



"El saber de mis hijos
hará mi grandeza"

UNIVERSIDAD DE SONORA

DEPARTAMENTO DE MATEMÁTICAS

Programa de Licenciatura en Matemáticas

Juegos estocásticos de suma cero con dinámica
lineal y pago cuadrático

T E S I S

Que para obtener el título de:

Licenciado en Matemáticas

Presenta:

Francisco Alejandro Bernal Cañez

Director de tesis: Dr. Jesús Adolfo Minjárez Sosa

Hermosillo, Sonora, México, 21 de Junio de 2024

SINODALES

Dra. Carmen Geraldi Higuera Chan

Universidad de Sonora, Hermosillo, México

Dra. Alejandra Fonseca Morales

Universidad de Sonora, Hermosillo, México

Dr. Jesús Adolfo Minjárez Sosa

Universidad de Sonora, Hermosillo, México

Dr. Óscar Vega Amaya

Universidad de Sonora, Hermosillo, México

A todos aquellos que me han acompañado a lo largo de este camino.

En especial, a mi madre.....

Agradecimientos

Quiero expresar mi más sincero agradecimiento a todas las personas que me han acompañado y apoyado durante esta travesía académica.

En primer lugar, a mi familia, quienes han sido mi pilar fundamental. A mi madre, por siempre creer en mí y brindarme su amor y apoyo incondicional. A mis tíos Javier, Angélica, Peter y Martha, por todo su apoyo y cariño a lo largo de estos años.

A Ana Sofía Salcido, por acompañarme durante los últimos ocho años. Su presencia y apoyo han sido fundamentales para convertirme en la persona que soy.

A mis grandes amigos, Luis Pablo y Hisaki, por acompañarme desde el inicio de la Licenciatura en Matemáticas y compartir conmigo cada paso de este camino. A Catherine, América, Marlon y Carlón, por ser mis compañeros de generación y por entender y compartir la experiencia única de haber comenzado nuestra carrera en plena pandemia. También quiero agradecer a Burruel, Gabriel y Chao por las valiosas pláticas y los buenos momentos de este último año.

A Héctor, Jorge y Tizado, por ser de gran importancia en esta travesía universitaria y por compartir varios de los momentos más memorables en mi formación académica y personal.

Finalmente, quiero expresar mi profundo agradecimiento a la planta docente del departamento de matemáticas, quienes han sido fundamentales en mi desarrollo académico. En especial, a mis sinodales, la Dra. Carmen Geraldi Higuera Chan, la Dra. Alejandra Fonseca Morales, el Dr. Jesús Adolfo Minjárez Sosa y el Dr. Oscar Vega Amaya, así como también a la Dra. Gudelia Figueroa Preciado, el Dr. Misael Avendaño Camacho y la Dra. Carolina Espinoza, por su guía y apoyo a lo largo de mi formación.

Índice general

Agradecimientos	III
Índice general	III
Introducción	1
Preliminares	4
0.1. Espacios de Borel y kernels estocásticos	4
0.2. Multifunciones y selectores	5
1. Juegos estocásticos de suma cero	6
1.1. Modelo de un juego estocástico GM_F	6
1.2. Estrategias	8
1.3. Criterio de optimalidad	9
1.4. Hipótesis y suposiciones	13
2. Juegos estocásticos con horizonte finito	14
2.1. Preliminares	14
2.2. Teorema de existencia del valor y del par óptimo para el caso de horizonte finito	20
3. Juegos estocásticos de horizonte infinito	26
3.1. Preliminares	26
3.2. Teorema de existencia del valor y del par óptimo para el caso de horizonte infinito	32
4. Juegos Lineales - Cuadráticos (Linear-Quadratic LQ)	36
4.1. Definición, Motivación y Revisión de Hipótesis en Juegos-LQ	36
4.2. Valor y Estrategias Óptimas para Juegos-LQ	42
4.3. Gráficas	51
A. Funciones semi-continuas	56
B. Teoremas minimax y selectores	57

Índice general v

C. Auxiliares 62

Bibliografía 64

Introducción

La teoría de juegos es una rama de las matemáticas que analiza las interacciones estratégicas entre individuos racionales llamados jugadores en situaciones donde los resultados dependen de las acciones o decisiones que ellos tomen. Por lo tanto, su objetivo principal es determinar las decisiones que lleven a obtener el mejor beneficio a los jugadores, ya sea que maximicen sus ganancias o que minimicen sus costos.

Con el notable avance de otras ciencias y de las áreas matemáticas como la probabilidad y los procesos estocásticos, en los últimos años, la teoría de juegos ha experimentado un desarrollo significativo, encontrando aplicaciones en campos tan diversos como la economía, la biología, la política y la informática.

Existen varias clases de juegos: cooperativos o no cooperativos, estáticos o dinámicos, deterministas o estocásticos. En particular, el estudio de juegos dinámicos estocásticos se puede dividir, por ejemplo:

- 1.- Según la cantidad de jugadores.
- 2.- Según el tipo de espacios de estados:
 - a) espacio numerable;
 - b) espacio de Borel (subconjunto de Borel de un espacio métrico separable y completo);
- 3.- Según el horizonte de planeación (cantidad de etapas):
 - a) Horizonte finito;
 - b) Horizonte infinito;
- 4.- Según la relación que exista entre las ganancias y pérdidas de cada jugador:

- a) Juegos de suma cero (La ganancia de un jugador es igual a la suma de las pérdidas de los otros);
- b) Juegos de suma no cero;

5.- Según el índice de funcionamiento:

- a) Pago total;
- b) Pago descontado;
- c) Pago promedio.

En este trabajo nos centramos en el estudio de juegos dinámicos estocásticos de suma cero con dos jugadores, considerando espacios de estados de Borel bajo el criterio de pago descontado con horizonte infinito, y total con horizonte finito. Específicamente estudiaremos las condiciones que garantizan la existencia del valor del juego y un par de estrategias óptimas. Todos los resultados los ejemplificaremos con una clase particular de juegos llamados **juegos-LQ**. Estos son juegos donde el proceso de estados evoluciona de acuerdo a una dinámica lineal y el pago es cuadrático. De aquí el nombre LQ (Linear-Quadratic).

El trabajo está estructurado de la siguiente manera:

En el primer capítulo se presenta el modelo de los juegos estocásticos que estudiaremos en esta tesis, así como su interpretación. Además, se define el concepto de estrategia y se establecen criterios de optimalidad, junto con el concepto de función valor de un juego. También se establecen hipótesis y suposiciones que los juegos estudiados deben cumplir. Esta sección se apoya principalmente en algunos resultados encontrados en [4].

En el segundo capítulo se estudian juegos con horizonte finito. En particular se demuestra la existencia del valor del juego, así como de estrategias óptimas. La principal fuente de referencia para este capítulo es [11].

De manera similar, el tercer capítulo presenta el estudio de juegos con horizonte infinito. Al igual que en el primer capítulo, la referencia principal es [4].

Finalmente, el cuarto capítulo aborda los Juegos-LQ. Se verifica que las hipótesis establecidas a lo largo del trabajo se cumplen para este caso específico y se resuelve completamente. Además, se incluyen gráficas que complementan la comprensión de este tipo de juego. Este capítulo se fundamenta principalmente en resultados de [8].

El resto de la bibliografía se emplea mayormente para remitir a los lectores a pruebas detalladas sobre resultados necesarios para la realización de las demostraciones presentadas en esta tesis.

Preliminares

En este capítulo introducimos conceptos y definiciones que usaremos durante el desarrollo del presente trabajo.

0.1. Espacios de Borel y kernels estocásticos

Un subconjunto de Borel X de un espacio métrico completo y separable es llamado **espacio de Borel**, y su σ -álgebra se denota por $\mathcal{B}(X)$. Un subconjunto de Borel de un espacio de Borel es también un espacio de Borel.

Sean X y Y espacios de Borel. Para este trabajo, cuando se dice “medible” nos referimos a “Borel medible”.

Definición 0.1.1. *Un **kérel estocástico** en X dado Y es una función $P(\cdot|\cdot)$ tal que*

- i) $P(\cdot|y)$ es una medida de probabilidad en X para cada $y \in Y$ fijo;*
- ii) $P(D|\cdot)$ es una función medible en Y para cada $D \in \mathcal{B}(X)$ fijo.*

$P(dx|y)$ denota la medida de probabilidad asociada a $y \in Y$ que, para $D \in \mathcal{B}(X)$, nos arroja la probabilidad condicional $P(D|y)$.

El conjunto de todos los kernels estocásticos en X dado Y se denota por $\mathbb{P}(X|Y)$.

Denotamos por $M(X)$ al conjunto de todas las funciones medibles $u : X \rightarrow \mathbb{R}$ y por $M_b(X)$ al subconjunto de todas las funciones acotadas en $M(X)$. Por otro lado, se denota por $C_b(X)$ al conjunto de todas las funciones continuas en $M_b(X)$. Así, se tiene que

$$C_b(X) \subset M_b(X) \subset M(X).$$

Dado un espacio de Borel X , denotemos por $\mathbb{P}(X)$ a la familia de medidas de probabilidad en X .

Observación 0.1.1. *A menos que se diga lo contrario, a lo largo del siguiente trabajo de tesis suponemos que $\mathbb{P}(X)$ está dotada con la topología débil, de modo que una sucesión μ_n en $\mathbb{P}(X)$ converge débilmente a una medida μ ($\mu_n \Rightarrow \mu$) si $\int u d\mu_n \rightarrow \int u d\mu$ para cada u en $C_b(X)$. Además, bajo la topología débil, para cualquier espacio de Borel X :*

i) $\mathbb{P}(X)$ es un espacio de Borel. (Consulte [6], página 91)

ii) Si X es compacto, $\mathbb{P}(X)$ también lo es. (Consulte [10], Teorema II 6.4.)

0.2. Multifunciones y selectores

Sean X y A espacios de Borel no vacíos.

Una **multifunción** F de X a A es una función tal que $F(x)$ es un subconjunto no vacío de A para cada $x \in X$.

Definición 0.2.1. a) Una multifunción F de X a A es **medible** si

$$F^{-1}(U) := \{x \in X : F(x) \cap U \neq \emptyset\}$$

es un subconjunto de Borel de X para cada conjunto abierto $U \subset A$. A una multifunción F se le llama **valuada en los cerrados** si $F(x)$ es un conjunto cerrado para toda $x \in X$. A su vez, esta es **valuada en los compactos** si $F(x)$ es un conjunto compacto para toda $x \in X$.

b) La **gráfica** de la multifunción F es el subconjunto de $X \times A$ definido como

$$GrF := \{(x, a) : x \in X, a \in F(x)\}$$

Decimos que F tiene una gráfica medible si GrF está en $\mathcal{B}(X \times A)$.

c) \mathbb{F}_F denota al conjunto (valuado singularmente) de funciones medibles $f : X \rightarrow A$ tales que $(x, f(x))$ está en GrF , esto es, $f(x) \in F(x)$ para cada $x \in X$. Una función $f \in \mathbb{F}_F$ es llamada **selector** de la multifunción F .

Capítulo 1

Juegos estocásticos de suma cero

En este capítulo introduciremos la clase de juegos estocásticos de suma cero de dos jugadores en los que estamos interesados. También presentaremos el criterio de optimalidad y las hipótesis que se utilizarán para el desarrollo de este trabajo. Denotaremos por J1 y J2 a los jugadores correspondientes.

1.1. Modelo de un juego estocástico GM_F

Consideremos el modelo de juego de suma cero de dos jugadores

$$GM_F := \{X, A, B, \mathbb{K}_A, \mathbb{K}_B, F, \rho, r\}$$

donde:

- i) X es el conjunto de todos los estados posibles del juego. Asumiremos que X es un espacio de Borel.
- ii) A y B son los espacios de acción para los jugadores J1 y J2, respectivamente, y también se asume que son espacios de Borel.
- iii) \mathbb{K}_A y \mathbb{K}_B son subconjuntos de Borel no vacíos de $X \times A$ y $X \times B$, respectivamente. Para cada $x \in X$

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

representa el conjunto de las acciones admisibles para J1 en el estado x del juego. Análogamente,

$$B(x) := \{b \in B : (x, b) \in \mathbb{K}_B\}$$

denota el conjunto de las acciones admisibles para J2 en el estado x del juego. Así, se define

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

que es un subconjunto de Borel de $X \times A \times B$. (Consulte el Lema 1.1 en [9])

iv) $F : X \times A \times B \times \mathbb{R} \rightarrow X$ es una función que define el proceso de estados del juego. Es decir, asumiremos que el juego evoluciona de acuerdo a la ecuación

$$x_{t+1} = F(x_t, a_t, b_t, \varepsilon_t), \quad t \in \mathbb{N}_0$$

donde $x_t \in X$ y $(a_t, b_t) \in A(x_t) \times B(x_t)$ representan el estado del juego y las acciones elegidas por los jugadores J1 y J2 al tiempo t , respectivamente. Además, $\{\varepsilon_t\}$ es una sucesión de variables aleatorias i.i.d con densidad ρ que toman valores en \mathbb{R} .

v) $r : \mathbb{K} \rightarrow \mathbb{R}$ es una función medible que denota la función de pago y representa una ganancia para J1 y, a su vez, un costo para J2.

Un juego estocástico GM_F es jugado de la siguiente manera:

En cada fase $t = 0, 1, \dots$, J1 y J2 observan el estado actual $x_t = x \in X$ del juego y, de manera independiente y con base en una estrategia (se define en la siguiente sección), escogen acciones $a_t = a \in A(x)$ y $b_t = b \in B(x)$ respectivamente. Como consecuencia de esto sucede lo siguiente:

- (1) J1 recibe una ganancia inmediata $r(x_t, a_t, b_t)$.
- (2) J2 paga un costo de $r(x_t, a_t, b_t)$.
- (3) El sistema se mueve a un nuevo estado $x_{t+1} = x' \in X$ conforme a la ley de evolución dada por el kernel estocástico

$$\begin{aligned}
Q(D|x, a, b) &:= P[x_{t+1} \in D | x_t = x, a_t = a, b_t = b], \quad D \in \mathcal{B}(X) \\
&= \int_{\mathbb{R}} 1_D[F(x, a, b, s)] \rho(s) ds, \quad D \in \mathcal{B}(X), \quad (1.1.1.1)
\end{aligned}$$

donde $1_D(\cdot)$ denota la función indicadora del conjunto D . Una vez que el juego se encuentra en el estado x' , J1 y J2 repiten el proceso.

La meta de J1 es maximizar su ganancia a lo largo de las fases del juego mientras que el objetivo de J2 es la de minimizar su costo.

Denotaremos $\mathbb{P}_A(x) := \mathbb{P}(A(x))$ y $\mathbb{P}_B(x) := \mathbb{P}(B(x))$ para cada $x \in X$. Entonces, $x \mapsto \mathbb{P}_A(x)$ y $x \mapsto \mathbb{P}_B(x)$ definen multifunciones de X a $\mathbb{P}(A)$ y de X a $\mathbb{P}(B)$, que serán denotadas por \mathbb{P}_A y \mathbb{P}_B , respectivamente.

1.2. Estrategias

Sea $H_0 = X$ y $H_t = \mathbb{K} \times H_{t-1}$ para $t = 1, 2, \dots$. Para cada t un elemento

$$h_t = (x_0, a_0, b_0, \dots, x_{t-1}, a_{t-1}, b_{t-1}, x_t)$$

de H_t representa la historia del juego hasta la etapa t . Una estrategia aleatorizada π para J1 es una sucesión $\pi = \{\pi_t, t = 0, 1, \dots\}$ de kernels estocásticos π_t en $\mathbb{P}(A|H_t)$ tal que

$$\pi_t(A(x_t)|h_t) = 1 \quad \forall h_t \in H_t, \quad t = 0, 1, \dots$$

Denotamos por Π a la familia de estrategias de J1.

Una estrategia $\pi = \{\pi_t\}$ es llamada **de Markov** si $\pi_t \in \mathbb{P}(A|X)$ para cada $t = 0, 1, \dots$, esto es, cada π_t depende solamente del estado actual x_t del sistema. El conjunto de todas las estrategias de Markov de J1 se denota por Π_M . Una estrategia de Markov $\pi = \{\pi_t\}$ se dice ser una estrategia **estacionaria** si existe $f \in \mathbb{P}(A|X)$ tal que $\pi_t = f$ para cada $t = 0, 1, \dots$. En este caso, la estrategia estacionaria π sera identificada por f . Denotamos por Π_S al conjunto de todas las estrategias estacionarias de J1. De este modo, claramente se tiene que

$$\Pi_S \subset \Pi_M \subset \Pi.$$

Los conjuntos de estrategias aleatorizadas, de Markov y estacionarias para J2 se definen de forma análoga y las denotaremos como $\Gamma, \Gamma_M, \Gamma_S$, respectivamente.

Sea (Ω, \mathcal{F}) el espacio medible que consiste del espacio muestral $\Omega := (X \times A \times B)^\infty$ y su σ -álgebra producto \mathcal{F} . Entonces para cada par de estrategias $(\pi, \gamma) \in \Pi \times \Gamma$ y cada estado inicial $x \in X$, por el Teorema de C. Ionescu-Tulcea (consulte [1], página 109, [6], página 80), existe una única medida de probabilidad $P_x^{\pi\gamma}$ y proceso estocástico $\{(x_t, a_t, b_t), t = 0, 1, \dots\}$ definido en (Ω, \mathcal{F}) en una forma canónica, donde x_t, a_t y b_t representan el estado y las acciones de J1 y J2, respectivamente, en la etapa $t = 0, 1, \dots$. El operador esperanza con respecto a $P_x^{\pi\gamma}$ se denota por $E_x^{\pi\gamma}$.

1.3. Criterio de optimalidad

Definición 1.3.1. Sea $\alpha \in (0, 1)$ un número fijo. Definimos la función de pago α -descuento esperado como

$$J_\alpha(x, \pi, \gamma) := E_x^{\pi\gamma} \left[\sum_{t=0}^{\infty} \alpha^t r(x_t, a_t, b_t) \right], \quad (1.1.3.1)$$

para cada par de estrategias (π, γ) y cada estado inicial x . Al número α se le llama “factor de descuento”.

