

Por lo tanto, $\rho^l = \rho^u$. ■

Lema 2.6.9 *Suponga que D1-D4 se satisfacen. Sea $h^* \in L_W(X)$ como en el Lema 2.6.6. Entonces $\nu(h^*) = 0$.*

Demostración. Procediendo como en la demostración del lema anterior, existe $\varphi_2^* \in \Phi_2$ tal que

$$\begin{aligned}h^*(x) &\geq R_{\varphi_1,\varphi_2^*}(x) - \rho^*\tau_{\varphi_1,\varphi_2^*}(x) \\ &+ Q_{\varphi_1,\varphi_2^*}h^*(x) - \nu(h^*)S_{\varphi_1,\varphi_2^*}(x) \quad \forall x \in X,\ \varphi_1 \in \Phi_1.\end{aligned}$$

Integrando respecto a $\mu_{\varphi_1,\varphi_2^*}(\cdot)$ obtenemos

$$\begin{aligned}\nu(h^*)\mu_{\varphi_1,\varphi_2^*}(S_{\varphi_1,\varphi_2^*}) &\geq \mu_{\varphi_1,\varphi_2^*}(\tau_{\varphi_1,\varphi_2^*})\left[\rho_{\varphi_1,\varphi_2^*} - \rho^*\right] \\ &= \mu_{\varphi_1,\varphi_2^*}(\tau_{\varphi_1,\varphi_2^*})\left[\rho_{\varphi_1,\varphi_2^*} - \inf_{\varphi_2\in\Phi_2}\sup_{\psi_1\in\Phi_1}\rho_{\psi_1,\varphi_2}\right] \\ &\geq \mu_{\varphi_1,\varphi_2^*}(\tau_{\varphi_1,\varphi_2^*})\left[\rho_{\varphi_1,\varphi_2^*} - \sup_{\psi_1\in\Phi_1}\rho_{\psi_1,\varphi_2^*}\right].\end{aligned}$$

Entonces

$$\nu(h^*)\frac{\mu_{\varphi_1,\varphi_2^*}(S_{\varphi_1,\varphi_2^*})}{\mu_{\varphi_1,\varphi_2^*}(\tau_{\varphi_1,\varphi_2^*})} \geq \rho_{\varphi_1,\varphi_2^*} - \sup_{\psi_1\in\Phi_1}\rho_{\psi_1,\varphi_2^*} \quad \forall\varphi_1 \in \Phi_1$$

Calculando el supremo sobre Φ_1 en ambos lados de la desigualdad obtenemos

$$\sup_{\varphi_1 \in \Phi_1} \left[\nu(h^*) \frac{\mu_{\varphi_1,\varphi_2^*}(S_{\varphi_1,\varphi_2^*})}{\mu_{\varphi_1,\varphi_2^*}(\tau_{\varphi_1,\varphi_2^*})} \right] \geq 0.$$

Por (2.29) tenemos $\sup_{\varphi_1 \in \Phi_1} \frac{\mu_{\varphi_1,\varphi_2^*}(S_{\varphi_1,\varphi_2^*})}{\mu_{\varphi_1,\varphi_2^*}(\tau_{\varphi_1,\varphi_2^*})} \geq 0$, lo cual implica $\nu(h^*) \geq 0$. Por lo tanto, por el Lema 2.6.7, $\nu(h^*) = 0$. ∎

A continuación se demuestra el Teorema 2.6.1.

Demostración del Teorema 2.6.1. Por los Lemas 2.6.6 y 2.6.9 existe $h^* = Th^* \in L_W(X)$ tal que $\nu(h^*) = 0$, y por el Teorema de selección minimax (Teorema 2.5.13), se satisface la ecuación

$$\begin{aligned} h^*(x) &= \sup_{\gamma_1 \in \mathbb{A}(x)} \inf_{\gamma_2 \in \mathbb{B}(x)} \left[R_{\gamma_1,\gamma_2}(x) - \rho^* \tau_{\gamma_1,\gamma_2}(x) + Q_{\gamma_1,\gamma_2} h^*(x) \right] \\ &= \inf_{\gamma_2 \in \mathbb{B}(x)} \sup_{\gamma_1 \in \mathbb{A}(x)} \left[R_{\gamma_1,\gamma_2}(x) - \rho^* \tau_{\gamma_1,\gamma_2}(x) + Q_{\gamma_1,\gamma_2} h^*(x) \right] \quad \forall x \in X. \end{aligned}$$

Además, existe $\varphi_2^* \in \Phi_2$ tal que

$$h^*(x) = \sup_{\gamma_1 \in \mathbb{A}(x)} \left[R_{\gamma_1,\varphi_2^*}(x) - \rho^* \tau_{\gamma_1,\varphi_2^*}(x) + Q_{\gamma_1,\varphi_2^*} h^*(x) \right]$$

y para cada $\varepsilon > 0$, existe $\varphi_1^{*,\varepsilon} \in \Phi_1$ tal que

$$h^*(x) \leq \inf_{\gamma_2 \in \mathbb{B}(x)} \left[R_{\varphi_1^{*,\varepsilon},\gamma_2}(x) - \rho^* \tau_{\varphi_1^{*,\varepsilon},\gamma_2}(x) + Q_{\varphi_1^{*,\varepsilon},\gamma_2} h^*(x) \right] + \varepsilon.$$

Por otra parte, por la Observación 2.5.2 se cumple

$$\lim_{n\to\infty} E_x^{\pi_1,\pi_2} \frac{1}{n} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k) \geq \sigma > 0,$$

y por el Lema 2.5.5 se tiene que

$$\lim_{n\to\infty} \frac{E_x^{\pi_1,\pi_2} h^*(x_n)}{E_x^{\pi_1,\pi_2} \sum_{k=0}^{n-1} \tau(x_k, a_k, b_k)} = 0.$$