Definición 1.3.2. Consideremos nuevamente un factor de descuento $\alpha \in (0, 1)$. Para $n = 1, 2, \dots$, definimos la función de pago esperado para un juego en n etapas como

$$J_n(x, \pi, \gamma) := E_x^{\pi\gamma} \left[\sum_{t=0}^{n-1} \alpha^t r(x_t, a_t, b_t) \right]. \quad (1.1.3.2)$$

Estudiaremos el caso en el cual el modelo del juego está bien definido, es decir,

$$\sup_{\pi \in \Pi} \sup_{\gamma \in \Gamma} |J_\alpha(x, \pi, \gamma)| < \infty, \quad \forall x \in X. \quad (1.1.3.3)$$

Observación 1.3.1. La condición (1.1.3.3) se cumple si r es acotada. En efecto, si $|r(x, a, b)| \leq M$ para alguna constante M , entonces

$$E_x^{\pi\gamma} \left[\sum_{t=0}^{\infty} \alpha^t |r(x_t, a_t, b_t)| \right] \leq E_x^{\pi\gamma} \left[\sum_{t=0}^{\infty} \alpha^t M \right] = E_x^{\pi\gamma} \left[\frac{M}{1 - \alpha} \right] = \frac{M}{1 - \alpha}, \quad \forall (\pi, \gamma) \in \Pi \times \Gamma, x \in X.$$

Dado que

$$|J_\alpha(x, \pi, \gamma)| \leq E_x^{\pi\gamma} \left[\sum_{t=0}^{\infty} \alpha^t |r(x_t, a_t, b_t)| \right]$$

se llega a que

$$|J_\alpha(x, \pi, \gamma)| \leq \frac{M}{1-\alpha} < \infty, \quad \forall (\pi, \gamma) \in \Pi \times \Gamma, x \in X.$$

Para introducir el primer criterio de optimalidad necesitamos los siguientes conceptos.

Definición 1.3.3. Para cada $n = 1, 2, \dots$, las funciones

$$L_n(x) := \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J_n(x, \pi, \gamma)$$

y

$$U_n(x) := \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J_n(x, \pi, \gamma)$$

son llamadas funciones de valor **inferior** y **superior**, respectivamente, para el juego en n etapas.

Observación 1.3.2. Notemos que para $\hat{\pi} \in \Pi$ y $\hat{\gamma} \in \Gamma$ arbitrarios

$$J_n(x, \hat{\pi}, \hat{\gamma}) \leq \sup_{\pi \in \Pi} J_n(x, \pi, \hat{\gamma}),$$

de donde

$$\inf_{\gamma \in \Gamma} J_n(x, \hat{\pi}, \gamma) \leq \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J_n(x, \pi, \gamma),$$

por lo que

$$\sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J_n(x, \pi, \gamma) \leq \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J_n(x, \pi, \gamma).$$

Así, podemos concluir que en general $L_n(\cdot) \leq U_n(\cdot)$.

Si se cumple que $L_n(x) = U_n(x)$ para toda $x \in X$, entonces la función común es llamada **valor** o **función valor del juego en n etapas**, y se denota por $V_n(\cdot)$.

Definición 1.3.4. Consideremos un juego en n etapas.

i) Una estrategia $\pi^* \in \Pi$ es **óptima para J1** si

$$U_n(x) \leq J_n(x, \pi^*, \gamma) \quad \text{para cada } \gamma \in \Gamma \text{ y } x \in X.$$

ii) Una estrategia $\gamma^* \in \Gamma$ es **óptima para J2** si

$$L_n(x) \geq J_n(x, \pi, \gamma^*) \quad \text{para cada } \pi \in \Pi \text{ y } x \in X.$$

Si se cumplen i) y ii), entonces (π^*, γ^*) es un **par de estrategias óptimas**, o bien **punto silla** para el juego.

Proposición 1.3.1. La definición anterior es equivalente a decir que (π^*, γ^*) es un par de estrategias óptimas si

$$U_n(x) = \inf_{\gamma \in \Gamma} J_n(x, \pi^*, \gamma) \quad \text{y} \quad L_n(x) = \sup_{\pi \in \Pi} J_n(x, \pi, \gamma^*).$$

Esto se deduce de lo siguiente:

(\Rightarrow)

Si

$$U_n(x) \leq J_n(x, \pi^*, \gamma) \quad \forall \gamma \in \Gamma, \quad x \in X$$

entonces

$$U_n(x) \leq \inf_{\gamma \in \Gamma} J_n(x, \pi^*, \gamma) \quad \forall x \in X;$$

y como

$$J_n(x, \pi^*, \gamma) \leq \sup_{\pi \in \Pi} J_n(x, \pi, \gamma) \quad \forall \gamma \in \Gamma, \quad x \in X,$$

entonces

$$\inf_{\gamma \in \Gamma} J_n(x, \pi^*, \gamma) \leq \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J_n(x, \pi, \gamma) = U_n(x) \quad \forall x \in X.$$

Por lo tanto $U_n(x) = \inf_{\gamma \in \Gamma} J_n(x, \pi^*, \gamma)$. Además, si

$$L_n(x) \geq J_n(x, \pi, \gamma^*) \quad \forall \pi \in \Pi, \quad x \in X$$

entonces

$$L_n(x) \geq \sup_{\pi \in \Pi} J_n(x, \pi, \gamma^*) \quad \forall x \in X;$$

y dado que

$$J_n(x, \pi, \gamma^*) \geq \inf_{\gamma \in \Gamma} J_n(x, \pi, \gamma) \quad \forall \pi \in \Pi, x \in X,$$

entonces

$$\sup_{\pi \in \Pi} J_n(x, \pi, \gamma^*) \geq \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J_n(x, \pi, \gamma) = L_n(x) \quad \forall x \in X.$$

Por ende $L_n(x) = \sup_{\pi \in \Pi} J_n(x, \pi, \gamma^*)$.

(\Leftarrow)

Por otro lado, si

$$U_n(x) = \inf_{\gamma \in \Gamma} J_n(x, \pi^*, \gamma) \quad \forall x \in X$$

entonces

$$U_n(x) \leq J_n(x, \pi^*, \gamma) \quad \forall \gamma \in \Gamma, x \in X,$$

y si

$$L_n(x) = \sup_{\pi \in \Pi} J_n(x, \pi, \gamma^*) \quad \forall x \in X$$

entonces

$$L_n(x) \geq J_n(x, \pi, \gamma^*) \quad \forall \pi \in \Pi, x \in X.$$

Esto prueba la equivalencia.

Observación 1.3.3. Conociendo el par de estrategias óptimas para un juego estocástico GM_F es sencillo alcanzar el valor del juego. Es decir, si (π^*, γ^*) son un par de estrategias

óptimas, por Definición 1.3.4

$$J_n(x, \pi^*, \gamma^*) \leq L_n(x) \leq U_n(x) \leq J_n(x, \pi^*, \gamma^*)$$

de donde se concluye que $L_n(x) = U_n(x) = V_n(x) = J_n(x, \pi^*, \gamma^*)$

Las funciones de valor y el par de estrategias óptimas para $J_\alpha(x, \pi, \gamma)$ se definen de forma similar. En particular, el valor del juego lo denotaremos como V_α .

1.4. Hipótesis y suposiciones

- (a) Para cada estado $x \in X$, los conjuntos no vacíos $A(x)$ y $B(x)$ de acciones admisibles son compactos.
- (b) Para cada $(x, a, b) \in \mathbb{K}$, $r(x, \cdot, b)$ es semi-continua superiormente en $A(x)$, y $r(x, a, \cdot)$ es semi-continua inferiormente en $B(x)$.
- (c) Para cada $(x, a, b) \in \mathbb{K}$ y cada función $v \in M_b(X)$, las funciones

$$\int_{\mathbb{R}} v [F(x, \cdot, b, s)] \rho(s) ds \quad \text{y} \quad \int_{\mathbb{R}} v [F(x, a, \cdot, s)] \rho(s) ds$$

son continuas en $A(x)$ y $B(x)$, respectivamente.

- (d) Existe una constante $M > 0$ y una función medible $w : X \rightarrow \mathbb{R}$ tal que $w(x) \geq 1$ para cada $x \in X$ y

$$|r(x, a, b)| \leq Mw(x) \quad \forall (x, a, b) \in \mathbb{K}, \quad (1.1.4.1)$$

y, además, (c) se cumple al sustituir v por w .

- (e) Existe una constante $1 \leq \nu \leq \frac{1}{\alpha}$ tal que

$$\int_{\mathbb{R}} w [F(x, a, b, s)] \rho(s) ds \leq \nu w(x), \quad \forall (x, a, b) \in \mathbb{K} \quad (1.1.4.2)$$

Capítulo 2

Juegos estocásticos con horizonte finito

En este capítulo se demuestra que los juegos estocásticos GM_F de suma cero con horizonte finito tienen valor. A partir de este hecho, mostraremos que existe un par de estrategias óptimas.

2.1. Preliminares

Sea n un entero positivo. El juego estocástico de n etapas en el cual los jugadores juegan hasta la etapa n es llamado **juego de horizonte finito**. Sean π y γ las estrategias de J1 y J2, respectivamente. Entonces el pago esperado en el juego está dado por $J_n(x, \pi, \gamma)$ como en la Definición 1.3.2.

Antes de introducir los resultados principales, es necesario establecer cierta notación y resultados preliminares. Primero que nada, por la Hipótesis 1.4(a) y la Observación 0.1.1 ii), notemos que las multifunciones $x \mapsto \mathbb{P}_A(x)$ y $x \mapsto \mathbb{P}_B(x)$ introducidas en el capítulo anterior son medibles y valuadas en los compactos.

Sean $x \in X$, $\mu \in \mathbb{P}_A(x)$ y $\lambda \in \mathbb{P}_B(x)$. Definimos

$$\tilde{r}(x, \mu, \lambda) := \int_{B(x)} \int_{A(x)} r(x, a, b) \mu(da) \lambda(db), \quad (2.2.1.1)$$

y para cada conjunto de Borel $D \in \mathcal{B}(X)$,

$$\tilde{Q}(D|x, \mu, \lambda) := \int_{B(x)} \int_{A(x)} Q(D|x, a, b) \mu(da) \lambda(db),$$

de modo que

$$\int_X u(y) \tilde{Q}(dy|x, \mu, \lambda) := \int_{B(x)} \int_{A(x)} \int_X u(y) Q(dy|x, a, b) \mu(da) \lambda(db). \quad (2.2.1.2)$$

Definición 2.1.1. Para cada función medible $u : X \rightarrow \mathbb{R}$ definimos su **w-norma** como

$$\|u\|_w := \sup_{x \in X} \frac{|u(x)|}{w(x)}$$

donde w es la función introducida en la Hipótesis 1.4(d). Denotamos por $\mathbb{B}_w(X)$ al espacio lineal normado de todas las funciones medibles u en X para las cuales $\|u\|_w$ es finita.

Observación 2.1.1. $\mathbb{B}_w(X)$ es un espacio de Banach. En efecto, consideremos $M_b(X)$ el espacio de Banach de funciones medibles $u : X \rightarrow \mathbb{R}$ acotadas bajo la **norma del supremo**

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

Notemos que $\mathbb{B}_w(X)$ es un espacio completo ya que si $\{u_n\}$ es una sucesión de Cauchy con la w -norma, entonces $\left\{\frac{u_n}{w}\right\}$ es de Cauchy en la norma del supremo. Así, dado que $M_b(X)$ es espacio de Banach, se puede deducir la existencia de una función $\frac{u}{w}$ tal que $\frac{u_n}{w} \rightarrow \frac{u}{w}$. Es posible notar que $\frac{wu}{w} = u \in \mathbb{B}_w(X)$ será el límite de la sucesión $\{u_n\}$ con la w -norma. De aquí se sigue que $\mathbb{B}_w(X)$ es un espacio de Banach.

Para cada $u \in \mathbb{B}_w(X)$ y $(x, a, b) \in \mathbb{K}$, definimos

$$H(u; x, a, b) := r(x, a, b) + \alpha \int_X u(y) Q(dy|x, a, b) \quad (2.2.1.3)$$

y

$$\tilde{H}(u; x, \mu, \lambda) = \tilde{r}(x, \mu, \lambda) + \alpha \int_X u(y) \tilde{Q}(dy|x, \mu, \lambda),$$

donde el primer y el segundo término son como en (2.2.1.1) y (2.2.1.2), respectivamente.

Observación 2.1.2. Verificando para funciones indicadoras, simples, positivas y finalmente para funciones $u \in \mathbb{B}_w(x)$ en general, veremos que en el contexto (1.1.1.1) se tendrá

que

$$\int_X u(y)Q(dy|x, a, b) = \int_{\mathbb{R}} u [F(x, a, b, s)] \rho(s)ds. \quad (2.2.1.4)$$

Además,

$$\tilde{Q}(D|x, \mu, \lambda) = \int_{B(x)} \int_{A(x)} \left[\int_{\mathbb{R}} 1_D [F(x, a, b, s)] \rho(s)ds \right] \mu(da)\lambda(db). \quad (2.2.1.5)$$

Consecuentemente

$$\int_X u(y)\tilde{Q}(dy|x, \mu, \lambda) = \int_{B(x)} \int_{A(x)} \left[\int_{\mathbb{R}} u [F(x, a, b, s)] \rho(s)ds \right] \mu(da)\lambda(db). \quad (2.2.1.6)$$

Por convención, a partir de aquí, usaremos

$$\int_X u(y)\tilde{Q}(dy|x, \mu, \lambda) = \int_{\mathbb{R}} u [\tilde{F}(x, \mu, \lambda, s)] \rho(s)ds.$$

Dicho esto, para el desarrollo de este trabajo podemos tomar

$$H(u; x, a, b) = r(x, a, b) + \alpha \int_{\mathbb{R}} u [F(x, a, b, s)] \rho(s)ds. \quad (2.2.1.7)$$

y

$$\tilde{H}(u; x, \mu, \lambda) = \tilde{r}(x, \mu, \lambda) + \alpha \int_{\mathbb{R}} u [\tilde{F}(x, \mu, \lambda, s)] \rho(s)ds \quad (2.2.1.8)$$

Observación 2.1.3. Agrupando de la manera adecuada, es posible obtener que

$$\begin{aligned} \tilde{H}(u; x, \mu, \lambda) &= \tilde{r}(x, \mu, \lambda) + \alpha \int_{\mathbb{R}} u [\tilde{F}(x, \mu, \lambda, s)] \rho(s)ds \\ &= \int_{B(x)} \int_{A(x)} H(u; x, a, b)\mu(da)\lambda(db). \end{aligned}$$

Definimos al **operador de Shapley** como

$$T_\alpha u(x) := \sup_{\mu \in \mathbb{P}_A(x)} \inf_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(u; x, \mu, \lambda), \quad \forall u \in \mathbb{B}_w(X),$$

o bien

$$T_\alpha u(x) = \sup_{\mu \in \mathbb{P}_A(x)} \inf_{\lambda \in \mathbb{P}_B(x)} \left\{ \tilde{r}(x, \mu, \lambda) + \alpha \int_{\mathbb{R}} u [\tilde{F}(x, \mu, \lambda, s)] \rho(s)ds \right\}, \quad \forall u \in \mathbb{B}_w(X).$$

Por las Hipótesis 1.4(b) y 1.4(c), así como por el Teorema A.0.2, sabemos que en $\mathbb{P}_A(x)$ y en $\mathbb{P}_B(x)$ se alcanzan en el máximo y el mínimo, respectivamente. De este modo,

es posible reemplazar supremo e ínfimo por máximo y mínimo, respectivamente, en la ecuación anterior, y obtener que

$$T_\alpha u(x) := \max_{\mu \in \mathbb{P}_A(x)} \min_{\lambda \in \mathbb{P}_B(x)} \left\{ \tilde{r}(x, \mu, \lambda) + \alpha \int_{\mathbb{R}} u \left[\tilde{F}(x, \mu, \lambda, s) \right] \rho(s) ds \right\}, \quad \forall u \in \mathbb{B}_w(X).$$

que también puede ser expresado como

$$T_\alpha u(x) = \max_{\mu \in \mathbb{P}_A(x)} \min_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(u; x, \mu, \lambda), \quad \forall u \in \mathbb{B}_w(X). \quad (2.2.1.9)$$

Dicho todo lo anterior, solo falta tener en cuenta el siguiente Lema para poder probar la existencia del valor y el par de estrategias óptimas en juegos estocásticos GM_F de horizonte finito:

Lema 2.1.1. *Suponga que se cumplen las Hipótesis 1.4. Entonces, para cada $u \in \mathbb{B}_w(X)$:*

$$(a) \quad T_\alpha u(x) = \min_{\lambda \in \mathbb{P}_B(x)} \max_{\mu \in \mathbb{P}_A(x)} \tilde{H}(u; x, \mu, \lambda).$$

(b) *Existen selectores medibles $f_0 \in \mathbb{F}_{\mathbb{P}_A}$ y $g_0 \in \mathbb{F}_{\mathbb{P}_B}$ tales que, para toda $x \in X$,*

$$\begin{aligned} T_\alpha u(x) &= \max_{\mu \in \mathbb{P}_A(x)} \tilde{H}(u; x, \mu, g_0(x)) \\ &= \min_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(u; x, f_0(x), \lambda) \\ &= \tilde{H}(u; x, f_0(x), g_0(x)). \end{aligned} \quad (2.2.1.10)$$

(c) $T_\alpha u$ está en $\mathbb{B}_w(X)$.

Demostración. Tomemos una función $u \in \mathbb{B}_w(X)$ arbitraria.