Entonces, por el Teorema 2.4.3, la estrategia $\varphi_1^{*,\varepsilon} \in \Phi_1$ es ε-óptima para el primer jugador, $\varphi_2^* \in \Phi_2$ es óptima para el segundo jugador y ρ^* es el valor del juego. ∎

Observación 2.6.10 *El juego markoviano (JM) es un caso particular del JSM al hacer*

$$G(t|x,a,b) = I_{[1,\infty)}(t) \quad \forall (x,a,b) \in \mathbb{K},$$

es decir, el tiempo de permanencia en los estados es siempre de una unidad. De esta forma se tiene que $\tau \equiv 1$ *y* $T_n = n$ *para cada* $n \in \mathbb{N}$.

2.7. Ejemplo

En esta sección se presenta un ejemplo de un juego markoviano que satisface las condiciones del Teorema 2.6.1. En este ejemplo el kérnel de transición no es fuertemente continuo.

Consideremos un proceso de inventarios que evoluciona de acuerdo a la ecuación en diferencias

$$x_{n+1} = (x_n + a_n - Z_{n+1})^+, \quad n \in \mathbb{N}_0,$$

donde

- $x_n \in X = \mathbb{R}_+$ es el inventario al inicio del periodo n;
- $a_n \in A(x_n) \equiv A = [0, a^*]$ es la cantidad de producto que ordena el controlador (el primer jugador) al inicio del periodo n;
- $Z_n \geq 0$ representa la demanda durante el n-ésimo periodo.

En este caso el segundo jugador será la "naturaleza", que en cada periodo n toma del conjunto

$$B(x_n) \equiv B = \{G_1, \ldots, G_N\} \subset \mathbb{P}([0, +\infty)),$$

una distribución G_i para Z_n. Si E_i representa el operador esperanza respecto a G_i, entonces

$$E_i Z_n = z^* \; \forall i = 1, \ldots N, \; n \in \mathbb{N}.$$

El kérnel de transición para este modelo queda definido de la siguiente manera:

$$Q(D|x,a,i) = \int_0^\infty I_D(x + a - z)^+ G_i(dz) \quad \forall D \in \mathcal{B}(X).$$

Supondremos que el juego satisface las siguientes condiciones

A1 $z^* > a^*$

A2 $\max_{1\le i\le N} G_i(z) < 1$ para todo $z \in [0,\infty)$

La función de pago por etapa está dada por

$$R(x,a,i) = \kappa_1 I_{\{a\neq 0\}} + \kappa_2 a + \kappa_3 \int_0^\infty (x+a-z)^+ G_i(dz),$$

para cada $(x,a,i)\in\mathbb{K}$, donde $\kappa_1,\kappa_1,\kappa_1\in\mathbb{R}_+$ son constantes.

Al ser un juego markoviano la condición D1 se satisface trivialmente. A continuación veremos que se satisfacen D2, D3 y D4.

Observemos lo siguiente. Sea

$$\Psi_i(p) := \int_0^\infty e^{p(a^*-z)} G_i(dz), \quad p\ge 0,$$

la función generadora de momentos de a^*-Z donde Z es una v.a. con distribución G_i. Observe que $\Psi_i(0)=1$, y por A1,

$$\frac{d\Psi_i}{dp}(0) = a^* - z^* < 0, \quad \forall i = 1,\ldots,N,$$

Dado que Ψ_i es continua en $\mathbb{R}_+$, existe $\beta_i>0$ tal que

$$\Psi_i(y) < 1 \ \ \forall y\in(0,\beta_i).$$

Si $r := \min\left\{\frac{\beta_1}{2},\ldots,\frac{\beta_N}{2}\right\}$ se tiene que

$$\Psi_i(r) =: \alpha_i < 1 \ \ \forall i = 1,\ldots,N.$$

Definamos

$$V(x) := e^{rx} \ge 1 \quad \text{y} \quad \alpha := \max\{\alpha_1,\ldots,\alpha_N\} \ \ \forall x\in X.$$

Ahora estamos preparados para demostrar que se satisface D3. Sean $D\in\mathcal{B}(X)$ y $(x,a,i)\in\mathbb{K}$. Por ser G_i distribución de probabilidad, existe $x^*\in X$ tal que

$$\max_{1\le i\le N}[1-G_i(x)] < \frac{1-\alpha}{2}, \quad \forall x > x^*, \tag{2.37}$$

y definamos

$$\delta := \min_{1\leq i\leq N}[1 - G_i(x^* + a^*)].$$

donde $\delta > 0$ por A2. Entonces

$$\begin{aligned} Q(D|x,a,i) &= \int_0^\infty I_D((x+a-z)^+)G_i(dz) \\ &= \int_0^{x+a} I_D(x+a-z)G_i(dz) + \int_{x+a}^\infty I_D(0)G_i(dz) \\ &\geq \int_{x+a}^\infty I_D(0)G_i(dz) = I_D(0)[1-G_i(x+a)] \\ &\geq I_{[0,x^*]}(x)I_D(0)[1-G_i(x+a)]. \end{aligned}$$

Ahora, si $x \in (x^*, \infty)$ se tiene que

$$Q(D|x,a,i) \geq I_{[0,x^*]}(x)I_D(0)[1-G_i(x+a)] = 0.$$

Por otra parte, si $x \in [0, x^*]$ se sigue que

$$\begin{aligned} Q(D|x,a,i) &\geq I_{[0,x^*]}(x)I_D(0)[1-G_i(x+a)] \\ &= I_D(0)[1-G_i(x+a)] \\ &\geq I_D(0)[1-G_i(x^*+a^*)] \\ &\geq \delta I_D(0). \end{aligned}$$

Por lo tanto

$$Q(D|x,a,i) \geq \delta I_{[0,x^*]}(x)I_D(0) \qquad (2.38)$$

$$(2.39)$$

$$= \phi(x,a,i)\nu(D),$$

donde $\phi(x,a,i) := \delta I_{[0,x^*]}(x)$ y $\nu(D) := I_D(0)$. Con esto hemos demostrado D3(b), y por la definición de $\nu(D)$ se satisface D3(a). Además, dado que $\nu(\delta I_{[0,x^*]}) = \delta I_{[0,x^*]}(0) = \delta$, D3(d) también se cumple.

Por otra parte

$$\begin{aligned}
\int_X V(y)Q(dy|x,a,i) &= \int_0^\infty V((x+a-z)^+)G_i(dz)\\
&= \int_0^{x+a} V(x+a-z)G_i(dz) + \int_{x+a}^\infty V(0)G_i(dz)\\
&= \int_0^{x+a} e^{r(x+a-z)}G_i(dz) + 1 - G_i(x+a)\\
&\leq V(x)\int_0^\infty e^{r(a-z)}G_i(dz) + [1-G_i(x)]\\
&\leq V(x)\int_0^\infty e^{r(a^*-z)}G_i(dz) + \frac{1-\alpha}{2}\\
&\leq \alpha V(x) + I_{[0,x^*]}(x)\\
&\leq \frac{1+\alpha}{2}V(x) + I_{[0,x^*]}(x).
\end{aligned}$$

Por lo tanto

$$\int_X V(y)Q(dy|x,a,i) \leq \frac{1+\alpha}{2}V(x) + I_{[0,x^*]}(x). \tag{2.40}$$

Definamos ahora

$$W(x) := V(x) + \frac{1}{\delta}, \quad \eta := \frac{1+\alpha}{2}.$$

Aplicando (2.40)

$$\begin{aligned}
\int_X W(y)Q(dy|x,a,i) &= \int_X V(y)Q(dy|x,a,i) + \frac{1}{\delta}\\
&\leq \left[\eta V(x) + \frac{1}{\delta}\right] + I_{[0,x^*]}(x)\\
&\leq \left[\eta V(x) + \frac{1}{\delta}\right]\\
&\quad +\delta I_{[0,x^*]}(x)\left[\frac{1}{\delta} + \int_X V(y)\nu(dy)\right].
\end{aligned}$$

Por otro lado, dado que $V(x) \geq 1$, se sigue que

$$\frac{1}{\delta}\left(\frac{1-\eta}{1+\frac{1}{\delta}}\right)V(x) \geq \frac{1}{\delta}\left(\frac{1-\eta}{1+\frac{1}{\delta}}\right). \tag{2.41}$$

Mediante cálculos algebraicos simples, (2.41) es equivalente a la desigualdad

$$\begin{aligned}\left(\frac{\eta+\frac{1}{\delta}}{1+\frac{1}{\delta}}\right)\left(V(x)+\frac{1}{\delta}\right) &= \frac{\eta+\frac{1}{\delta}}{1+\frac{1}{\delta}}W(x)\\ &\geq \eta V(x)+\frac{1}{\delta}.\end{aligned}$$

Entonces

$$\begin{aligned}\int_X W(y)Q(dy|x,a,i) &\leq \frac{\eta+\frac{1}{\delta}}{1+\frac{1}{\delta}}W(x)+\delta I_{[0,x^*]}(x)\int_X W(y)\nu(dy)\\ &= \lambda W(x)+\phi(x,a,i)\int_X W(y)\nu(dy),\end{aligned}$$

donde $\lambda := \frac{\eta+\frac{1}{\delta}}{1+\frac{1}{\delta}}$. Observemos que $\eta \in (0,1)$ ya que $\alpha \in (0,1)$, por lo cual $\lambda \in (0,1)$. De este modo hemos demostrado que se satisface D3(c). Por lo tanto, el sistema cumple D3.

Para verificar D2 basta observar lo siguiente

$$\begin{aligned}|R(x,a,i)| &= \left|\kappa_1 I_{\{a\neq 0\}}+\kappa_2 a+\kappa_3\int_0^\infty (x+a-z)^+G_i(dz)\right|\\ &\leq \kappa_1+\kappa_2 a^*+\kappa_3\int_0^\infty (x+a-z)^+G_i(dz)\\ &= \kappa_1+\kappa_2 a^*+\kappa_3\int_0^{x+a}(x+a-z)G_i(dz)\\ &\leq \kappa_1+\kappa_2 a^*+\kappa_3\int_0^\infty (x+a)G_i(dz)-\kappa_3\int_0^{x+a} zG_i(dz)\\ &\leq \kappa_1+\kappa_2 a^*+\kappa_3\left(x+a^*\right)\\ &\leq \left[\kappa_3+\left(\kappa_2+\kappa_3\right)a^*+\frac{\kappa_3}{r}\right]V(x)\\ &\leq \left[\kappa_1+\left(\kappa_2+\kappa_3\right)a^*+\frac{\kappa_3}{r}\right]W(x).\end{aligned}$$

Por lo tanto, el sistema cumple D2.