(a) Por las Hipótesis 1.4(c) y 1.4(d), la integral en (2.2.1.3) es continua en $a \in A(x)$ y $b \in B(x)$. Es por este hecho y por la Hipótesis 1.4(b) que para cada $(x, a, b) \in \mathbb{K}$ la función $H(u; x, \cdot, b)$ es semi-continua superiormente en $A(x)$ y $H(u; x, a, \cdot)$ es semi-continua inferiormente en $B(x)$. De aquí, además considerando la topología débil que asumimos para las medidas de probabilidad y aprovechando el Teorema de Convergencia Monótona (vea el Teorema C.0.2), podemos obtener que la función $\tilde{H}(u; x, \mu, \lambda)$ es semi-continua superiormente en $\mu \in \mathbb{P}_A(x)$ y semi-continua inferiormente en $\lambda \in \mathbb{P}_B(x)$ [consulte [7], Lema 2.3]. Más aún, podemos notar que $\tilde{H}(u; x, \mu, \lambda)$ es cóncava en μ :

$$\begin{aligned}
 \tilde{H}(u; x, t\mu_1 + (1-t)\mu_2, \lambda) &= \int_{B(x)} \left[\int_{A(x)} H(u; x, a, b)(t\mu_1(da) + (1-t)\mu_2(da)) \right] \lambda(db) \\
 &= \int_{A(x)} \left[\int_{B(x)} H(u; x, a, b)\lambda(db) \right] (t\mu_1(da) + (1-t)\mu_2(da)) \\
 &= t \int_{A(x)} \left[\int_{B(x)} H(u; x, a, b)\lambda(db) \right] \mu_1(da) \\
 &\quad + (1-t) \int_{A(x)} \left[\int_{B(x)} H(u; x, a, b)\lambda(db) \right] \mu_2(da) \\
 &= t\tilde{H}(u; x, \mu_1, \lambda) + (1-t)\tilde{H}(u; x, \mu_2, \lambda) \\
 \Rightarrow \tilde{H}(u; x, t\mu_1 + (1-t)\mu_2, \lambda) &\leq t\tilde{H}(u; x, \mu_1, \lambda) + (1-t)\tilde{H}(u; x, \mu_2, \lambda) \quad ,
 \end{aligned}$$

donde el cambio en el orden de integración en la segunda línea es posible gracias al Teorema de Fubini (vea el Teorema C.0.3). De forma análoga, se puede probar que $\tilde{H}(u; x, \mu, \lambda)$ es convexa en λ . Así, viendo que $\tilde{H}(u; x, \mu, \lambda)$ satisface todas las condiciones que pide el Teorema de Ky Fan (vea el Teorema B.0.3), se obtiene que

$$T_\alpha u(x) = \max_{\mu \in \mathbb{P}_A(x)} \min_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(u; x, \mu, \lambda) = \min_{\lambda \in \mathbb{P}_B(x)} \max_{\mu \in \mathbb{P}_A(x)} \tilde{H}(u; x, \mu, \lambda)$$

(b) Definamos

$$\tilde{H}_1(x, \mu) := \min_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(u; x, \mu, \lambda) \quad (2.2.1.11)$$

para cada $x \in X$ y $\mu \in \mathbb{P}_A(x)$. De lo establecido en el inciso anterior, $\tilde{H}_1(x, \cdot)$ es semi-continua superiormente en $\mathbb{P}_A(x)$. Por lo tanto, por la Observación 0.1.1 ii) y el Teorema B.0.2, existe $f_0 \in \mathbb{F}_{\mathbb{P}_A}$ tal que

$$\tilde{H}_1(x, f_0(x)) = \max_{\mu \in \mathbb{P}_A(x)} \tilde{H}_1(x, \mu) \quad \forall x \in X.$$

Así, obtenemos que

$$\tilde{H}_1(x, f_0(x)) = \max_{\mu \in \mathbb{P}_A(x)} \min_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(u; x, \mu, \lambda). \quad (2.2.1.12)$$

De este modo, combinando (2.2.1.9) y (2.2.1.12), se obtiene que

$$T_\alpha u(x) = \min_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(u; x, f_0(x), \lambda).$$

Similarmente, podemos definir

$$\tilde{H}_2(x, \lambda) := \max_{\mu \in \mathbb{P}_A(x)} \tilde{H}(u; x, \mu, \lambda),$$

y por argumentos similares, sabemos de la existencia de $g_0 \in \mathbb{F}_{\mathbb{P}_B}$ tal que

$$T_\alpha u(x) = \max_{\mu \in \mathbb{P}_A(x)} \tilde{H}(u; x, \mu, g_0(x)).$$

Luego, es claro que

$$\min_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(u; x, f_0(x), \lambda) \leq \tilde{H}(u; x, f_0(x), g_0(x)) \leq \max_{\mu \in \mathbb{P}_A(x)} \tilde{H}(u; x, \mu, g_0(x)),$$

pero como

$$\min_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(u; x, f_0(x), \lambda) = T_\alpha u(x) = \max_{\mu \in \mathbb{P}_A(x)} \tilde{H}(u; x, \mu, g_0(x))$$

se llega a que

$$T_\alpha u(x) = \tilde{H}(u; x, f_0(x), g_0(x))$$

- (c) Como $u \in \mathbb{B}_w(X)$, $\|u\|_w \geq \frac{|u(\cdot)|}{w(\cdot)}$ o bien, $\|u\|_w w(\cdot) \geq |u(\cdot)|$. Por esto, y por (1.1.4.1) y (1.1.4.2), tenemos que para cualquier $(x, a, b) \in \mathbb{K}$,

$$\begin{aligned} |H(u; x, a, b)| &= \left| r(x, a, b) + \alpha \int_{\mathbb{R}} u [F(x, a, b, s)] \rho(s) ds \right| \\ &\leq |r(x, a, b)| + \left| \alpha \int_{\mathbb{R}} u [F(x, a, b, s)] \rho(s) ds \right| \\ &\leq |r(x, a, b)| + \alpha \int_{\mathbb{R}} |u [F(x, a, b, s)]| \rho(s) ds \\ &\leq |r(x, a, b)| + \alpha \int_{\mathbb{R}} \|u\|_w w [F(x, a, b, s)] \rho(s) ds \\ &\leq Mw(x) + \|u\|_w \alpha \int_{\mathbb{R}} w [F(x, a, b, s)] \rho(s) ds \\ &\leq (M + \alpha v \|u\|_w) w(x) \end{aligned} \tag{2.2.1.13}$$

Luego, aprovechando (2.2.1.10) y (2.2.1.13), tenemos que para toda $x \in X$

$$\begin{aligned}
 |T_\alpha u(x)| &= |\tilde{H}(u; x, f_0(x), g_0(x))| \\
 &= \left| \int_{B(x)} \int_{A(x)} H(u; x, a, b)(f_0(x))(da)(g_0(x))(db) \right| \\
 &\leq \int_{B(x)} \int_{A(x)} |H(u; x, a, b)|(f_0(x))(da)(g_0(x))(db) \\
 &\leq \int_{B(x)} \int_{A(x)} (M + \alpha v \|u\|_w) w(x) (f_0(x)(da))(g_0(x)(db)) = (M + \alpha v \|u\|_w) w(x)
 \end{aligned}$$

lo que implica que

$$\sup_{x \in X} \frac{|T_\alpha u(x)|}{w(x)} = \|T_\alpha u\|_w \leq M + \alpha v \|u\|_w < \infty.$$

De aquí se concluye que $T_\alpha u$ está en $B_w(X)$.

■

2.2. Teorema de existencia del valor y del par óptimo para el caso de horizonte finito

Teorema 2.2.1. *Suponga que se cumplen las Hipótesis 1.4. Entonces el juego estocástico GM_F de horizonte finito tiene valor y $J1$ y $J2$ tienen estrategias óptimas de Markov. Más aún, si V_n es la función valor para el juego de n etapas, entonces $V_n \in \mathbb{B}_w(X)$ y $V_n(x) = T_\alpha V_{n-1}(x)$ para cada $n \geq 1$.*

Demostración. Definamos las siguientes funciones:

$$\begin{aligned}
 V_0^*(x) &:= 0 \\
 V_n^*(x) &:= T_\alpha V_{n-1}^*(x) \\
 &= \min_{\lambda \in \mathbb{P}_{B(x)}} \max_{\mu \in \mathbb{P}_{A(x)}} \left\{ \tilde{r}(x, \mu, \lambda) + \alpha \int_{\mathbb{R}} V_{n-1}^* \left[\tilde{F}(x, \mu, \lambda, s) \right] \rho(s) ds \right\}, \quad n \in \mathbb{N}.
 \end{aligned}$$

Como $V_0^* \equiv 0 \in \mathbb{B}_w(X)$, por Lema 2.1.1 (c) se tiene que para cada $n \in \mathbb{N}_0$, $V_n^* \in \mathbb{B}_w(X)$. Entonces, para cada $n \in \mathbb{N}_0$, por Lema 2.1.1 (a),

$$\begin{aligned} V_n^*(x) &= T_\alpha V_{n-1}^*(x) \\ &= \min_{\lambda \in \mathbb{P}_B(x)} \max_{\mu \in \mathbb{P}_A(x)} \tilde{H}(V_{n-1}^*; x, \mu, \lambda) \\ &= \max_{\mu \in \mathbb{P}_A(x)} \min_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(V_{n-1}^*; x, \mu, \lambda) \end{aligned} \quad (2.2.2.1)$$

y por Lema 2.1.1 (b), existen selectores medibles $f_{n-1} \in \mathbb{F}_{\mathbb{P}_A}$ y $g_{n-1} \in \mathbb{F}_{\mathbb{P}_B}$ tales que

$$\begin{aligned} V_n^*(x) &= T_\alpha V_{n-1}^*(x) \\ &= \max_{\mu \in \mathbb{P}_A(x)} \tilde{H}(V_{n-1}^*; x, \mu, g_{n-1}(x)) \\ &= \min_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(V_{n-1}^*; x, f_{n-1}(x), \lambda) \\ &= \tilde{H}(V_{n-1}^*; x, f_{n-1}(x), g_{n-1}(x)). \end{aligned} \quad (2.2.2.2)$$

Ahora para $n \in \mathbb{N}$ definimos

$$\pi_n := \{f_{n-1}, f_{n-2}, \dots, f_0\} \quad (2.2.2.3)$$

$$\gamma_n := \{g_{n-1}, g_{n-2}, \dots, g_0\} \quad (2.2.2.4)$$

donde $f_i \in \mathbb{F}_{\mathbb{P}_A}$ y $g_i \in \mathbb{F}_{\mathbb{P}_B}$ son los respectivos maximizadores y minimizadores para V_{i+1}^* como en (2.2.2.2) para $i = 0, 1, \dots, n-1$.

A continuación probaremos que, para cada n , $V_n^* = L_n = U_n = V_n$, y que π_n y γ_n como en (2.2.2.3) y (2.2.2.4), respectivamente, son estrategias de Markov óptimas para J_1 y J_2 , correspondientemente.

Procederemos por inducción matemática:

- Para $n = 1$, $x_0 = x$ y cualquier par de estrategias arbitrarias $\hat{\pi}_1 = \{\hat{f}_0\}$, $\hat{\gamma}_1 = \{\hat{g}_0\}$, por la definición de J_1 en (1.1.3.2),

$$\begin{aligned} J_1(x, \hat{\pi}_1, \hat{\gamma}_1) &= E_x^{\hat{\pi}_1 \hat{\gamma}_1} [r(x_0, a_0, b_0)] = \dots \\ &= \int_{B(x_0)} \int_{A(x_0)} r(x_0, a_0, b_0) (\hat{f}_0(x_0)(da_0)) (\hat{g}_0(x_0)(db_0)) = \tilde{r}(x, \hat{f}_0(x), \hat{g}_0(x)). \end{aligned} \quad (2.2.2.5)$$

Notemos que ambas estrategias tienen un solo componente que toma la forma $\hat{f}_0 = \mu \in \mathbb{P}_A(x)$ y $\hat{g}_0 = \lambda \in \mathbb{P}_B(x)$. Es decir, en ambos casos las estrategias son un conjunto que consta de una sola medida de probabilidad que considera únicamente el estado inicial del juego. Por esto, cualquier estrategia en este caso es de Markov. Dicho esto, por (2.2.2.5) y por la definición de \tilde{H} en (2.2.1.8),

$$\tilde{H}(0; x, \mu, \lambda) = \tilde{H}(0; x, \hat{f}_0(x), \hat{g}_0(x)) = \tilde{r}(x, \hat{f}_0(x), \hat{g}_0(x)) = J_1(x, \hat{\pi}_1, \hat{\gamma}_1).$$

Combinado lo dicho anteriormente con (2.2.2.1),

$$\begin{aligned} V_1^*(x) &= \max_{\mu \in \mathbb{P}_A} \min_{\lambda \in \mathbb{P}_B} \tilde{H}(0; x, \mu, \lambda) \\ &= \max_{f \in \mathbb{F}_{\mathbb{P}_A}} \min_{g \in \mathbb{F}_{\mathbb{P}_B}} \tilde{H}(0; x, f(x), g(x)) \\ &= \max_{\pi \in \Pi} \min_{\gamma \in \Gamma} J_1(x, \pi, \gamma) \\ &= \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J_1(x, \pi, \gamma) = L_1(x). \end{aligned}$$

Más aún,

$$\begin{aligned} V_1^*(x) &= \min_{\lambda \in \mathbb{P}_B} \max_{\mu \in \mathbb{P}_A} \tilde{H}(0; x, \mu, \lambda) \\ &= \min_{g \in \mathbb{F}_{\mathbb{P}_B}} \max_{f \in \mathbb{F}_{\mathbb{P}_A}} \tilde{H}(0; x, f(x), g(x)) \\ &= \min_{\gamma \in \Gamma} \max_{\pi \in \Pi} J_1(x, \pi, \gamma) \\ &= \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J_1(x, \pi, \gamma) = U_1(x). \end{aligned}$$

Por lo tanto,

$$V_1^*(x) = L_1(x) = U_1(x) = V_1(x),$$

es decir, la función valor para el juego de 1 etapa y es $V_1^*(x)$. Por otro lado, por (2.2.2.2),

$$V_1(x) = \tilde{H}(0; x, f_0(x), g_0(x)) = J_1(x, \pi_1, \gamma_1),$$

lo que implica que (π_1, γ_1) es un par de estrategias óptimas de Markov para el juego de 1 etapa.

- **(Hipótesis de Inducción)** Supongamos que para $n = k - 1$ se cumple que

$$V_{k-1}^*(x) = L_{k-1}(x) = U_{k-1}(x) = V_{k-1}(x) = J_{k-1}(x, \pi_{k-1}, \gamma_{k-1}),$$

es decir, $V_{k-1}^*(x)$ es la función valor para un juego de $k - 1$ etapas y $(\pi_{k-1}, \gamma_{k-1})$ es un par de estrategias de Markov óptimas para este.

- Ahora probaremos que lo anterior se cumple para $n = k$. En efecto, sea $\hat{\gamma}_k = (\hat{g}_{n-1}, \hat{g}_{n-2}, \dots, \hat{g}_0)$ una estrategia arbitraria para J2. Entonces,

$$\begin{aligned} V_k^*(x) &= TV_{k-1}^*(x) = \min_{\lambda \in \mathbb{P}_B(x)} \max_{\mu \in \mathbb{P}_A(x)} \left\{ \tilde{r}(x, \mu, \lambda) + \alpha \int_{\mathbb{R}} V_{k-1}^* \left[\tilde{F}(x, \mu, \lambda, s) \right] \rho(s) ds \right\} \\ &= \min_{\lambda \in \mathbb{P}_B(x)} \left\{ \tilde{r}(x, f_{k-1}(x), \lambda) + \alpha \int_{\mathbb{R}} V_{k-1}^* \left[\tilde{F}(x, f_{k-1}(x), \lambda, s) \right] \rho(s) ds \right\} \\ &\leq \tilde{r}(x, f_{k-1}(x), \hat{g}_{k-1}(x)) + \alpha \int_{\mathbb{R}} V_{k-1}^* \left[\tilde{F}(x, f_{k-1}(x), \hat{g}_{k-1}(x), s) \right] \rho(s) ds \end{aligned}$$

Iteramos esta desigualdad:

$$\begin{aligned} V_k^*(x) &\leq \tilde{r}(x, f_{k-1}(x), \hat{g}_{k-1}(x)) + \alpha \int_{\mathbb{R}} V_{k-1}^* \left[\tilde{F}(x, f_{k-1}(x), \hat{g}_{k-1}(x), s) \right] \rho(s) ds \\ &\leq \tilde{r}(x, f_{k-1}(x), \hat{g}_{k-1}(x)) + \alpha \int_X V_{k-1}^*(y) \tilde{Q}(dy|x, f_{k-1}, \hat{g}_{k-1}) \\ &\leq \tilde{r}(x, f_{k-1}(x), \hat{g}_{k-1}(x)) \\ &\quad + \alpha \int_X \left[\tilde{r}(y, f_{k-2}(y), \hat{g}_{k-2}(y)) + \alpha \int_X V_{k-2}^*(z) \tilde{Q}(dz|x, f_{k-2}, \hat{g}_{k-2}) \right] \tilde{Q}(dy|x, f_{k-1}, \hat{g}_{k-1}) \\ &\leq \tilde{r}(x, f_{k-1}(x), \hat{g}_{k-1}(x)) + \alpha \int_X \tilde{r}(y, f_{k-2}(y), \hat{g}_{k-2}(y)) \tilde{Q}(dy|x, f_{k-1}, \hat{g}_{k-1}) \\ &\quad + \alpha^2 \int_X \int_X V_{k-2}^*(z) \tilde{Q}(dz|x, f_{k-2}, \hat{g}_{k-2}) \tilde{Q}(dy|x, f_{k-1}, \hat{g}_{k-1}) \\ &\leq \tilde{r}(x, f_{k-1}(x), \hat{g}_{k-1}(x)) + \alpha E_x^{\pi_k \hat{\gamma}_k} [r(x_1, a_1, b_1)] \\ &\quad + \alpha^2 \int_X \int_X \left[\tilde{r}(y, f_{k-2}(y), \hat{g}_{k-2}(y)) + \alpha \int_X V_{k-2}^*(z) \tilde{Q}(dz|x, f_{k-2}, \hat{g}_{k-2}) \right] \\ &\quad \quad \quad \tilde{Q}(dz|x, f_{k-2}, \hat{g}_{k-2}) \tilde{Q}(dy|x, f_{k-1}, \hat{g}_{k-1}) \\ &\quad \quad \quad \vdots \\ &\leq \tilde{r}(x, f_{k-1}(x), \hat{g}_k(x)) + \alpha E_x^{\pi_k \hat{\gamma}_k} [r(x_1, a_1, b_1)] + \alpha^2 E_x^{\pi_k \hat{\gamma}_k} [r(x_2, a_2, b_2)] + \\ &\quad \dots + \alpha^k E_x^{\pi_k \hat{\gamma}_k} [r(x_k, a_k, b_k)] \\ &\leq E_x^{\pi_k \hat{\gamma}_k} [r(x_0, a_0, b_0)] + \alpha E_x^{\pi_k \hat{\gamma}_k} [r(x_1, a_1, b_1)] + \alpha^2 E_x^{\pi_k \hat{\gamma}_k} [r(x_2, a_2, b_2)] + \\ &\quad \dots + \alpha^k E_x^{\pi_k \hat{\gamma}_k} [r(x_k, a_k, b_k)], \end{aligned}$$

de modo que se obtiene que

$$V_k^*(x) \leq E_x^{\pi_k \hat{\gamma}_k} \left[\sum_{t=0}^{k-1} \alpha^t r(x_t, a_t, b_t) \right] = J_k(x, \pi_k, \hat{\gamma}_k). \quad (2.2.2.6)$$

Similarmente, para una estrategia arbitraria para J1 $\hat{\pi}_k = (\hat{f}_{k-1}, \hat{f}_{k-2}, \dots, \hat{f}_0)$,

$$V_k^*(x) \geq J_k(x, \hat{\pi}_k, \gamma_k). \quad (2.2.2.7)$$

De (2.2.2.6) se tiene

$$\sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J_k(x, \pi, \gamma) \geq \inf_{\gamma \in \Gamma} J_k(x, \pi_k, \gamma) \geq V_k^*(x), \quad (2.2.2.8)$$

y de (2.2.2.7)

$$\inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J_k(x, \pi, \gamma) \leq \sup_{\pi \in \Pi} J_k(x, \pi, \gamma_k) \leq V_k^*(x). \quad (2.2.2.9)$$

Así, de (2.2.2.8) y (2.2.2.9) obtenemos

$$U_k(x) = \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J_k(x, \pi, \gamma) \leq V_k^*(x) \leq \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J_k(x, \pi, \gamma) = L_k(x),$$

de donde llegamos a que

$$V_k^*(x) = L_k(x) = U_k(x) = V_k(x),$$

es decir, el valor para el juego de n etapas existe y, por la hipótesis de inducción,

$$V_k(x) = TV_{k-1}(x).$$

Por otro lado, por (2.2.2.6) y (2.2.2.7)

$$J_k(x, \pi_k, \hat{\gamma}_k) \geq V_k(x) \geq J_k(x, \hat{\pi}_k, \gamma_k), \quad \forall \hat{\pi}_k \in \Pi, \hat{\gamma}_k \in \Gamma,$$

en particular para $(\pi_k, \gamma_k) \in \Pi \times \Gamma$

$$V_k(x) = J_k(x, \pi_k, \gamma_k),$$

esto es, (π_k, γ_k) es un par de estrategias de Markov óptimas. Por lo tanto,

$$V_k(x) = TV_{k-1}(x) = J_k(x, \pi_k, \gamma_k),$$

lo que completa la prueba.

■

Capítulo 3

Juegos estocásticos de horizonte infinito

En este capítulo estudiaremos el caso de juegos con horizonte infinito. Específicamente mostraremos la existencia del valor del juego, así como un par de estrategias óptimas. Antes de abordar estos problemas, primero introduciremos una serie de resultados que utilizaremos en nuestro estudio.

3.1. Preliminares

Consideremos nuevamente el modelo GM_F introducido en el Capítulo 1 y la función de pago α -descontado esperado $J_\alpha(x, \pi, \gamma)$ como en (1.1.3.1). Las funciones de valor inferior y superior α -descontado son

$$L_\alpha(x) := \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J_\alpha(x, \pi, \gamma),$$
$$U_\alpha(x) := \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J_\alpha(x, \pi, \gamma).$$

Buscamos probar que $L_\alpha(\cdot) = U_\alpha(\cdot)$ de modo que la función valor $V_\alpha(\cdot)$ para el juego α -descontado exista. Para hacer esto, es necesario mostrar propiedades de contracción del correspondiente operador de Shapley T_α .

Definición 3.1.1. Sea (S, d) un espacio métrico. Un mapeo $T : S \rightarrow S$ es de contracción si existe un número real $0 \leq \tau < 1$ tal que

$$d(Ts_1, Ts_2) \leq \tau d(s_1, s_2)$$

para cada $s_1, s_2 \in S$. En este caso τ es llamado **módulo** de T

Proposición 3.1.1. (Teorema del Punto Fijo de Banach) Un mapeo de contracción T en un espacio métrico completo (S, d) tiene un único punto fijo s^* . Más aún, $d(T^n s, s^*) \leq \tau^n d(s, s^*)$ para cada $s \in S$, $n = 0, 1, \dots$, donde τ es el módulo de T , y $T^n := T(T^{n-1})$ para cada $n = 1, 2, \dots$, donde T^0 es el mapeo identidad.

Lema 3.1.1. Bajo las Hipótesis 1.4, el operador T_α definido en (2.2.1.9) es un mapeo de contracción en $\mathbb{B}_w(X)$, con módulo $\tau := \nu\alpha$ (Con ν como en (1.1.4.2)).

Demostración. Para empezar, notemos que T_α es un operador monótono, es decir, si u y \tilde{u} son funciones en $\mathbb{B}_w(X)$, y $u \geq \tilde{u}$, entonces $T_\alpha u(x) \geq T_\alpha \tilde{u}(x)$ para cada $x \in X$. Esto se deduce de las siguientes relaciones:

$$\begin{aligned} u \geq \tilde{u} &\Rightarrow \alpha \int_{\mathbb{R}} u[F(x, a, b, s)]\rho(s)ds \geq \alpha \int_{\mathbb{R}} \tilde{u}[F(x, a, b, s)]\rho(s)ds && \forall (x, a, b) \in \mathbb{K} \\ &\Rightarrow r(x, a, b) + \alpha \int_{\mathbb{R}} u[F(x, a, b, s)]\rho(s)ds \geq r(x, a, b) + \alpha \int_{\mathbb{R}} \tilde{u}[F(x, a, b, s)]\rho(s)ds \\ &&& \forall (x, a, b) \in \mathbb{K} \\ &\Rightarrow H(u; x, a, b) \geq H(\tilde{u}; x, a, b) && \forall (x, a, b) \in \mathbb{K} \\ &\Rightarrow \int_{B(x)} \int_{A(x)} H(u; x, a, b) \mu(da) \lambda(db) \geq \int_{B(x)} \int_{A(x)} H(\tilde{u}; x, a, b) \mu(da) \lambda(db) \\ &&& \forall (x, a, b) \in \mathbb{K} \\ &\Rightarrow \tilde{H}(u; x, \mu, \lambda) \geq \tilde{H}(\tilde{u}; x, \mu, \lambda) \\ &\Rightarrow \sup_{\mu \in \mathbb{P}_{A(x)}} \inf_{\lambda \in \mathbb{P}_{B(x)}} \tilde{H}(u; x, \mu, \lambda) \geq \sup_{\mu \in \mathbb{P}_{A(x)}} \inf_{\lambda \in \mathbb{P}_{B(x)}} \tilde{H}(\tilde{u}; x, \mu, \lambda) \\ &\Rightarrow T_\alpha u(x) \geq T_\alpha \tilde{u}(x). \end{aligned}$$

Por otro lado, por (1.1.4.2), para cada número real $k \geq 0$

$$T_\alpha(u + kw)(x) \leq T_\alpha u(x) + \nu\alpha kw(x) \quad \forall x \in X, u \in \mathbb{B}_w(X). \quad (3.3.1.1)$$

En efecto:

$$\begin{aligned}
 T_\alpha(u + kw)(x) &= \max_{\mu \in \mathbb{P}_{A(x)}} \min_{\lambda \in \mathbb{P}_{B(x)}} \tilde{H}(u + kw; x, \mu, \lambda) \\
 &= \max_{\mu \in \mathbb{P}_{A(x)}} \min_{\lambda \in \mathbb{P}_{B(x)}} \int_{B(x)} \int_{A(x)} H(u + kw; x, a, b) \mu(da) \lambda(db) \\
 &= \max_{\mu \in \mathbb{P}_{A(x)}} \min_{\lambda \in \mathbb{P}_{B(x)}} \int_{B(x)} \int_{A(x)} \left(r(x, a, b) + \alpha \int_{\mathbb{R}} (u + kw) [F(x, a, b, s)] \rho(s) ds \right) \mu(da) \lambda(db) \\
 &= \max_{\mu \in \mathbb{P}_{A(x)}} \min_{\lambda \in \mathbb{P}_{B(x)}} \int_{B(x)} \int_{A(x)} \left(r(x, a, b) + \alpha \int_{\mathbb{R}} u [F(x, a, b, s)] \rho(s) ds \right. \\
 &\quad \left. + \alpha \int_{\mathbb{R}} kw [F(x, a, b, s)] \rho(s) ds \right) \mu(da) \lambda(db) \\
 &\leq \max_{\mu \in \mathbb{P}_{A(x)}} \min_{\lambda \in \mathbb{P}_{B(x)}} \int_{B(x)} \int_{A(x)} \left(r(x, a, b) + \alpha \int_{\mathbb{R}} u [F(x, a, b, s)] \rho(s) ds \right. \\
 &\quad \left. + \nu \alpha kw(x) \right) \mu(da) \lambda(db) \\
 &\leq \max_{\mu \in \mathbb{P}_{A(x)}} \min_{\lambda \in \mathbb{P}_{B(x)}} \int_{B(x)} \int_{A(x)} \left(r(x, a, b) + \alpha \int_{\mathbb{R}} u [F(x, a, b, s)] \rho(s) ds \right) \mu(da) \lambda(db) \\
 &\quad + \nu \alpha kw(x) \\
 &\leq \max_{\mu \in \mathbb{P}_{A(x)}} \min_{\lambda \in \mathbb{P}_{B(x)}} \tilde{H}(u; x, \nu, \lambda) + \nu \alpha kw(x) \\
 &\leq T_\alpha u(x) + \nu \alpha kw(x).
 \end{aligned}$$

Ahora, para verificar que T_α es una contracción, escogemos arbitrariamente u y \tilde{u} en $\mathbb{B}_w(X)$. Como $u \leq \tilde{u} + w \|u - \tilde{u}\|_w$, por la monotonicidad de T_α y (3.3.1.1) con $k = \|u - \tilde{u}\|_w$, se sigue que

$$T_\alpha u(x) \leq T_\alpha(\tilde{u} + kw)(x) \leq T_\alpha \tilde{u}(x) + \nu \alpha kw(x),$$

es decir,

$$T_\alpha u(x) - T_\alpha \tilde{u}(x) \leq \nu \alpha \|u - \tilde{u}\|_w w(x).$$

Si ahora intercambiamos u y \tilde{u} obtenemos

$$T_\alpha u(x) - T_\alpha \tilde{u}(x) \geq -\nu \alpha \|u - \tilde{u}\|_w w(x),$$

de manera que

$$|T_\alpha u(x) - T_\alpha \tilde{u}(x)| \leq \nu \alpha \|u - \tilde{u}\|_w w(x).$$

Por tanto, tomando $\tau := \nu \alpha$, obtenemos que $\|T_\alpha u - T_\alpha \tilde{u}\|_w \leq \tau \|u - \tilde{u}\|_w$, lo cual demuestra lo que se quería. ■

Lema 3.1.2. Sean M , w y ν como en las Hipótesis 1.4, sean $\pi \in \Pi$ y $\gamma \in \Gamma$ estrategias arbitrarias para $J1$ y $J2$, respectivamente, y $x \in X$ un estado inicial. Entonces para cada $t = 0, 1, \dots$

- (a) $E_x^{\pi\gamma}[w(x_t)] \leq \nu^t w(x)$,
- (b) $|E_x^{\pi\gamma}[r(x_t, a_t, b_t)]| \leq M \nu^t w(x)$, y
- (c) $\lim_{t \rightarrow \infty} \alpha^t E_x^{\pi\gamma}[u(x_t)] = 0$ para cada $u \in \mathbb{B}_w(X)$.

Demostración. (a) Para $t = 0$ se sigue directamente. Como $x_0 = x$ es dado

$$E_x^{\pi\gamma}[w(x)] = w(x) = \nu^0 w(x).$$

Ahora, si $t \geq 1$, por Hipótesis 1.4(e) y la propiedad de Markov de $P_x^{\pi\gamma}$ se tiene

$$\begin{aligned} E_x^{\pi\gamma}[w(x_t)|h_{t-1}, a_{t-1}, b_{t-1}] &= \int_{\mathbb{R}} w[F(x_{t-1}, a_{t-1}, b_{t-1}, s)] \rho(s) ds \\ &\leq \nu w(x_{t-1}). \end{aligned}$$

Por lo tanto, aplicando la propiedad de la doble esperanza,

$$E_x^{\pi\gamma}[E_x^{\pi\gamma}[w(x_t)|h_{t-1}, a_{t-1}, b_{t-1}]] \leq E_x^{\pi\gamma}[\nu w(x_{t-1})],$$

es decir,

$$E_x^{\pi\gamma}[w(x_t)] \leq \nu E_x^{\pi\gamma}[w(x_{t-1})].$$

Iterando esta desigualdad obtenemos

$$E_x^{\pi\gamma}[w(x_t)] \leq \nu E_x^{\pi\gamma}[w(x_{t-1})] \leq \nu^2 E_x^{\pi\gamma}[w(x_{t-2})] \leq \dots \leq \nu^t E_x^{\pi\gamma}[w(x_0)] = \nu^t E_x^{\pi\gamma}[w(x)] = \nu^t w(x),$$

llegando finalmente a (a).

(b) Observemos que por Hipótesis 1.4(d) se tiene que

$$|r(x_t, a_t, b_t)| \leq Mw(x_t) \quad \forall t = 0, 1, \dots,$$

de manera que, por (a),

$$E_x^{\pi\gamma}[|r(x_t, a_t, b_t)|] \leq E_x^{\pi\gamma}[Mw(x_t)] = ME_x^{\pi\gamma}[w(x_t)] \leq M\nu^t w(x)$$

concluyendo que (b) se cumple.

(c) Por definición de w -norma y por la parte (a) de este Lema, tenemos

$$\alpha^t E_x^{\pi\gamma}[u(x_t)] \leq \alpha^t E_x^{\pi\gamma}[|u(x_t)|] \leq \alpha^t \|u\|_w E_x^{\pi\gamma}[w(x_t)] \leq \alpha^t \|u\|_w \nu^t w(x) = (\alpha\nu)^t \|u\|_w w(x),$$

de donde, como $0 < \alpha \leq \alpha\nu < 1$,

$$\lim_{t \rightarrow \infty} (\alpha\nu)^t \|u\|_w w(x) = 0$$

Análogamente podemos obtener que

$$\alpha^t E_x^{\pi\gamma}[u(x_t)] \geq -(\alpha\nu)^t \|u\|_w w(x)$$

y que

$$\lim_{t \rightarrow \infty} -(\alpha\nu)^t \|u\|_w w(x) = 0.$$

Así,

$$0 = \lim_{t \rightarrow \infty} -(\alpha\nu)^t \|u\|_w w(x) \leq \lim_{t \rightarrow \infty} \alpha^t E_x^{\pi\gamma}[u(x_t)] \leq \lim_{t \rightarrow \infty} (\alpha\nu)^t \|u\|_w w(x) = 0$$

por lo que (c) se cumple

■

Definición 3.1.2. Sean $f \in \mathbb{F}_{\mathbb{P}_A}$, $g \in \mathbb{F}_{\mathbb{P}_B}$, y sea \tilde{H} como en (2.2.1.8). Definimos el operador

$$R_{fg} : \mathbb{B}_w(X) \rightarrow \mathbb{B}_w(X), \quad u \mapsto R_{fg}u,$$

como

$$R_{fg}u(x) := \tilde{H}(u; x, f(x), g(x)) \quad \forall x \in X. \quad (3.3.1.2)$$

Lema 3.1.3. *El operador R_{fg} es un operador contracción en $\mathbb{B}_w(X)$ y $J_\alpha(x, f, g)$ es su único punto fijo en $\mathbb{B}_w(X)$.*

Demostración. El hecho de que R_{fg} es un operador de contracción en $\mathbb{B}_w(X)$ con módulo $\tau := \alpha\nu$ se sigue al utilizar argumentos similares a los del Lema 3.1.1:

Por lo tanto, R_{fg} tiene un único punto fijo u_{fg} en $\mathbb{B}_w(X)$, es decir,

$$u_{fg} = R_{fg}u_{fg} \quad (3.3.1.3)$$

De (3.3.1.2) y (3.3.1.3) tenemos entonces que u_{fg} es la única solución en $\mathbb{B}_w(X)$ de la ecuación

$$\begin{aligned} u_{fg}(x) &= \tilde{r}(x, f(x), g(x)) + \alpha \int_{\mathbb{R}} u_{fg}[\tilde{F}(x, f(x), g(x))] \rho(s) ds, \quad \forall x \in X \\ &= \tilde{r}(x, f(x), g(x)) + \alpha \int_X u_{fg}(y) Q(dy|x, f(x), g(x)), \quad \forall x \in X. \end{aligned} \quad (3.3.1.4)$$

Además, al iterar (3.3.1.3) y (3.3.1.4) se cumple que

$$u_{fg}(x) = R_{fg}^n u_{fg}(x) = E_x^{fg} \left[\sum_{t=0}^{n-1} \alpha^t r(x_t, f(x_t), g(x_t)) \right] + \alpha^n E_x^{fg} [u_{fg}(x_n)]$$

para toda $x \in X$ y $n \geq 1$, donde $E_x^{fg}[u(x_n)] = \int_X u(y) Q^n(dy|x, f, g)$, y $Q^n(\cdot|x, f, g)$ es el kernel de transición de n pasos del proceso de Markov $\{x_t\}$ al usar f y g . Finalmente, por el Lema 3.1.2 (c) y tomando $n \rightarrow \infty$, vemos que por (1.1.3.1), $u_{fg}(x) = J_\alpha(x, f, g)$ para cada $x \in X$. ■

3.2. Teorema de existencia del valor y del par óptimo para el caso de horizonte infinito

Para presentar el resultado principal de este capítulo, primero debemos recordar del Teorema 2.2.1 que

$$V_n(x) = T_\alpha V_{n-1}(x)$$

para cada $n \geq 1$ y $x \in X$, con $V_0(\cdot) \equiv 0$. Esto es, de la Definición 1.3.3,

$$\begin{aligned} V_n(x) &= \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J_n(x, \pi, \gamma) \\ &= \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J_n(x, \pi, \gamma) \quad \forall x \in X. \end{aligned}$$

Ahora consideraremos el caso cuando $n \rightarrow \infty$. El siguiente teorema nos dice, entre otras cosas, que la sucesión $\{V_n\}$ converge geoméricamente a V_α en la w -norma.