Ahora demostraremos que se satisface D4. Dado que B es finito, para verificar que este conjunto de condiciones se satisface, podemos suponer sin pérdida de generalidad que $B = \{G_1\}$. Trivialmente se satisfacen la partes (a), (b) y (c) de D4, ya que para cada $(x, a, i) \in \mathbb{K}$, $\tau(x, a, i) = 1$ y además $A(x) = [0, a^*]$ y $B(x) = \{G_1, \ldots, G_N\}$.

Demostremos que Q es débilmente continuo y R es semicontinua sobre $\mathbb{K}$. Sea $u \in C_b(X)$ y $\{(y_k, c_k, 1)\}_{n\in\mathbb{N}}$ una sucesión convergente en $\mathbb{K}$ tal que $(y_k, c_k) \to (y, c)$. Observe que que existe $M > 0$ tal que $|y_k| \leq M$ para cada $k \in \mathbb{N}$. El mapeo $z \mapsto (y + c - z)^+$ es continuo y dominado por $M + a^*$, entonces por el teorema de convergencia dominada tenemos

$$\lim_{k\to\infty} \int_0^\infty (y_k + c_k - z)^+ G_1(dz) = \int_0^\infty (y + c - z)^+ G_1(dz),$$

y también

$$\begin{aligned}
\lim_{k\to\infty} \int_X u(z) Q(dz|y_k, c_k, 1) &= \lim_{k\to\infty} \int_0^\infty u((y_k + c_k - z)^+) G_1(dz) \\
&= \int_0^\infty u((y + c - z)^+) G_1(dz) \\
&= \int_X u(z) Q(dz|y, c, 1).
\end{aligned}$$

Además, como $I_{\{c\neq 0\}}$ es una función s.c.i., se tiene que

$$\begin{aligned}
\liminf_{k\to\infty} R(y_k, c_k, 1) &= \liminf_{k\to\infty} \left[\kappa_1 I_{\{c_k\neq 0\}} + \kappa_2 c_k + \kappa_3 \int_0^\infty (y_k + c_k - z)^+ G_1(dz) \right] \\
&= \liminf_{k\to\infty} \kappa_1 I_{\{c_k\neq 0\}} + \kappa_2 c + \kappa_3 \int_0^\infty (y + c - z)^+ G_1(dz) \\
&\geq \kappa_1 I_{\{c\neq 0\}} + \kappa_2 c + \kappa_3 \int_0^\infty (y + c - z)^+ G_1(dz) \\
&= R(y, c, 1).
\end{aligned}$$

Por lo tanto Q es débilmente continuo y R es s.c.i. en $\mathbb{K}$.

Veamos ahora que QW es continua en $\mathbb{K}$. Para cada $z \in [0, \infty)$ y $k \in \mathbb{N}$,

$$\begin{aligned} W((y_k + c_k - z)^+) &= V((y_k + c_k - z)^+) + \frac{1}{\delta} \\ &\leq V(M + a^*) + \frac{1}{\delta}. \end{aligned}$$

Entonces, por la continuidad de W y el teorema de convergencia dominada, se tiene que

$$\begin{aligned} \lim_{k \to \infty} \int_X W(z) Q(dz|y_k, c_k, 1) &= \lim_{k \to \infty} \int_0^\infty W((y_k + c_k - z)^+) G_1(dz) \\ &= \int_0^\infty W((y + c - z)^+) G_1(dz) \\ &= \int_X W(z) Q(dz|y, c, 1). \end{aligned}$$

Con esto hemos demostrado D4(d) y D4(e). Por lo tanto, el sistema satisface D4.

Observación 2.7.1 *Si no se satisface A2, el kérnel Q podría no ser fuertemente continuo*

Por ejemplo, tomemos

$$G_1 = I_{[z^*, \infty)}$$

donde $z^* > 0$, $c \in (0, a^*)$ y $c_k = \text{mín} \left\{ c + \frac{1}{k}, a^* \right\}$, $k \in \mathbb{N}$. Note que $c_k \to c$ y

$$\max_{1 \leq i \leq N} G_i(z^*) = G_1(z^*) = 1,$$

es decir, la condición A2 no se satisface.

Ahora consideremos la función $u = I_{[0,c]}$ y observe que

$$\begin{aligned} \int_X u(z) Q(dz|z^*, c_k, 1) &= \int_0^\infty u((z^* + c_k - z)^+) G_1(dz) \\ &= u(z^* + c_k - z^*) \\ &= I_{[0,c]}(c_k) = 0, \end{aligned}$$

y

$$\begin{aligned}
\int_X u(z)Q(dz|z^*,c,1) &= \int_0^\infty u((z^*+c-z)^+)G_1(dz) \\
&= I_{[0,c]}(c) = 1.
\end{aligned}$$

Por lo tanto

$$\lim_{k\to\infty} Qu(z^*,c_k,1) = 0 \neq 1 = Qu(z^*,c,1),$$

i.e., la función $Qu(z^*,\cdot,1)$ no es continua sobre $A(z^*)$. Para mas detalles acerca de este ejemplo se puede consultar [6] y [13].

2.8. Observaciones finales

Un problema que no abordamos explícitamente en este trabajo fue el problema de control óptimo semi-markoviano (PCOSM) en costo promedio. Esto se debe a que se puede deducir un teorema análogo al Teorema 1.6.1 para este modelo sin hacer grandes cambios a lo ya hecho en los capítulos 1 y 2.

Otra observación que es importante remarcar es que Jaskiewicz en [13] hace uso de una serie de condiciones distintas a las que usamos en este capítulo para resolver el mismo problema: la existencia de estrategias óptimas (ε-óptimas) para el juego semi-markoviano en pago promedio. La diferencia entre las condiciones usadas en [13] y las que se vieron en la sección 2.5 es que Jaskiewicz usa la condición que se muestra a continuación.