Teorema 3.2.1. *Supongamos que las Hipótesis 1.4 se cumplen. Sea ν la constante en (1.1.4.2), y definamos $\tau := \alpha\nu$. Entonces:*

(a) *La función de valor α -descontado V_α es la única función en $\mathbb{B}_w(X)$ que satisface $T_\alpha V_\alpha = V_\alpha$, y*

$$\|V_n - V_\alpha\|_w \leq \frac{M\tau^n}{1 - \tau} \quad \forall n = 1, 2, \dots$$

(b) *Existe un par de estrategias óptimas.*

Demostración. Por Lema 3.1.1 y el Teorema de Punto Fijo de Banach (Proposición 3.1.1), T_α tiene un único punto fijo V^* en $\mathbb{B}_w(X)$, es decir,

$$T_\alpha V^* = V^*$$

y

$$\|T_\alpha^n u - V^*\|_w \leq \tau^n \|u - V^*\|_w \quad \forall u \in \mathbb{B}_w(X), n = 0, 1, \dots \quad (3.3.2.1)$$

Por lo tanto, para probar (a) y (b) tenemos que mostrar que

- (i) V^* está en $\mathbb{B}_w(X)$, con norma $\|V^*\|_w \leq \frac{M}{1-\tau}$, y
- (ii) $V_\alpha = V^*$.

Para probar (i), sean $\pi \in \Pi$ y $\gamma \in \Gamma$ estrategias arbitrarias para J1 y J2, respectivamente, y sea $x \in X$ un estado inicial arbitrario, entonces (i) se cumple del Lema 3.1.2 dado que

$$\alpha^t E_x^{\pi\gamma} |r(x_t, a_t, b_t)| \leq M \alpha^t \nu^t w(x), \quad \forall t \in \mathbb{N}_0,$$

lo cual implica

$$\sum_{t=0}^{\infty} \alpha^t E_x^{\pi\gamma} |r(x_t, a_t, b_t)| = \sum_{t=0}^{\infty} E_x^{\pi\gamma} |\alpha^t r(x_t, a_t, b_t)| \leq \sum_{t=0}^{\infty} M(\alpha\nu)^t w(x) = \frac{Mw(x)}{1-\alpha\nu} = \frac{Mw(x)}{1-\tau}$$

con $\tau := \alpha\nu$. Ahora, recordando que

$$|J_\alpha(x, \pi, \gamma)| \leq E_x^{\pi\gamma} \left[\sum_{t=0}^{\infty} \alpha^t |r(x_t, a_t, b_t)| \right] = \sum_{t=0}^{\infty} E_x^{\pi\gamma} |\alpha^t r(x_t, a_t, b_t)|$$

y puesto que $\pi \in \Pi$ y $\gamma \in \Gamma$ son arbitrarios, tenemos

$$|V^*(x)| \leq \frac{Mw(x)}{1-\tau}, \quad \forall x \in X,$$

es decir,

$$\frac{|V^*(x)|}{w(x)} \leq \frac{M}{1-\tau}, \quad \forall x \in X.$$

Entonces

$$\sup_{x \in X} \frac{|V^*(x)|}{w(x)} \leq \frac{M}{1-\tau},$$

por lo cual se cumple

$$\|V^*\|_w \leq \frac{M}{1-\tau}$$

donde $\frac{M}{1-\tau}$ es finito, lo que implica que $V^* \in \mathbb{B}_w(X)$. Más aún, $\tau^n \|V^*\|_w \leq \frac{M\tau^n}{1-\tau}$. Luego, utilizando que

$$V_n = T_\alpha V_{n-1} = T_\alpha^n V_0 \quad \forall n = 0, 1, \dots, \quad V_0 = 0,$$

y por (3.3.2.1) con $u \equiv 0$, llegamos a que

$$\|V_n - V^*\|_w = \|T_\alpha^n 0 - V^*\|_w \leq \tau^n \|0 - V^*\|_w = \tau^n \| - V^*\|_w = \tau^n \|V^*\|_w \leq \frac{M\tau^n}{1-\tau},$$

donde por transitividad

$$\|V_n - V^*\|_w \leq \frac{M\tau^n}{1-\tau}.$$

Para probar (ii), notemos que por la igualdad $T_\alpha V^* = V^*$ y por Lema 2.1.1, existen $f_* \in \mathbb{F}_{\mathbb{P}_A}$ y $g_* \in \mathbb{F}_{\mathbb{P}_B}$ tales que, para cada $x \in X$,

$$V^*(x) = \sup_{\mu \in \mathbb{P}_A(x)} \tilde{H}(V^*; x, \mu, g_*(x)) \tag{3.3.2.2}$$

$$= \inf_{\lambda \in \mathbb{P}_B(x)} \tilde{H}(V^*; x, f_*(x), \lambda) \tag{3.3.2.3}$$

$$= \tilde{H}(V^*; x, f_*(x), g_*(x)). \tag{3.3.2.4}$$

Observemos que (3.3.2.4) puede ser escrita como

$$\begin{aligned} V^*(x) &= \tilde{r}(x, f_*(x), g_*(x)) + \alpha \int_{\mathbb{R}} V^* \left[\tilde{F}(x, f_*(x), g_*(x), s) \right] \rho(s) ds \\ &= \tilde{r}(x, f_*(x), g_*(x)) + \alpha \int_X V^*(x_1) \tilde{Q}(dx_1 | x, f_*(x), g_*(x)). \end{aligned}$$

Ahora bien, del Lema 3.1.3 se sigue que $V^*(x) = J_\alpha(x, f_*, g_*)$. Por lo tanto, y a su vez considerando (3.3.2.2), se tiene

$$J_\alpha(x, f_*, g_*) = \sup_{\mu \in \mathbb{P}_A(x)} \left[\tilde{r}(x, \mu, g_*(x)) + \alpha \int_X J_\alpha(x_1, f_*(x), g_*(x)) \tilde{Q}(dx_1 | x, \mu, g_*) \right]$$

para toda $x \in X$. Aquí podemos hacer uso resultados estándar de programación dinámica, de modo que llegaremos a que

$$J_\alpha(x, f_*, g_*) = \sup_{\pi \in \Pi} J_\alpha(x, \pi, g_*).$$

Similarmente, pero ahora considerando el ínfimo como en (3.3.2.3) tenemos

$$J_\alpha(x, f_*, g_*) = \inf_{\gamma \in \Gamma} J_\alpha(x, f_*, \gamma).$$

Consecuentemente

$$J_\alpha(x, f_*, g_*) = \sup_{\pi \in \Pi} J_\alpha(x, \pi, g_*) \geq \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J_\alpha(x, \pi, \gamma).$$

Además,

$$J_\alpha(x, f_*, g_*) = \inf_{\gamma \in \Gamma} J_\alpha(x, f_*, \gamma) \leq \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J_\alpha(x, \pi, \gamma).$$

Por lo tanto,

$$\sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} J_\alpha(x, \pi, \gamma) = J_\alpha(x, f_*, g_*) = \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} J_\alpha(x, \pi, \gamma),$$

es decir,

$$L_\alpha(x) = J_\alpha(x, f_*, g_*) = U_\alpha(x),$$

lo que implica que la función $V_\alpha(x)$ existe, y que además

$$V_\alpha(x) = V^*(x) = J_\alpha(x, f_*, g_*)$$

para toda $x \in X$, donde f_*, g_* son las estrategias óptimas para J1 y J2 respectivamente. ■

Capítulo 4

Juegos Lineales - Cuadráticos (Linear-Quadratic LQ)

En este capítulo se estudiará una clase particular de juegos estocásticos de suma cero, los llamados juegos-LQ. La característica principal de estos juegos es que el proceso de estados $\{x_t\}$ evoluciona de acuerdo a una ecuación en diferencias lineal, y la función de pago por etapa es cuadrática. Específicamente demostraremos que los juegos-LQ satisfacen las Hipótesis 1.4, y por lo tanto podemos garantizar que existe tanto el valor del juego como el par de estrategias óptimas. De hecho, mediante un proceso recursivo, las calcularemos explícitamente. Concluimos el capítulo presentando algunos gráficos que ilustran todos los elementos que hemos visto para esta clase de juegos.

4.1. Definición, Motivación y Revisión de Hipótesis en Juegos-LQ

Los juegos lineales-cuadráticos, o mejor conocidos como juegos-LQ son juegos dinámicos que evolucionan de acuerdo a una ecuación lineal y tienen un pago cuadrático. Para fines de esta tesis, consideraremos que el proceso de estados $\{x_t\}$ evoluciona como

$$x_{t+1} = F(x_t, a_t, b_t, \varepsilon_t) = x_t + a_t + b_t + \varepsilon_t \quad \text{para } t \in \mathbb{N}_0 \quad (4.4.1.1)$$

donde $x_0 = x$ es dado, $X = A = B = \mathbb{R}$ y $\{\varepsilon_t\}$ es una sucesión de variables aleatorias i.i.d que tienen una distribución normal con media cero y varianza $\sigma^2 \in (0, 3)$, esto es

$$\rho(s) := \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) \quad \forall s \in \mathbb{R}, \quad E(\varepsilon_t) = 0 \quad \text{y} \quad \sigma^2 = E(\varepsilon_t^2) \quad \forall t \in \mathbb{N}_0.$$

Los conjuntos de acciones admisibles para J1 y J2 son $A(x) = B(x) = \left[-\frac{|x|}{2}, \frac{|x|}{2}\right]$, y la función de pago r es la función cuadrática

$$r(x, a, b) = x^2 - a^2 + b^2 \tag{4.4.1.2}$$

Podemos notar que este juego-LQ es un tipo particular de juego GM_F . Una aplicación del mismo es en los llamados procesos de rastreo. En efecto, supongamos que el proceso de estados $\{x_t\}$ evoluciona como (4.4.1.1) y sea x^* un estado fijo. El objetivo de J1 es elegir acciones $a \in A(x)$ tal que la trayectoria $\{x_t\}$ esté lo más lejos posible de x^* , mientras que el objetivo de J2 es el de tomar acciones que acerquen la trayectoria $\{x_t\}$ a x^* . Entonces, una posible forma de la función de pago r que modela esta situación es

$$r(x, a, b) = \varphi(d(x, x^*)) - a^2 + b^2,$$

donde d es una métrica en X y φ una función monótona creciente. En particular, dado que $X = \mathbb{R}$, si tomamos $x^* = 0$, $d(x, x^*) = |x - x^*|$ y $\varphi(x) = x^2$, entonces r toma la forma (4.4.1.2).

Notemos que es posible ver que este juego-LQ cumple con lo establecido en las Hipótesis 1.4:

a) Para cada estado $x \in X$, los conjuntos no vacíos $A(x)$ y $B(x)$ de acciones admisibles son compactos.

$A(x)$ y $B(x)$ fueron definidos como intervalos cerrados y acotados, por lo que son conjuntos compactos no vacíos.

b) Para cada $(x, a, b) \in \mathbb{K}$, $r(x, \cdot, b)$ es semi-continua superiormente en $A(x)$, y $r(x, a, \cdot)$ es semi-continua inferiormente en $B(x)$.

Es claro que $r(x, a, b) = x^2 - a^2 + b^2$ cumple que $r(x, \cdot, b)$ es continua en $A(x)$ y por tanto semi-continua superiormente en el conjunto, y además $r(x, a, \cdot)$ es continua en $B(x)$ y en consecuencia es semi-continua inferiormente en el mismo.

c) Para cada $(x, a, b) \in \mathbb{K}$ y cada función $v \in M_b(X)$, las funciones

$$\int_{\mathbb{R}} v [F(x, \cdot, b, s)] \rho(s) ds \quad \text{y} \quad \int_{\mathbb{R}} v [F(x, a, \cdot, s)] \rho(s) ds$$

son continuas en $A(x)$ y $B(x)$, respectivamente.

Veamos que para cualquier función acotada y medible v en \mathbb{R} ,

$$\int_{\mathbb{R}} v [F(x, a, b, s)] \rho(s) ds = \int_{\mathbb{R}} v(x + a + b + s) \rho(s) ds = \int_{\mathbb{R}} v(y) \rho(y - x - a - b) dy$$

al considerar el cambio de variable $y = x + a + b + s$. Ahora consideremos una sucesión en los reales $\{a_n\}$ tal que $a_n \rightarrow a$. Dado que ρ es la densidad de una distribución normal, es una función continua. Por ende

$$\rho(y - x - a_n - b) \rightarrow \rho(y - x - a - b).$$

Además, como v es acotada, existe una $m \geq 0$ tal que $|v(y)| \leq m$ para toda $y \in \mathbb{R}$.

Dicho lo anterior, veamos que

$$\begin{aligned}
 0 &\leq \left| \int_{\mathbb{R}} v(y)\rho(y-x-a_n-b)dy - \int_{\mathbb{R}} v(y)\rho(y-x-a-b)dy \right| \\
 &\leq \left| \int_{\mathbb{R}} v(y) \left(\rho(y-x-a_n-b) - \rho(y-x-a-b) \right) dy \right| \\
 &\leq \int_{\mathbb{R}} |v(y)| \left| \rho(y-x-a_n-b) - \rho(y-x-a-b) \right| dy \\
 &\leq \int_{\mathbb{R}} |v(y)| \left| \rho(y-x-a_n-b) - \rho(y-x-a-b) \right| dy \\
 &\leq m \int_{\mathbb{R}} \left| \rho(y-x-a_n-b) - \rho(y-x-a-b) \right| dy,
 \end{aligned}$$

donde, por el Teorema de Scheffé (Véase C.0.1),

$$\int_{\mathbb{R}} \left| \rho(y-x-a_n-b) - \rho(y-x-a-b) \right| dy \rightarrow 0.$$

Así,

$$\left| \int_{\mathbb{R}} v(y)\rho(y-x-a_n-b)dy - \int_{\mathbb{R}} v(y)\rho(y-x-a-b)dy \right| \rightarrow 0,$$

o bien,

$$\int_{\mathbb{R}} v(y)\rho(y-x-a_n-b)dy \rightarrow \int_{\mathbb{R}} v(y)\rho(y-x-a-b)dy,$$

implicando así la continuidad de la primera integral de nuestra hipótesis. Para verificar la continuidad en la segunda integral se realiza un procedimiento análogo.

d) Existe una constante $M > 0$ y una función medible $w : X \rightarrow \mathbb{R}$ tal que $w(x) \geq 1$ para cada $x \in X$ y

$$|r(x, a, b)| \leq Mw(x) \quad \forall (x, a, b) \in \mathbb{K},$$

y, además, (c) se cumple al sustituir v por w .

Observemos que para $a, b \in \left[-\frac{|x|}{2}, \frac{|x|}{2}\right]$ tenemos

$$\begin{aligned} |r(x, a, b)| &= r(x, a, b) = x^2 - a^2 + b^2 \\ &\leq x^2 + a^2 + b^2 \\ &\leq x^2 + \frac{x^2}{4} + \frac{x^2}{4} = \frac{3}{2}x^2 \\ &\leq \frac{3}{2}(x^2 + 1) = M(x^2 + 1). \end{aligned}$$

Tomando $M = \frac{3}{2}$ y $w(x) = x^2 + 1$ se cumple la primera parte de nuestra hipótesis. Para la segunda parte, recordamos que debe existir una sucesión creciente de funciones acotadas $\{w_n\}$ tal que $w_n \nearrow w$ para toda $x \in \mathbb{R}$. Por lo dicho en la hipótesis anterior, para cualquier $n \in \mathbb{N}$,

$$\int_{\mathbb{R}} w_n [F(x, a, b, s)] \rho(s) ds$$

es continua en $A(x)$ y $B(x)$. En particular,

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} w_n [F(x, a, b, s)] \rho(s) ds$$

es continua en ambos conjuntos. Sin embargo, aprovechamos el Teorema de la Convergencia Monótona (Véase C.0.2) para observar que

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} w_n [F(x, a, b, s)] \rho(s) ds = \int_{\mathbb{R}} \lim_{n \rightarrow \infty} w_n [F(x, a, b, s)] \rho(s) ds = \int_{\mathbb{R}} w [F(x, a, b, s)] \rho(s) ds,$$

de donde concluimos que la hipótesis anterior se cumple al sustituir v por w .

e) Existe una constante $1 \leq \nu \leq \frac{1}{\alpha}$ tal que

$$\int_{\mathbb{R}} w [F(x, a, b, s)] \rho(s) ds \leq \nu w(x) \quad \forall (x, a, b) \in \mathbb{K}$$

Tomando nuevamente $w(x) = x^2 + 1$, para cada $(x, a, b) \in \mathbb{K}$ tenemos

$$\begin{aligned}
\int_{\mathbb{R}} w [F(x, a, b, s)] \rho(s) ds &= \int_{\mathbb{R}} w(x + a + b + s) \rho(s) ds \\
&= \int_{\mathbb{R}} [(x + a + b + s)^2 + 1] \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \\
&= \int_{\mathbb{R}} (x + a + b + s)^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \\
&\quad + \int_{\mathbb{R}} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \\
&= \int_{\mathbb{R}} (x + a + b + s)^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds + 1 \\
&= \int_{\mathbb{R}} y^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y - (x + a + b))^2}{2\sigma^2}\right) dy + 1 \\
&= (x + a + b)^2 + \sigma^2 + 1 \\
&\leq 4x^2 + 4 = 4w(x)
\end{aligned}$$

donde la quinta igualdad se obtiene al realizar un cambio de variable $y = x + a + b + s$, dejándonos una integral que correspondería a la del segundo momento de una variable aleatoria que se distribuye normalmente con media $x + a + b$ y varianza σ^2 . De lo anterior, sabemos que nuestra hipótesis cumple al tomar $\nu = 4$ y cualquier factor de descuento $\alpha < \frac{1}{4}$.

Dicho todo lo anterior, podemos asegurar que los resultados obtenidos en los capítulos 2 y 3 de esta tesis se cumplen para este juego-LQ, es decir, este tiene un valor y además existen estrategias de Markov óptimas para J1 y J2 que nos permiten alcanzarlo.

4.2. Valor y Estrategias Óptimas para Juegos-LQ

Como se mencionó anteriormente, este juego-LQ consta de una función valor para la el juego en n etapas $V_n(x)$. Esto implica que

$$L_n(x) = V_n(x) = U_n(x) \quad \forall n \in \mathbb{N}_0.$$

Observemos que la base fundamental para la obtención de los resultados en los capítulos 2 y 3 de esta tesis es la aplicación del Teorema de Ky Fan. En particular, el considerar los conjuntos $\mathbb{P}_A(x)$ y $\mathbb{P}_B(x)$ de medidas y estrategias aleatorizadas nos permitió aplicar dicho Teorema ya que, como se expuso, las hipótesis necesarias para su utilización se cumplen para $\tilde{H}(u; x, \mu, \gamma)$ al considerar cualesquier $u \in \mathbb{B}_w(X)$, $\mu \in \mathbb{P}_A(x)$ y $\lambda \in \mathbb{P}_B(x)$. Sin embargo, en este caso particular de un juego-LQ, las hipótesis del Teorema de Fan se satisfacen sin necesidad de considerar los espacios de medidas $\mathbb{P}_A(x)$ y $\mathbb{P}_B(x)$. Esto es, $H(u; x, a, b)$ satisface todas las condiciones necesarias para aplicar el Teorema mencionado directamente sin importar como sean las $u \in \mathbb{B}_w(X)$, $a \in A(x)$ y $b \in B(x)$ que se contemplen. Es por esto que, aún cuando la función de valor de acuerdo a lo establecido en el capítulo 2 tendría la forma

$$\begin{aligned} V_n(x) &= \min_{\lambda \in \mathbb{P}_B(x)} \max_{\mu \in \mathbb{P}_A(x)} \tilde{H}(V_{n-1}; x, \mu, \lambda) \\ &= \min_{\lambda \in \mathbb{P}_B(x)} \max_{\mu \in \mathbb{P}_A(x)} \left[\int_{B(x)} \int_{A(x)} H(V_{n-1}; x, a, b) \mu(da) \lambda(db) \right] \\ &= \min_{\lambda \in \mathbb{P}_B(x)} \max_{\mu \in \mathbb{P}_A(x)} \left[\int_{B(x)} \int_{A(x)} \left(x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} V_{n-1}(x + a + b + s) \rho(s) ds \right) \mu(da) \lambda(db) \right], \end{aligned}$$

podemos obtener que, a partir de que se tome el operador de Shapley como

$$T_\alpha u(x) = \max_{a \in A(x)} \min_{b \in B(x)} H(u; x, a, b), \quad \forall u \in \mathbb{B}_w(X)$$

para este caso particular y de un desarrollo análogo al realizado en el capítulo mencionado, la función de valor puede ser expresada como

$$\begin{aligned} V_n(x) &= \min_{b \in B(x)} \max_{a \in A(x)} H(V_{n-1}; x, a, b) \\ &= \min_{b \in B(x)} \max_{a \in A(x)} \left[x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} V_{n-1}(x + a + b + s) \rho(s) ds \right], \end{aligned} \quad (4.4.2.1)$$

donde $V_0(x) = 0$. De aquí observaremos que es posible trabajar con lo que llamaremos acciones óptimas, que consisten en una clase particular de estrategias aleatorizadas denominadas puras o deterministas. La principal característica de estas estrategias es que concentran toda la probabilidad en valores específicos, en este caso en aquellos que nos permiten alcanzar el punto silla de H . Denotaremos por a_n^* y b_n^* a las acciones óptimas de J1 y J2, respectivamente, en la etapa n del juego.

Ahora, veamos que

$$H(V_0; x, a, b) = x^2 - a^2 + b^2,$$

de modo que

$$V_1(x) = \min_{b \in B(x)} \max_{a \in A(x)} [x^2 - a^2 + b^2], \quad x \in \mathbb{R}. \quad (4.4.2.2)$$

Encontrando que el único punto crítico de $H(V_{n-1}; x, a, b)$ (considerando a y b variables) es $(0, 0)$, verificando que este se trata de un punto silla y sustituyendo en (4.4.2.2), obtenemos que

$$V_1(x) = x^2$$

y por lo tanto $a_1^*(x) = b_1^*(x) = 0$.

Ahora observemos que

$$\begin{aligned} H(V_1; x, a, b) &= x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} V_1(x + a + b + s) \rho(s) ds \\ &= x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} (x + a + b + s)^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \\ &= x^2 - a^2 + b^2 + \alpha \left((x + a + b)^2 + \sigma^2 \right) \\ &= x^2 - a^2 + b^2 + \alpha(x^2 + a^2 + b^2 + 2ax + 2bx + 2ab + \sigma^2) \\ &= (1 + \alpha)x^2 + (-1 + \alpha)a^2 + (1 + \alpha)b^2 + 2\alpha ax + 2\alpha bx + 2\alpha ab + \alpha\sigma^2. \end{aligned}$$

Entonces

$$V_2(x) = \min_{b \in B(x)} \max_{a \in A(x)} [(1 + \alpha)x^2 + (-1 + \alpha)a^2 + (1 + \alpha)b^2 + 2\alpha ax + 2\alpha bx + 2\alpha ab + \alpha\sigma^2]. \quad (4.4.2.3)$$

Buscando los puntos críticos de $H(V_1; x, a, b)$ y empleando el criterio de la segunda derivada parcial podemos obtener que $(\alpha x, -\alpha x)$ es el único punto crítico y que además es punto silla. Por lo tanto, es posible sustituir en (4.4.2.3) para obtener que

$$V_2(x) = (\alpha + 1)x^2 + \alpha\sigma^2.$$

y que $a_2^*(x) = \alpha x$ y $b_2^*(x) = -\alpha x$.

Similarmente,

$$\begin{aligned} H(V_2; x, a, b) &= x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} V_2(x + a + b + s)\rho(s)ds \\ &= x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} \left((\alpha + 1)(x + a + b + s)^2 + \alpha\sigma^2 \right) \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \\ &= x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} (\alpha + 1)(x + a + b + s)^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \\ &\quad + \alpha \int_{\mathbb{R}} \alpha\sigma^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \\ &= x^2 - a^2 + b^2 + \alpha(\alpha + 1) \left((x + a + b)^2 + \sigma^2 \right) + \alpha^2\sigma^2 \\ &= [x^2 - a^2 + b^2 + \alpha(\alpha + 1)(x^2 + a^2 + b^2 + 2ax + 2bx + 2ab + \sigma^2) + \alpha^2\sigma^2] \\ &= \left(1 + \alpha(\alpha + 1)\right)x^2 + \left(-1 + \alpha(\alpha + 1)\right)a^2 + \left(1 + \alpha(\alpha + 1)\right)b^2 + 2\alpha(\alpha + 1)ax \\ &\quad + 2\alpha(\alpha + 1)bx + 2\alpha(\alpha + 1)ab + \left(\alpha^2 + \alpha(\alpha + 1)\right)\sigma^2, \end{aligned}$$

de forma que

$$\begin{aligned} V_3(x) &= \min_{b \in B(x)} \max_{a \in A(x)} \left[\left(1 + \alpha(\alpha + 1)\right)x^2 + \left(-1 + \alpha(\alpha + 1)\right)a^2 + \left(1 + \alpha(\alpha + 1)\right)b^2 + 2\alpha(\alpha + 1)ax \right. \\ &\quad \left. + 2\alpha(\alpha + 1)bx + 2\alpha(\alpha + 1)ab + \left(\alpha^2 + \alpha(\alpha + 1)\right)\sigma^2 \right]. \end{aligned} \quad (4.4.2.4)$$

Consistente a lo que se ha estado haciendo, vemos que $\left((1 + \alpha)x, -(1 + \alpha)x\right)$ es el único punto crítico de $H(V_2; x, a, b)$ y verificamos que este se trata de un punto silla, de modo que podemos sustituir en (4.4.2.4) para obtener que

$$V_3(x) = \left(1 + \alpha(\alpha + 1)\right)x^2 + \left(\alpha^2 + \alpha(\alpha + 1)\right)\sigma^2.$$

y además concluir que $a_3^*(x) = (1 + \alpha)x$ y $b_3^*(x) = -(1 + \alpha)x$.

Ahora bien, probaremos por inducción el siguiente resultado.

Teorema 4.2.1. Para el juego-LQ que cumple con las ecuaciones (4.4.1.1) y (4.4.1.2) se tiene que, para $t \in \mathbb{N}$:

$$V_t(x) = \frac{\beta_t}{\alpha} x^2 + \delta_t \sigma^2, \quad (4.4.2.5)$$

y

$$a_t^*(x) = \beta_{t-1} x \quad b_t^*(x) = -\beta_{t-1} x \quad (4.4.2.6)$$

donde $\beta_0 = 0$,

$$\beta_t = \alpha(1 + \beta_{t-1}) = \frac{\alpha - \alpha^{t+1}}{1 - \alpha}$$

y

$$\delta_t = \beta_{t-1} + \alpha \delta_{t-1} = \sum_{i=1}^{t-1} \alpha^{t-i-1} \beta_i.$$

Demostración. ■ Veamos que para $t = 1$,

$$a_1^*(x) = \beta_0 x = 0 \quad b_1^* = -\beta_0 x = 0$$

y

$$\begin{aligned} V_1(x) &= \left(\frac{\alpha - \alpha^{1+1}}{1 - \alpha} \right) \frac{x^2}{\alpha} + \left(\sum_{i=1}^{1-1} \alpha^{t-i-1} \beta_i \right) \sigma^2 \\ &= \left(\frac{\alpha - \alpha^2}{1 - \alpha} \right) \frac{x^2}{\alpha} + \left(\sum_{i=1}^0 \alpha^{t-i-1} \beta_i \right) \sigma^2 \\ &= \frac{\alpha}{\alpha} x^2 + (0) \sigma^2 \\ &= x^2, \end{aligned}$$

coincidiendo con lo obtenido con anterioridad. Además, verificando para $t = 2$ y $t = 3$ podremos observar que se obtiene lo previamente calculado para $a_2^*(x), b_2^*(x), V_2(x)$ y $a_3^*(x), b_3^*(x), V_3(x)$, respectivamente.

- **(Hipótesis de Inducción)** Supongamos que lo que buscamos probar es cierto para $t = k - 1$ (con $k \geq 2$), es decir,

$$V_{k-1}(x) = \frac{\beta_{k-1}}{\alpha}x^2 + \delta_{k-1}\sigma^2$$

y

$$a_{k-1}^*(x) = \beta_{k-2}x \quad b_{k-1}^*(x) = -\beta_{k-2}x$$

- Vemos que

$$H(V_{k-1}; x, a, b) = x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} V_{k-1}(x + a + b + s)\rho(s)ds,$$

y al aplicar la hipótesis de inducción

$$\begin{aligned} H(V_{k-1}; x, a, b) &= x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} \left(\frac{\beta_{k-1}}{\alpha}(x + a + b + s)^2 + \delta_{k-1}\sigma^2 \right) \rho(s)ds \\ &= x^2 - a^2 + b^2 \\ &\quad + \alpha \int_{\mathbb{R}} \left(\frac{\beta_{k-1}}{\alpha}(x + a + b + s)^2 + \delta_{k-1}\sigma^2 \right) \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \\ &= x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} \frac{\beta_{k-1}}{\alpha}(x + a + b + s)^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \\ &\quad + \alpha \int_{\mathbb{R}} \delta_{k-1}\sigma^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \\ &= x^2 - a^2 + b^2 + \beta_{k-1} \left((x + a + b)^2 + \sigma^2 \right) + \alpha\delta_{k-1}\sigma^2 \\ &= x^2 - a^2 + b^2 + \beta_{k-1}(x^2 + a^2 + b^2 + 2ax + 2bx + 2ab + \sigma^2) + \alpha\delta_{k-1}\sigma^2 \\ &= (1 + \beta_{k-1})x^2 + (-1 + \beta_{k-1})a^2 + (1 + \beta_{k-1})b^2 + 2\beta_{k-1}ax \\ &\quad + 2\beta_{k-1}bx + 2\beta_{k-1}ab + \beta_{k-1}\sigma^2 + \alpha\delta_{k-1}\sigma^2, \end{aligned}$$

de donde

$$\begin{aligned} V_k(x) &= \min_{b \in B(x)} \max_{a \in A(x)} [(1 + \beta_{k-1})x^2 + (-1 + \beta_{k-1})a^2 + (1 + \beta_{k-1})b^2 + 2\beta_{k-1}ax \\ &\quad + 2\beta_{k-1}bx + 2\beta_{k-1}ab + \beta_{k-1}\sigma^2 + \alpha\delta_{k-1}\sigma^2], \quad (4.4.2.7) \end{aligned}$$

Ahora, calculamos el punto crítico para $H(V_{k-1}; x, a, b)$:

$$\begin{aligned} \frac{\partial H(V_{k-1}; x, a, b)}{\partial a} &= 0 \\ \Rightarrow 2(-1 + \beta_{k-1})a + 2\beta_{k-1}x + 2\beta_{k-1}b &= 0 \\ \Rightarrow (-1 + \beta_{k-1})a + \beta_{k-1}x + \beta_{k-1}b &= 0, \end{aligned} \quad (4.4.2.8)$$

y análogamente

$$\begin{aligned} \frac{\partial H(V_{k-1}; x, a, b)}{\partial b} &= 0 \\ \Rightarrow 2(1 + \beta_{k-1})b + 2\beta_{k-1}x + 2\beta_{k-1}a &= 0 \\ \Rightarrow (1 + \beta_{k-1})b + \beta_{k-1}x + \beta_{k-1}a &= 0. \end{aligned} \quad (4.4.2.9)$$

Igualando (4.4.2.8) y (4.4.2.9)

$$\begin{aligned} (-1 + \beta_{k-1})a + \beta_{k-1}x + \beta_{k-1}b &= (1 + \beta_{k-1})b + \beta_{k-1}x + \beta_{k-1}a \\ \Rightarrow -a + \beta_{k-1}a + \beta_{k-1}x + \beta_{k-1}b &= b + \beta_{k-1}a + \beta_{k-1}x + \beta_{k-1}b \\ \Rightarrow -a &= b. \end{aligned}$$

Sustituyendo en (4.4.2.8) o en (4.4.2.9) llegamos a que $(\beta_{k-1}x, -\beta_{k-1}x)$ es el único punto crítico de $H(V_{k-1}; x, a, b)$, y dado que

$$\begin{aligned} \det \begin{pmatrix} \frac{\partial^2 H(V_{k-1}; x, a, b)}{\partial a^2} & \frac{\partial^2 H(V_{k-1}; x, a, b)}{\partial a \partial b} \\ \frac{\partial^2 H(V_{k-1}; x, a, b)}{\partial a \partial b} & \frac{\partial^2 H(V_{k-1}; x, a, b)}{\partial b^2} \end{pmatrix} &= \det \begin{pmatrix} -1 + \beta_{k-1} & \beta_{k-1} \\ \beta_{k-1} & 1 + \beta_{k-1} \end{pmatrix} \\ &= (-1 + \beta_{k-1}^2) - (\beta_{k-1}^2) \\ &= -1 \\ &< 0 \end{aligned}$$

sabemos que este se trata de un punto silla. Sustituyendo $a = \beta_{k-1}x$ y $b = -\beta_{k-1}x$ en (4.4.2.7) para tener el mínimo sobre $B(x)$ y el máximo sobre $A(x)$ se obtiene que

$$\begin{aligned} V_k(x) &= (1 + \beta_{k-1})x^2 + (-1 + \beta_{k-1})(\beta_{k-1}x)^2 + (1 + \beta_{k-1})(-\beta_{k-1}x)^2 + 2\beta_{k-1}(\beta_{k-1}x)x \\ &\quad + 2\beta_{k-1}(-\beta_{k-1}x)x + 2\beta_{k-1}(\beta_{k-1}x)(-\beta_{k-1}x) + \beta_{k-1}\sigma^2 + \beta_{k-1}\delta_{k-1}\sigma^2 \\ &= (1 + \beta_{k-1})x^2 + (\beta_{k-1} + \alpha\delta_{k-1})\sigma^2 \\ &= \frac{\beta_k}{\alpha}x^2 + \delta_k\sigma^2 \end{aligned}$$

Adicionalmente, tomando $\alpha < \frac{1}{4}$ conforme a lo dicho en la comprobación de las hipótesis para este juego-LQ,

$$\beta_t = \frac{\alpha - \alpha^{t+1}}{1 - \alpha} < \frac{\alpha}{1 - \alpha} \leq \frac{1}{3} < \frac{1}{2}$$

de manera que para toda $t \in \mathbb{N}_0$

$$a_t^*(x) \in A(x) \quad b_t^*(x) \in B(x)$$

donde $A(x) = B(x) = \left[-\frac{|x|}{2}, \frac{|x|}{2}\right]$, confirmando que para cada una de las fases del juego, la función valor y las acciones óptimas para los jugadores son consistentes con lo establecido en (4.4.2.5) y (4.4.2.6)

■

Observación 4.2.1. *Es importante no confundirse con las funciones de valor y las de acciones óptimas para J1 y J2. El valor para los juegos en t etapas está en función del estado inicial del juego. Por otro lado, las acciones óptimas para J1 y J2 se toman de acuerdo a una función que considera el estado del juego al tiempo t , siendo consistente con lo establecido en el capítulo 2, donde se dijo que las estrategias óptimas para el caso de horizonte finito son de Markov.*

Finalmente, estableceremos el resultado principal para el caso de horizonte infinito.