Condición 2.8.1 (D5) *Existe un conjunto $C \in \mathcal{B}(X)$, un conjunto abierto $\tilde{C} \subset C$, medida no trivial $\nu \in \mathbb{P}(C)$, $\lambda, \delta \in (0,1)$, $\eta, \kappa > 0$ y $\beta < 1$ tales que, para cada $D \in \mathcal{B}(C)$ se cumple lo siguiente:*

(a) $QW(x,a,b) \leq \lambda W(x) + \eta I_C(x) \quad \forall (x,a,b) \in \mathbb{K}$,

(b) $\nu(\tilde{C}) > 0$,

(c) $Q(D|x,a,b) \geq \delta\nu(D)$ *para cada $x \in C$, $a \in A(x)$ y $b \in B(x)$,*

(d) $\sup_{x \in C} W(x) < \infty$,

(e) $H(\kappa|x, a, b) \leq \beta$ *para cada* $x \in C$, $a \in A(x)$ *y* $b \in B(x)$.

Si se cumple D5 entonces se cumplen las condiciones D3 y D1. La condición D3 es menos restrictiva que la condición D5 en [13]. Además, el enfoque de punto fijo desarrollado en la presente tesis es más directo y simple que el de [13].

Apéndice A

Funciones Semicontinuas

Las demostraciones de los resultados que aquí se presentan sin demostración se pueden encontrar en el apéndice 2 de [3].

Definición A.0.2 *Sea (Y, d) un espacio métrico. Decimos que la función $f : Y \to \bar{\mathbb{R}}$ es semicontinua inferiormente (s.c.i.) sobre Y si para cada $a \in \bar{\mathbb{R}}$, $\{y \in Y : f(x) > a\}$ es abierto. Y es semicontinua superiormente (s.c.s.) si $\{y \in Y : f(x) < a\}$ es abierto.*

Observación A.0.3 *La función v es continua si y solo si es v s.c.i. y s.c.s.*

Observación A.0.4 *f es s.c.i. si y solo si $-f$ es s.c.s..*

Teorema A.0.5 *La función $f : Y \to \bar{\mathbb{R}}$ es s.c.s si y sólo si para cada sucesión $(y_n)_{n\in\mathbb{N}}$ que converge a $y \in Y$*

$$f(y) \leq \liminf_{n\to\infty} f(y_n)$$

y f es s.c.s. si y solo si $f(y) \geq \limsup_{n\to\infty} f(y_n)$.

Teorema A.0.6 *Si $u : Y \to \mathbb{R}$ es semicontinua inferiormente y acotada inferiormente, entonces existe una sucesión $(u_n)_{n\in\mathbb{N}}$ de funciones en $C_b(Y)$ tal que $u_n \uparrow u$ (si u es s.c.s. y acotada superiormente, entonces existe una sucesión $(u_n)_{n\in\mathbb{N}}$ en $C_b(Y)$ tal que $u_n \downarrow u$).*

Demostración. Sin pérdida de generalidad supongamos que $u \geq 0$. Definamos $d' : Y \times Y \to \mathbb{R}$ como

$$d'(x,y) := \frac{d(x,y)}{1 + d(x,y)},$$

La función d' es una métrica en Y que es acotada por 1. Más aún, d' es una métrica en Y equivalente a d, i.e., podemos usar d' en lugar de d en lo que a topología se refiere.

Ahora, para cada $n \in \mathbb{N}$ fijo, definimos la función $u_n : Y \to \mathbb{R}$ como

$$u_n(y) := \operatorname*{\acute{i}nf}_{x \in Y} \left[u(x) + nd'(x,y) \right].$$

Observemos que

$$\begin{aligned} 0 &\leq u_n(y) = \operatorname*{\acute{i}nf}_{x \in Y} \left[u(x) + nd'(x,y) \right] \\ &\leq \operatorname*{\acute{i}nf}_{x \in Y} \left[u(x) + n \right] \\ &= \operatorname*{\acute{i}nf}_{x \in Y} u(x) + n \quad \forall y \in Y, \end{aligned}$$

i.e., u_n es acotada. Por otra parte

$$u_n(y) \leq u(y) + nd'(y,y) = u(y) \quad \forall y \in Y,$$

entonces $u_n \leq u$. Además, por definición $u_{n-1} \leq u_n$, i.e., la sucesión $(u_n)_{n \in \mathbb{N}}$ es monótona. Por lo tanto, existe $h : Y \to \mathbb{R}$ tal que $u_n \uparrow h \leq u$.

Veamos ahora que u_n es continua para cada $n \in \mathbb{N}$. Sean $x, y \in Y$ arbitrarios fijos, entonces

$$u(z) + nd'(z,y) \leq u(z) + nd'(z,x) + nd'(x,y) \quad \forall z \in Y,$$

por lo cual

$$\begin{aligned} u_n(y) &= \operatorname*{\acute{i}nf}_{z \in Y} \left[u(z) + nd'(z,y) \right] \\ &\leq \operatorname*{\acute{i}nf}_{z \in Y} \left[u(z) + nd'(z,x) \right] + nd'(x,y) \\ &= u_n(x) + nd'(x,y). \end{aligned}$$

Por simetría

$$|u_n(x) - u_n(y)| \leq nd'(x,y)$$

Como d' es una métrica en Y equivalente a d, se sigue que u_n es continua en Y.