Teorema 4.2.2. *Para el juego-LQ que cumple con las ecuaciones (4.4.1.1) y (4.4.1.2) con horizonte infinito, la función de valor y las acciones óptimas están dadas como:*

$$V_\alpha(x) = \frac{1}{1 - \alpha}x^2 + \frac{\alpha}{(1 - \alpha)^2}\sigma^2,$$

y

$$a_*(x) = \frac{\alpha}{1 - \alpha}x \quad b_*(x) = -\frac{\alpha}{1 - \alpha}x$$

Demostración. Primero observemos que para toda $t \in \mathbb{N}_0$,

$$\begin{aligned}
\delta_t &= \sum_{i=1}^{t-1} \alpha^{t-i-1} \beta_i \\
&= \sum_{i=1}^{t-1} \alpha^{t-i-1} \left(\frac{\alpha - \alpha^{i+1}}{1 - \alpha} \right) \\
&= \sum_{i=1}^{t-1} \frac{\alpha^{t-i} - \alpha^t}{1 - \alpha} \\
&= \frac{1}{1 - \alpha} \sum_{i=1}^{t-1} \alpha^{t-i} - \frac{(t-1)\alpha^t}{1 - \alpha} \\
&= \frac{1}{1 - \alpha} \sum_{j=1}^{t-1} \alpha^j - \frac{(t-1)\alpha^t}{1 - \alpha} \\
&= \frac{1}{1 - \alpha} \left(\frac{\alpha - \alpha^t}{1 - \alpha} \right) - \frac{(t-1)\alpha^t}{1 - \alpha} \\
&= \frac{\alpha - \alpha^t}{(1 - \alpha)^2} - \frac{(t-1)\alpha^t}{1 - \alpha}
\end{aligned}$$

de donde es sencillo observar que, nuevamente para $\alpha < \frac{1}{4}$,

$$\lim_{t \rightarrow \infty} \beta_t = \frac{\alpha}{1 - \alpha} \quad \text{y} \quad \lim_{t \rightarrow \infty} \delta_t = \frac{\alpha}{(1 - \alpha)^2}.$$

Esto implica

$$\lim_{t \rightarrow \infty} V_t(x) = \frac{1}{1 - \alpha} x^2 + \frac{\alpha}{(1 - \alpha)^2} \sigma^2.$$

Luego, por lo demostrado en el capítulo 3, sabemos que se cumple la convergencia con la w -norma, $w = x^2 + 1$, esto es,

$$\lim_{t \rightarrow \infty} \|V_t - V_\alpha\|_w \leq \lim_{t \rightarrow \infty} \frac{M\tau^t}{1 - \tau} = 0$$

de donde se puede obtener que

$$V_\alpha(x) = \frac{1}{1 - \alpha} x^2 + \frac{\alpha}{(1 - \alpha)^2} \sigma^2.$$

Finalmente, definimos

$$a_*(x) := \lim_{t \rightarrow \infty} a_t^*(x) = \lim_{t \rightarrow \infty} \beta_{t-1}x = \frac{\alpha}{1-\alpha}x \in A(x) \quad (4.4.2.10)$$

$$b_*(x) := \lim_{t \rightarrow \infty} b_t^*(x) = \lim_{t \rightarrow \infty} -\beta_{t-1}x = -\frac{\alpha}{1-\alpha}x \in B(x) \quad (4.4.2.11)$$

y verificamos que cumple la relación establecida en Teorema 3.2.1 (a):

$$\begin{aligned} T_\alpha V_\alpha(x) &= \min_{b \in B(x)} \max_{a \in A(x)} \left[x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} V_\alpha(x + a + b + s) \rho(s) ds \right] \\ &= \min_{b \in B(x)} \max_{a \in A(x)} \left[x^2 - a^2 + b^2 \right. \\ &\quad \left. + \alpha \int_{\mathbb{R}} \left(\frac{1}{1-\alpha} (x + a + b + s)^2 + \frac{\alpha}{(1-\alpha)^2} \sigma^2 \right) \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \right] \\ &= \min_{b \in B(x)} \max_{a \in A(x)} \left[x^2 - a^2 + b^2 + \alpha \int_{\mathbb{R}} \frac{1}{1-\alpha} (x + a + b + s)^2 \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \right. \\ &\quad \left. + \alpha \int_{\mathbb{R}} \frac{\alpha}{(1-\alpha)^2} \sigma^2 \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds \right] \\ &= \min_{b \in B(x)} \max_{a \in A(x)} \left[x^2 - a^2 + b^2 + \frac{\alpha}{1-\alpha} \left((x + a + b)^2 + \sigma^2 \right) + \frac{\alpha^2}{(1-\alpha)^2} \sigma^2 \right] \\ &= \min_{b \in B(x)} \max_{a \in A(x)} \left[x^2 - a^2 + b^2 + \frac{\alpha}{1-\alpha} (x^2 + a^2 + b^2 + 2ax + 2bx + 2ab + \sigma^2) + \frac{\alpha^2}{(1-\alpha)^2} \sigma^2 \right] \\ &= \min_{b \in B(x)} \max_{a \in A(x)} \left[\frac{1}{1-\alpha} x^2 + \frac{2\alpha-1}{1-\alpha} a^2 + \frac{1}{1-\alpha} b^2 + \frac{2\alpha}{1-\alpha} ax + \frac{2\alpha}{1-\alpha} bx + \frac{2\alpha}{1-\alpha} ab \right. \\ &\quad \left. + \frac{\alpha}{1-\alpha} \sigma^2 + \frac{\alpha^2}{(1-\alpha)^2} \sigma^2 \right]. \end{aligned}$$

Finalmente, sustituyendo de acuerdo a (4.4.2.10) y (4.4.2.11),

$$\begin{aligned} T_\alpha V_\alpha(x) &= \frac{1}{1-\alpha} x^2 + \frac{2\alpha-1}{1-\alpha} \left(\frac{\alpha}{1-\alpha} x \right)^2 + \frac{1}{1-\alpha} \left(-\frac{\alpha}{1-\alpha} x \right)^2 + \frac{2\alpha}{1-\alpha} \left(-\frac{\alpha}{1-\alpha} x \right) x \\ &\quad + \frac{2\alpha}{1-\alpha} \left(\frac{\alpha}{1-\alpha} x \right) x + \frac{2\alpha}{1-\alpha} \left(-\frac{\alpha}{1-\alpha} x \right) \left(\frac{\alpha}{1-\alpha} x \right) + \frac{\alpha}{1-\alpha} \sigma^2 + \frac{\alpha^2}{(1-\alpha)^2} \sigma^2 \\ &= \frac{1}{1-\alpha} x^2 + \frac{\alpha}{(1-\alpha)^2} \sigma^2 \\ &= V_\alpha(x). \end{aligned}$$

■

4.3. Gráficas

A continuación, se presentan algunas gráficas que ilustran como varía el valor del juego en las fases $t = 1, 2, 3, 10$ y cuando $t \rightarrow \infty$ con respecto al estado inicial x :

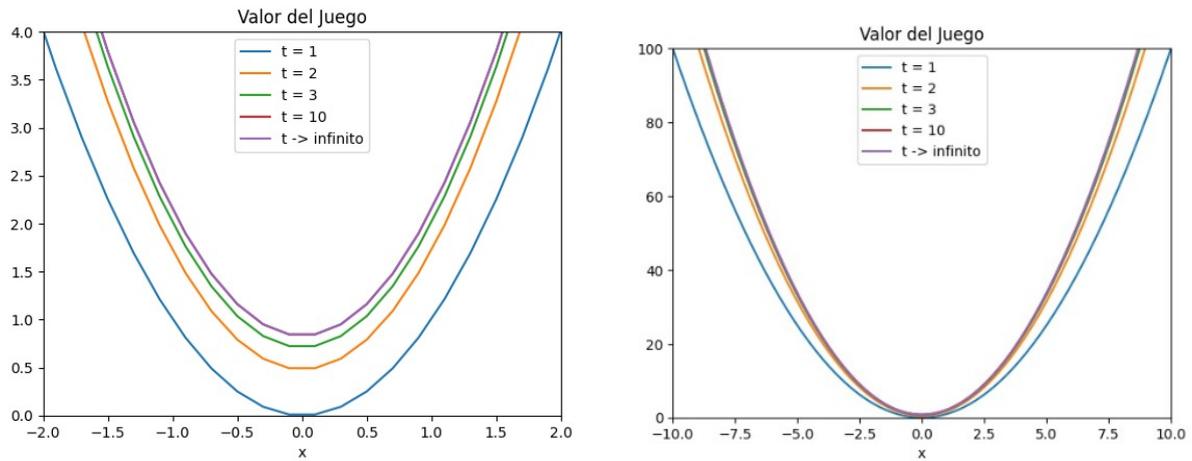


FIGURA 4.1: Se fijan $\alpha = 0.24$ y $\sigma^2 = 2$

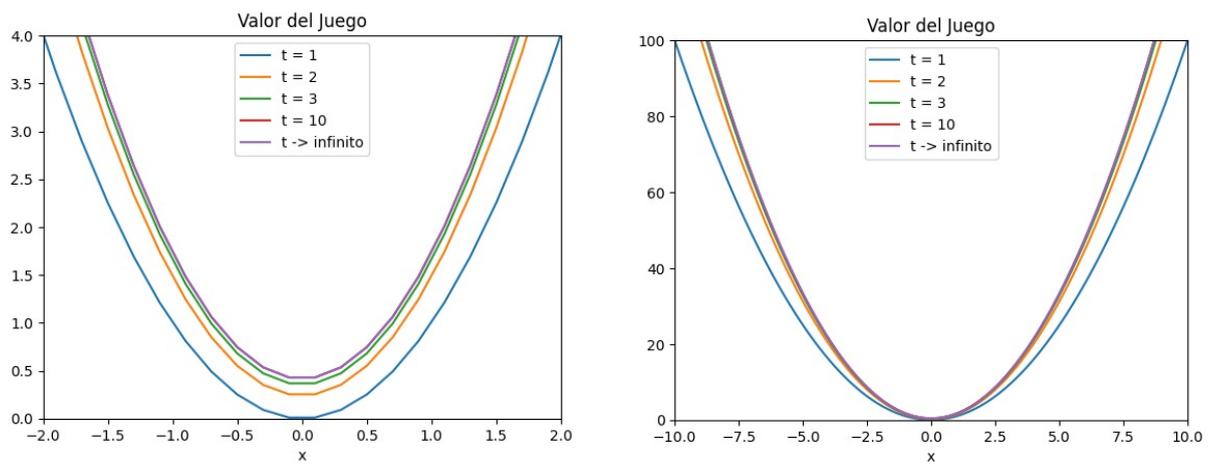


FIGURA 4.2: Se fijan $\alpha = 0.24$ y $\sigma^2 = 1$

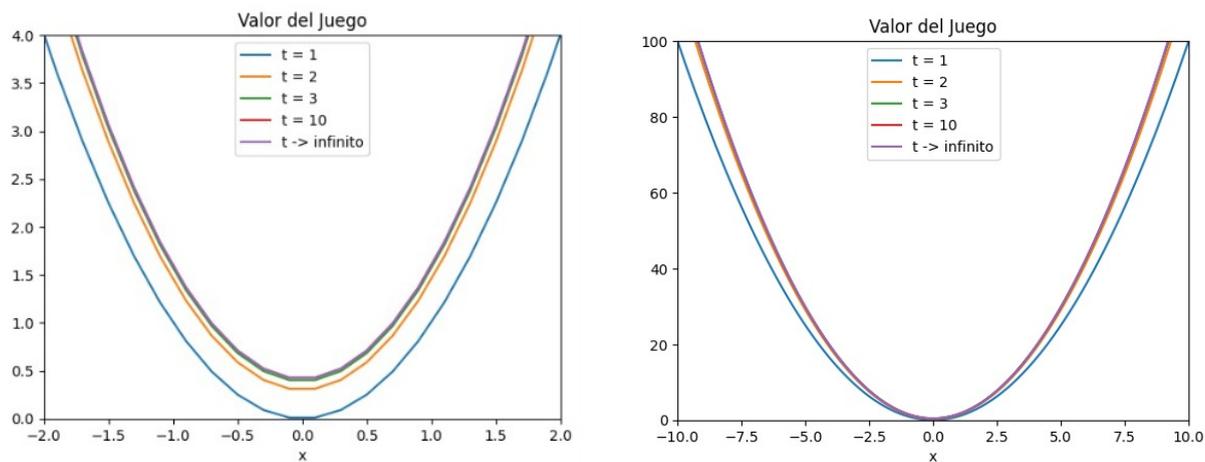


FIGURA 4.3: Se fijan $\alpha = 0.15$ y $\sigma^2 = 2$

Ahora se muestran un par de gráficas donde podemos ver a $H(V_0; 10, a, b)$ y $H(V_1; 5, a, b)$ con sus respectivos punto silla.

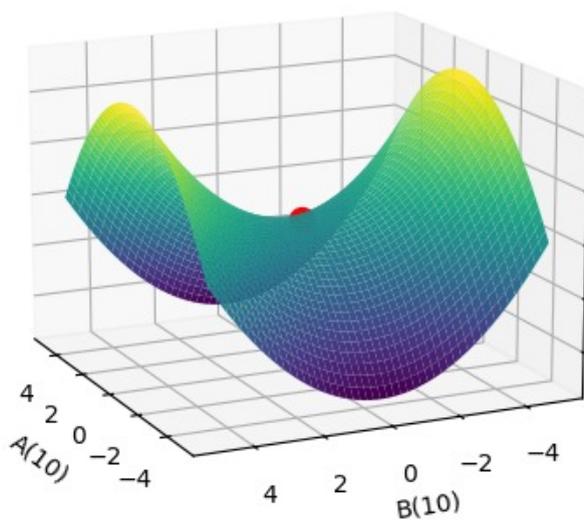


FIGURA 4.4: $H(V_0; 10, a, b)$ y punto $(0, 0, 100)$
(Se fijan $\alpha = 0.24$ y $\sigma^2 = 2$)

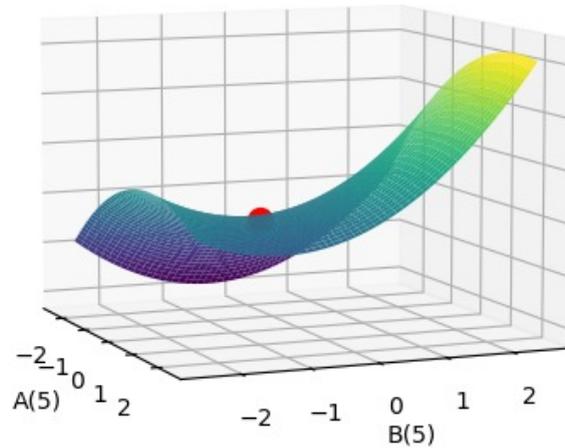


FIGURA 4.5: $H(V_1; 5, a, b)$ y punto $(1.2, -1.2, 33.24)$
(Se fijan $\alpha = 0.2$ y $\sigma^2 = 1.5$)

Finalmente, se presentan las gráficas de una serie de simulaciones de cuatro juegos-LQ (como los estudiados en este capítulo) distintos, cuya característica común es el ruido aleatorio presente en cada etapa. En cada par de gráficas, la de la izquierda ilustra el cambio de estado en cada etapa según la dinámica del sistema dada por (4.4.1.1). La gráfica de la derecha muestra el pago acumulado α -descontado en cada etapa según (4.4.1.2).

Las representaciones gráficas incluyen:

- **En azul:** Muestra las trayectorias de los estados en cada etapa y los pagos acumulados α -descontados cuando J1 y J2 siguen sus respectivas estrategias óptimas.
- **En amarillo:** Muestra las trayectorias de los estados y los pagos acumulados α -descontados cuando J1 sigue sus estrategias óptimas y J2 escoge acciones $b_t \in B(x_t)$ al azar con distribución uniforme.
- **En verde:** Muestra las trayectorias de los estados y los pagos acumulados α -descontados cuando J2 sigue sus estrategias óptimas y J1 escoge acciones $a_t \in A(x_t)$ al azar con distribución uniforme.
- **En rojo:** En las gráficas de la izquierda, la línea roja indica el estado inicial, mientras que en las gráficas de la derecha, el rojo representa el valor del juego en cada etapa.