Demostremos que $u_n \to u$. Sea $\varepsilon > 0$. Por la definición de ínfimo, para cada $n \in \mathbb{N}$ y $y \in Y$ fijos, existe $z_n \in Y$ tal que

$$u_n(y) + \varepsilon > u(z_n) + nd'(y, z_n) \geq nd'(y, z_n). \tag{A.1}$$

Pero $u_n(y) + \varepsilon \leq u(y) + \varepsilon$, entonces

$$0 = \lim_{n\to\infty} \frac{u(y) + \varepsilon}{n} \geq \limsup_{n\to\infty} d'(y, z_n) \geq 0,$$

por lo tanto $\lim_{n\to\infty} d'(y, z_n) = 0$, i.e., $z_n \to y$. Dado que u es s.c.i.,

$$\liminf_{n\to\infty} u(z_n) = \sup_{n\in\mathbb{N}} \inf_{k\geq n} u(z_n) \geq u(y).$$

Entonces, existe $N \in \mathbb{N}$ tal que, para cada $n > N$, $u(z_n) > u(y) - \varepsilon$. Si además aplicamos (A.1), se sigue que

$$\begin{aligned} u_n(y) &> u(z_n) + nd'(y, z_n) - \varepsilon \geq u(z_n) - \varepsilon \\ &> u(y) - 2\varepsilon. \end{aligned}$$

Por lo tanto $0 \leq u_n \uparrow u$. ∎

Teorema A.0.7 *Sea $(u_n)_{n\in\mathbb{N}}$ una sucesión de funciones s.c.i. sobre (Y, d).*
(a) *si u_n es s.c.i. para cada $n \in \mathbb{N}$, entonces $\sup_{n\in\mathbb{N}} u_n$ es s.c.i.;*
(b) *si $u_n \uparrow u$, entonces u es s.c.i..*

Definición A.0.8 *Sea (Y, d) un espacio métrico. Para la función $v : Y \to \mathbb{R}$ definamos*

$$\begin{aligned} v^l(y) := & \sup_{r>0} \inf_{t\in B_r(y)} v(t) \\ = & \lim_{r\to 0^+} \inf_{t\in B_r(y)} v(t), \end{aligned} \tag{A.2}$$

*donde $B_r(y) := \{t \in Y : d(y, t) < r\}$ para cada $r > 0$. A v^l le llamaremos la **envolvente inferior** de v.*

Lema A.0.9 *La función definida en (A.2) satisface lo siguiente*

(a) *v^l es semicontinua inferiormente.*

(b) *si $w \geq v$ entonces $w^l \geq v^l$.*

(c) *v es s.c.i. si y solo si $v = v^l$.*

(d) *si $w \geq v$ y v es s.c.i., entonces $w^l \geq v$.*

Asi como se definió v^l, también podemos definir la **envolvente superior** de v como

$$\begin{aligned} v^u(s) := & \inf_{r>0} \sup_{t \in B_r(y)} v(t) \\ = & \lim_{r \to 0^+} \sup_{t \in B_r(y)} v(t) \end{aligned} \tag{A.3}$$

De este modo tenemos el siguiente lema

Lema A.0.10 *La función definida en (A.3) satisface lo siguiente*

(a) *v^u es semicontinua superiormente.*

(b) *si $w \leq v$ entonces $w^u \leq v^u$.*

(c) *v es s.c.s. si y solo si $v = v^u$.*

(d) *si $w \leq v$ y v es s.c.s., entonces $w^u \leq v$.*

Lema A.0.11 *Sea $w : Y \to [1, \infty)$ una función continua definida sobre el espacio métrico (Y, d). Entonces*

(a) *$u^l \in L_w(Y)$ para cada $u \in B_w(Y)$*

(b) *$\left\| u^l - v^l \right\|_w \leq \| u - v \|_w$ para cada $u, v \in B_w(Y)$*

Demostración. La parte (**a**) se deduce por (A.0.9). Para demostrar la parte (b) observemos que para cada $u, v \in B_w(Y)$, las desigualdades

$$\begin{aligned} \left| \inf_{t \in B_r(y)} u(t) - \inf_{t \in B_r(y)} v(t) \right| & \leq \sup_{t \in B_r(y)} |u(t) - v(t)| \\ & \leq \| u - v \|_w \sup_{t \in B_r(y)} w(t) \quad \forall y \in Y, \end{aligned}$$

se satisfacen para cada $r > 0$. Por (A.0.3), $w = w^u$, entonces

$$\begin{aligned} \left| u^l(y) - v^l(y) \right| &\leq \|u - v\|_w \lim_{r \to 0^+} \sup_{t \in B_r(y)} w(t) \\ &\leq \|u - v\|_w \, w(y) \quad \forall y \in Y, \end{aligned}$$

y con esto hemos demostrado (**b**). ■

Apéndice B

Correspondencias

Definición B.0.12 *Sean X y Y conjuntos no vacíos. Una correspondencia φ de X a Y es una función que asigna a cada x en X un elemento $\varphi(x)$ del conjunto potencia 2^Y, es decir, φ es una función con dominio X y contradominio 2^Y.*

Para representar una correspondencia de X a Y usaremos la notación $\varphi : X \twoheadrightarrow Y$.

Definición B.0.13 *Sea $\varphi : X \twoheadrightarrow Y$ una correspondencia. Definimos a la gráfica de φ como el conjunto*

$$Gr\varphi := \{(x, y) \in X \times Y : y \in \varphi(x)\}.$$

Definición B.0.14 *Sea X un espacio topológico, $x \in X$ y $F \subset X$. Diremos que V es una vecindad de x, si existe un abierto A tal que $x \in A \subset V$. Análogamente, S es una vecindad de F si existe un abierto O tal que $F \subset O \subset S$. A los conjuntos A y O se les llama vecindad abierta de x y F respectivamente.*

Definición B.0.15 *Sean X, Y espacios topológicos y $\varphi : X \twoheadrightarrow Y$ una correspondencia.*

(a) *Se dice que φ es **semicontinua superiormente** en $x \in X$, si para cada vecindad U de $\varphi(x)$ existe una vecindad V de x tal que*

$$\varphi(z) \subset U \quad \forall z \in V.$$

(b) *Se dice que φ es* ***semicontinua inferiormente*** *en $x \in X$, si para cada conjunto abierto U tal que $U \cap \varphi(x) \neq \varnothing$, existe una vecindad V de x, tal que*

$$U \cap \varphi(x) \neq \varnothing \quad \forall z \in V.$$

(c) *Se dice que φ es* ***continua*** *en $x \in X$, si es semicontinua superiormente e inferiormente en x.*

La correspondencia φ es ***semicontinua superiormente (inferiormente)*** *en X si lo es en cada punto $x \in X$, análogamente φ es* ***continua*** *en X si lo es en cada $x \in X$.*

Teorema B.0.16 *Sean X, Y espacios métricos y $\varphi : X \twoheadrightarrow Y$ una correspondencia. Los siguientes enunciados son equivalentes:*

(a) *φ es semicontinua inferiormente en x.*

(b) *Para cada sucesión $(x_n)_{n\in\mathbb{N}}$ que converge a x y para cada $y \in \varphi(x)$ existe una sucesión $(y_n)_{n\in\mathbb{N}}$ que converge a y y tal que*

$$y_n \in \varphi(x_n) \quad \forall n \in \mathbb{N}.$$

Demostración. La demostración se puede ver en [7, Teorema 1.2.3]. ■

Teorema B.0.17 *Sean X, Y espacios métricos y $\varphi : X \twoheadrightarrow Y$ una correspondencia. Los siguientes enunciados son equivalentes:*

(a) *φ es semicontinua superiormente en x y $\varphi(x)$ es compacto.*

(b) *Para cada sucesión $((x_n, y_n))_{n\in\mathbb{N}} \subset Gr\varphi$, tal que $(x_n)_{n\in\mathbb{N}}$ es convergente a x, se tiene que la sucesión $(y_n)_{n\in\mathbb{N}}$ tiene un punto de acumulación en $\varphi(x)$.*

Demostración. La demostración se puede ver en [7, Teorema 1.2.4] ■

Apéndice C

Procesos de Markov

Consideremos un proceso de Markov en tiempo discreto $\{x_t\}$ con valores en un espacio de Borel X y kérnel de transición P sobre X dado X. Las probabilidades de transición en n-pasos se definen recursivamente de la siguiente manera:

$$
\begin{aligned}
P^{n+1}(B|x) :=& \quad \int_X P^n(B|y)P(dy|x), \quad n \in \mathbb{N}, \\
P^0(B|x) :=& \quad I_B(x),
\end{aligned}
$$

para cada $x \in X$ y $B \in \mathcal{B}(X)$.

Teorema C.0.18 *Sea $\Omega := X^\infty$ y $\mathcal{F}$ la σ-álgebra producto correspondiente. Para cada medida de probabilidad ν en X existe una medida de probabilidad P_ν definida en el espacio medible $(\Omega, \mathcal{F})$ que satisface las siguientes propiedades para cada $n \in \mathbb{N}_0$ y $B \in \mathcal{B}(X)$:*

- $P_\nu[x_0 \in B] = \nu(B)$
- $P_\nu[x_{n+1} \in B|x_0, x_1, \ldots, x_n] = P(B|x_n)$

Demostración. Véase [3, Theorem 2.7.2]. ∎

La esperanza con respecto a la medida de probabilidad P_ν será denotada por E_ν. Si ν está concentrada en un estado $x \in X$, es decir,

$$\nu(B) = I_B(x) \;\; \forall B \in \mathcal{B}(X),$$

entonces escribiremos P_x en lugar de P_ν.

Definición C.0.19 *Sea φ una medida σ-finita en X. El proceso de Markov $\{x_t\}$ es **φ-irreducible** si para cada $x \in X$ y $B \in \mathcal{B}(X)$, con $\varphi(B) > 0$, existe un número natural n tal que $P^n(B|x) > 0$. En este caso también se dice que la medida φ es una **medida de irreducibilidad** para el proceso $\{x_t\}$.*

Observación C.0.20 *Si $\{x_t\}$ es φ-irreducible y $\varphi' \ll \varphi$, entonces también es φ'-irreducible.*

Definición C.0.21 *Una medida σ-finita ψ en X es una **medida máxima de irreducibilidad** para el proceso $\{x_t\}$ si cualquier otra medida de irreducibilidad φ es absolutamente continua respecto a ψ, es decir, $\varphi \ll \psi$.*

Teorema C.0.22 *Suponga que el proceso $\{x_t\}$ es φ-irreducible. Entonces, existe una medida máxima de irreducibilidad ψ.*

Demostración. Véase [17, Proposition 4.2.2] y [20, Proposition 2.4]. ■

Cuando el proceso $\{x_t\}$ sea φ-irreducible, usaremos la siguiente notación:

$$\mathcal{B}^+(X) := \{B \in \mathcal{B}(X) : \psi(B) > 0\},$$

donde ψ es una medida de máxima irreducibilidad.

Definición C.0.23 *El proceso $\{x_t\}$ es **Harris recurrente** si es φ-irreducible y*

$$P_x \left[\bigcap_{N=1}^{\infty} \bigcup_{k=N}^{\infty} \{x_k \in B\} \right] = 1,$$

para cada $x \in X$ y $B \in \mathcal{B}^+$.

Definición C.0.24 *Una medida σ-finita μ es una **medida invariante** para la cadena de Markov $\{x_t\}$ si*

$$\mu(B) = \int_X P(B|x)\mu(dx).$$

*Si además μ satisface $\mu(X) = 1$, entonces diremos que μ es una **medida de probabilidad invariante** para la cadena de Markov $\{x_t\}$.*

Teorema C.0.25 *Si $\{x_t\}$ es Harris recurrente, entonces existe una medida invariante μ, la cual es única excepto por constantes multiplicativas.*

Demostración. Véase [17, Theorem 10.0.1]. ■

Definición C.0.26 *La cadena de Markov $\{x_t\}$ es **Harris recurrente positiva** si es Harris recurrente y admite una única medida de probabilidad invariante μ.*

Teorema C.0.27 *La cadena de Markov $\{x_t\}$ es Harris recurrente positiva si y solo si, para cada $B \in \mathcal{B}(X)$, existe una constante a_B tal que*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} P^k(B|x) = a_B \quad \forall x \in X.$$

Demostración. Véase [10, Theorem 1.1]. ■

Observación C.0.28 *Si la cadena de Markov $\{x_t\}$ es Harris recurrente positiva y μ es su medida de probabilidad invariante, entonces*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} E_x u(x_k) = \int_X u(y)\mu(dy) \quad \forall x \in X,$$

para cada función medible y acotada $u : X \to \mathbb{R}$.

Bibliografía

[1] C.D. Aliprantis y K.C. Border (2006), *Infinite Dimensional Analysis*, Springer-Verlag, Berlin.

[2] A. Araposthatis, V.S. Borkar, E. Fernández-Gaucherand, M.K. Gosh, S.I. Marcus (1993), *Discrete-time controlled Markov processes with average cost criterion: a survey,* SIAM J. Control Optim 31, 238-344.

[3] R.B. Ash, C. Doléans-Dade (2000), Probability and Measure Theory (Second Edition), Academic Press.

[4] D.P. Bertsekas, S.E. Shreve (1976), *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, Inc.

[5] E.A. Feinberg, P.O. Kasyanov, N.V. Zadoianchuk (2013), *Berge's theorem for noncompact images sets*, J. Math. Anal. Appl. 397, 255–259.

[6] E. A. Feinberg, M. E. Lewis (2005), *Optimality of four-threshold policies in inventory systems with customer returns borrowing/storage options.* Probab. Eng. Inf. Sci. 19, 45-71.

[7] A. Fonseca-Morales (2010), *Semicontinuidad y medibilidad de correspondencias y la existencia de selectores continuos y medibles*, Tesis de Licenciatura, Universidad de Sonora.

[8] O. Hernández-Lerma y J.B. Lasserre (1996), *Discrete-Time Markov Control Processes. Basic Optimality Criteria, Springer-Verlag, N.Y.*

[9] O. Hernández-Lerma y J.B. Lasserre (1999), *Further Topics on Discrete-time Markov Control Processes,* Springer-Verlag, New York.

[10] O. Hernández-Lerma y J.B. Lasserre (2000), *Further criteria for positive Harris recurrence of Markov chains*, Proc. Amer. Math. Soc. 129, 1521-1524.

[11] O. Hernández-Lerma, R. Montes-de-Oca, R. Cavazos-Cadena (1991), *Recurrence conditions for Markov decision processes with Borel spaces: a survey*, Ann. Oper. Res. 28, 29–46.

[12] C.J. Himmelberg, F.S. Van Vleck (1975), *Multifunctions with values in a space of probability measures*, J. Math. Anal. Appl. 50, 108-112.

[13] A. Jaskiewicz (2009), *Zero-Sum Ergodic Semi-Markov Games with Weakly Contiuous Transition Probabilities,* J- Optim Theory Appl 141, 321-347.

[14] A. Jaskiewicz, A.S. Nowak (2006), *On the optimality equation for average cost Markov control processes with Feller transition probabilities,* J. Math. Anal. Appl. 316, 495–509.

[15] A. Jaskiewicz, A.S. Nowak (2016), *Zero-sum stochastic games*, in Handbook of Dynamic Game Theory, T. Başar, G. Zaccour (eds), 1-65, DOI 10.1007/978-3-319-27335-8_8-1

[16] H.-U. Küenle (2007), *On Markov games with average reward criterion and weakly continuous transition probabilities,* SIAM J. Control Optim. 45, 2156-2168.

[17] S.P. Meyn, R.L. Tweedie (1993), *Markov chains and stochastic stability*, Springer-Verlag, London.

[18] A.S. Nowak (1985), *Measurable selection theorems for minimax stochastic optimization problems*, SIAM J. Control Optim. 23, 466-476.

[19] A.S. Nowak (1986), *Semicontinuous nonstationary stochastic games.* J. Math. Anal. Appl. 117, 84-99.

[20] E. Nummelin (1983), *General Irreductible Markov Chains and Nonnegative Operators*, Cambridge University Press, Cambridge.

[21] S.M. Ross (1983), *Introduction to Stochastic Dynamic Programming: Probability and Mathematical*, Academic Press.

[22] O. Vega-Amaya (2003), *The average cost optimality equation: A fixed point approach*. Bol. Soc. Mat. Mexicana (3), 9, 185-195.

[23] O. Vega-Amaya (2003), *Zero-sum average semi-Markov games: fixed-point solutions of the Shapley equation,* SIAM J. Control Optim. 42, 1876-1894.

[24] O. Vega-Amaya (2018), *Solutions of the average cost optimality equation for Markov decision processes with weakly continuous kernel: The fixed-point approach revisited,* J. Math. Anal. Appl. 464, 152–163.