Fijando diferentes estados iniciales, valores para α y para σ^2 , las simulaciones nos arrojaron los siguientes resultados:

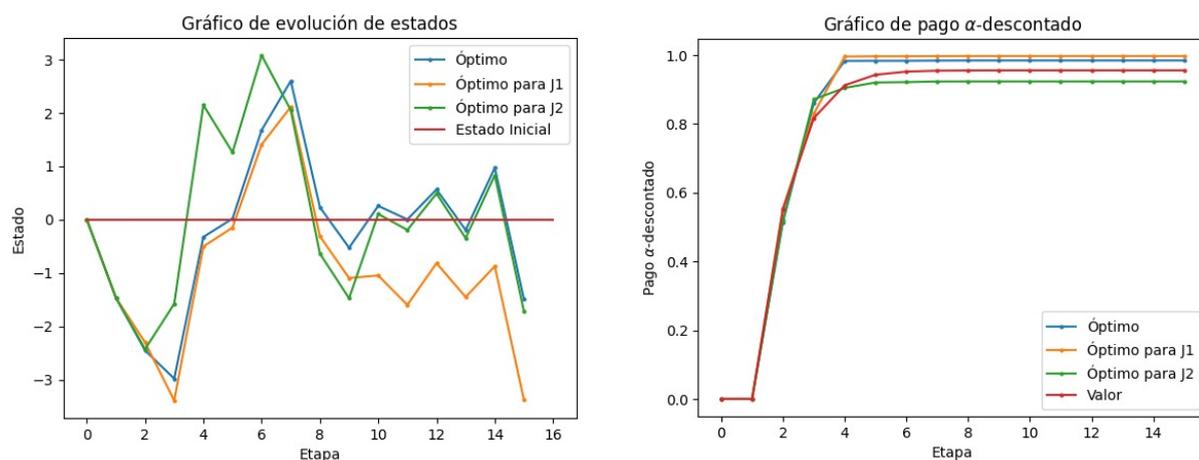


FIGURA 4.6: Se fijan $x_0 = 0$, $\alpha = 0.24$ y $\sigma^2 = 2.3$

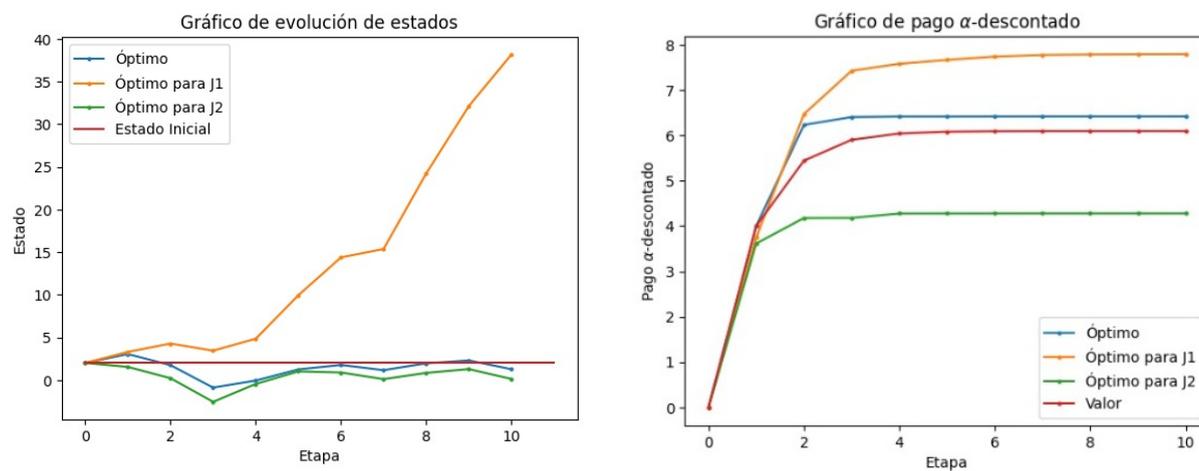


FIGURA 4.7: Se fijan $x_0 = 2$, $\alpha = 0.24$ y $\sigma^2 = 2$

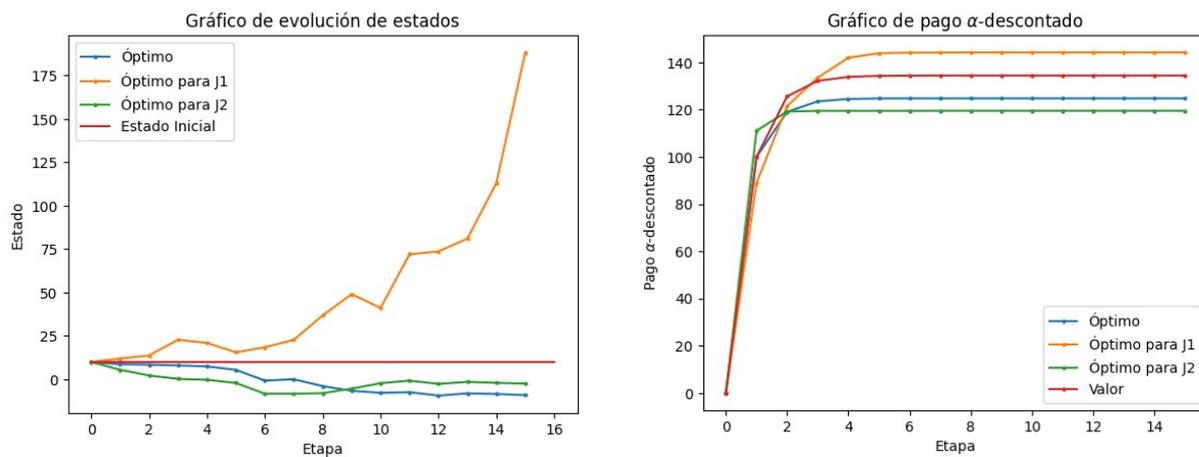


FIGURA 4.8: Se fijan $x_0 = 10$, $\alpha = 0.2499$ y $\sigma^2 = 2.5$

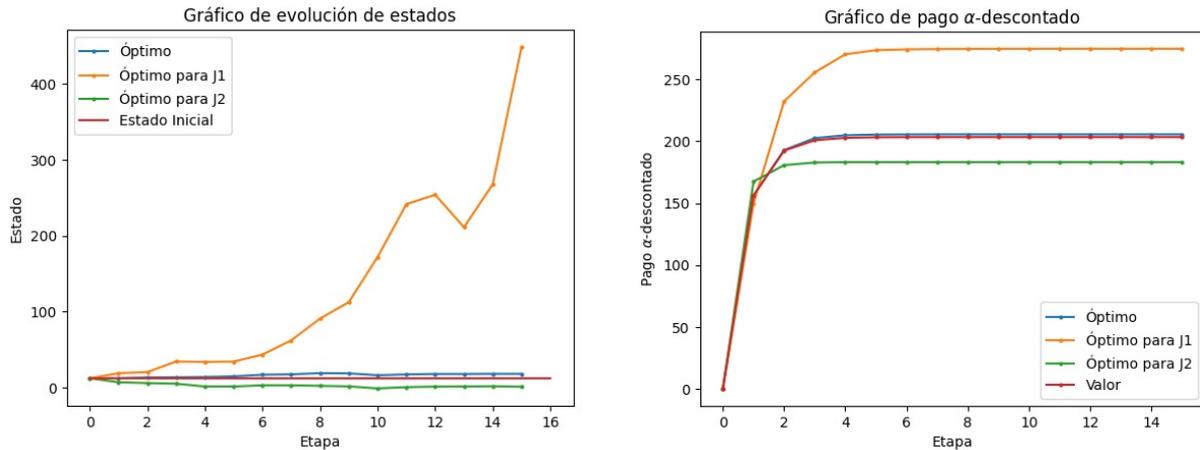


FIGURA 4.9: Se fijan $x_0 = 12.5$, $\alpha = 0.23$ y $\sigma^2 = 1.2$

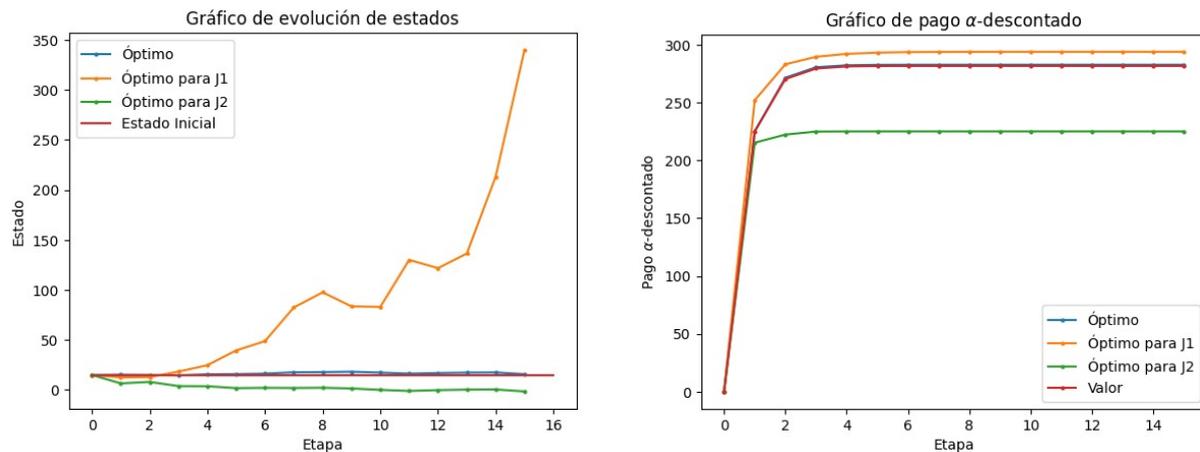


FIGURA 4.10: Se fijan $x_0 = 15$, $\alpha = 0.2$ y $\sigma^2 = 1$

Estos gráficos son especialmente útiles para comprender los conceptos de valor del juego en n etapas y las estrategias óptimas para J1 y J2. En las gráficas de la derecha se puede observar que cuando J1 sigue su estrategia óptima, obtiene un pago acumulado α -descontado mayor, lo cual se evidencia al ver que las trayectorias amarillas y azules están por encima de las verdes. Del mismo modo, cuando J2 sigue su estrategia óptima, su pérdida, representada por el pago acumulado α -descontado, es menor. Esto se refleja en que las trayectorias verdes y azules están por debajo de las amarillas. Además, el valor del juego, representado por las trayectorias rojas, se encuentra entre las trayectorias amarillas y verdes, lo cual es coherente con la idea de que el valor del juego es la menor de las ganancias esperadas para J1 y la mayor de las pérdidas esperadas para J2 al emplear cada uno sus estrategias óptimas. Finalmente, se puede observar que la trayectoria azul suele acercarse más a la roja, lo que es consistente con la idea de que, al seguir ambos sus estrategias óptimas, el comportamiento esperado será igual al valor del juego en cada etapa.

Apéndice A

Funciones semi-continuas

Sea X un espacio métrico. Una función $v : X \rightarrow \mathbb{R}$ se dice ser **semi-continua inferiormente** en X si $\{x \in X : v(x) > a\}$ es abierto en X para cada $a \in \mathbb{R} \cup \{+\infty\}$. Por otro lado, v es llamada **semi-continua superiormente** en X si $-v$ es semi-continua inferiormente. Note que v es continua si y solo si es tanto **semi-continua inferiormente** y **semi-continua superiormente**.

Tenemos el siguiente criterio para semi-continuidad.

Teorema A.0.1. *La función v es semi-continua inferiormente en X si y solo si, para cada sucesión $\{x_n\}$ en X tal que converge a $x \in X$, tenemos*

$$\liminf_{n \rightarrow \infty} v(x_n) \geq v(x).$$

Así, v es semi-continua superiormente si y solo si

$$\limsup_{n \rightarrow \infty} v(x_n) \leq v(x) \quad \text{cuando } x_n \rightarrow x$$

Demostración. Ver [1], página 388 ■

Teorema A.0.2. *Sea v semi-continua inferiormente en un espacio métrico topológico compacto. Entonces v alcanza su ínfimo. Por lo tanto, si v es semi-continua superiormente en un espacio compacto X , entonces v alcanza su supremo.*

Demostración. Ver [1], página 389 ■

Apéndice B

Teoremas minimax y selectores

Sea X un espacio métrico separable dotado con la σ -álgebra de Borel $\mathcal{B}(X)$.

Lema B.0.1. *Si u es semi-continua inferiormente y acotada inferiormente en X , entonces el mapeo $\bar{u} : \mathbb{P}(X) \rightarrow \mathbb{R}$ definido por*

$$\bar{u}(\mu) := \int u d\mu$$

es semi-continua inferiormente en $\mathbb{P}(X)$.

Demostración. Ver Lema 6.3 en [3] ■

Observación B.0.1. *Si una función u es semi-continua superiormente y acotada superiormente, entonces por Lema B.0.1,*

$$\bar{u}(\mu) := \int u d\mu$$

es semi-continua superiormente en $\mathbb{P}(X)$.

Teorema B.0.1. *Sea F una multifunción valuada en los compactos de X a A . Entonces los siguientes enunciados son equivalentes:*

- i) F es Borel medible.*
- ii) GrF es un subconjunto de Borel de X en A .*

Demostración. Ver [5] ■

Teorema B.0.2. Sean X y A espacios de Borel, sea F una multifunción valuada en los compactos Borel medible de X a A , y sea $u : GrF \rightarrow \mathbb{R}$ una función Borel medible tal que $u(x, \cdot)$ es una función semi-continua en $F(x)$ para cada $x \in X$. Entonces existe un selector Borel medible $f : X \rightarrow A$ para F tal que

$$u(x, f(x)) = \max_{a \in F(x)} u(x, a) \quad \text{para cada } x \in X.$$

Mas aún, la función v definida por $v(x) = \max_{a \in F(x)} u(x, a)$ es Borel medible. Similarmente, si $u(x, \cdot)$ es una función semi-continua inferiormente en $F(x)$ para cada $x \in X$, entonces existe un selector Borel medible $g : X \rightarrow A$ para F tal que

$$u(x, g(x)) = \min_{a \in F(x)} u(x, a) \quad \text{para cada } x \in X,$$

y, la función w definida por $w(x) = \min_{a \in F(x)} u(x, a)$ es Borel medible.

Demostración. Ver [5] ■

Teorema B.0.3. (Teorema de Ky Fan) Sean X, Y dos espacios compactos de Hausdorff y f una función valuada en los reales definida en $X \times Y$. Suponga que

- (a) $f(x, y)$ es semi-continua superiormente en x para cada $y \in Y$, y semi-continua inferiormente en y para cada $x \in X$,
- (b) f es cóncava en X , i.e., para cada $x_1, x_2 \in X$ y $0 \leq t \leq 1$, existe un elemento $x_0 \in X$ tal que

$$tf(x_1, y) + (1 - t)f(x_2, y) \leq f(x_0, y) \quad \text{para cada } y \in Y.$$

- (c) f es convexa en Y , i.e., para cada $y_1, y_2 \in Y$ y $0 \leq t \leq 1$, existe un elemento $y_0 \in Y$ tal que

$$tf(x, y_1) + (1 - t)f(x, y_2) \geq f(x, y_0) \quad \text{para cada } x \in X.$$

Entonces se cumple que

$$\max_{x \in X} \min_{y \in Y} f(x, y) = \min_{y \in Y} \max_{x \in X} f(x, y)$$

Demostración. (Ver [2]) Primero que nada notemos que al satisfacerse (a), independientemente si se cumplen o no (b) y (c), las expresiones

$$\max_{x \in X} \min_{y \in Y} f(x, y), \quad \text{y} \quad \max_{y \in Y} \min_{x \in X} f(x, y)$$

tienen sentido. De hecho, al ser $f(x, y)$ semi-continua superiormente para toda $x \in X$ en el espacio compacto Y , $\max_{y \in Y} f(x, y)$ existe. Luego, como el máximo de una familia de funciones semi-continuas inferiormente en X , $\max_{y \in Y} f(x, y)$ es una función semi-continua inferiormente en el espacio compacto X , y por lo tanto $\min_{x \in X} \max_{y \in Y} f(x, y)$ existe. Un argumento análogo nos sirve para asegurar que $\max_{y \in Y} \min_{x \in X} f(x, y)$ también existe.

Ahora, supongamos que f es convexa en X y cóncava en Y . Sean $\{x_1, x_2, \dots, x_n\} \subset X$ y $\{y_1, y_2, \dots, y_m\} \subset Y$ conjuntos arbitrarios. Por el teorema minimax de Von Neumann existen dos conjuntos $\{\xi_1, \xi_2, \dots, \xi_n\}$, $\{\eta_1, \eta_2, \dots, \eta_m\}$ de números no negativos tales que

$$\sum_{i=1}^n \xi_i = 1 \qquad \sum_{k=1}^m \eta_k = 1$$

y que

$$\max_{1 \leq k \leq m} \sum_{i=1}^n \xi_i f(x_i, y_k) \leq \min_{1 \leq i \leq n} \sum_{k=1}^m \eta_k f(x_i, y_k). \tag{B.B.0.1}$$

Como f es convexa en X y concava en Y , existen $x_0 \in X$, $y_0 \in Y$ tales que

$$f(x_0, y_k) \leq \sum_{i=1}^n \xi_i f(x_i, y_k), \qquad (1 \leq k \leq m) \tag{B.B.0.2}$$

y

$$f(x_i, y_0) \geq \sum_{k=1}^m \eta_k f(x_i, y_k), \qquad (1 \leq i \leq n). \tag{B.B.0.3}$$

De [B.B.0.1](#), [B.B.0.2](#) y [B.B.0.3](#) se cumple que

$$f(x_0, y_k) \leq f(x_i, y_0), \qquad (1 \leq i \leq n, 1 \leq k \leq m).$$

De acuerdo con esto, es cierto que

$$\min_{x \in X} \max_{1 \leq k \leq m} f(x, y_k) \leq \max_{y \in Y} \min_{1 \leq i \leq n} f(x_i, y).$$

Entonces, cualquier número real r satisface al menos una de las dos desigualdades siguientes

$$\min_{x \in X} \max_{1 \leq k \leq m} f(x, y_k) \leq r, \quad \max_{y \in Y} \min_{1 \leq i \leq n} f(x_i, y) \geq r.$$

Sean $L(y; r) = \{x \in X | f(x, y) \leq r\}$, $U(x; r) = \{y \in Y | f(x, y) \geq r\}$, que son subconjuntos cerrados de X y Y , respectivamente. De aquí, para cada real r y cualesquiera dos conjuntos finitos $\{x_1, x_2, \dots, x_n\} \subset X$ y $\{y_1, y_2, \dots, y_m\} \subset Y$, las intersecciones

$$\bigcap_{k=1}^m L(y_k; r), \quad \bigcap_{i=1}^n U(x_i; r)$$

no pueden ser vacías al mismo tiempo. Como X, Y son compactos, se sigue que para cada real r al menos una de las dos intersecciones

$$\bigcap_{y \in Y} L(y; r), \quad \bigcap_{x \in X} U(x; r)$$

es no vacía. Esto es, existe $x_0 \in X$ tal que $f(x_0, y) \leq r$ para toda $y \in Y$, o existe $y_0 \in Y$ tal que $f(x, y_0) \geq r$ para toda $x \in X$. En otras palabras, todo real r satisface al menos una de las dos desigualdades

$$\min_{x \in X} \max_{y \in Y} f(x, y) \leq r, \quad \max_{y \in Y} \min_{x \in X} f(x, y) \geq r$$

de manera que

$$\min_{x \in X} \max_{y \in Y} f(x, y) \leq \max_{y \in Y} \min_{x \in X} f(x, y). \tag{B.B.0.4}$$

Sin embargo, notemos que, para $\hat{x} \in X$, $\hat{y} \in Y$ arbitrarios,

$$f(\hat{x}, \hat{y}) \leq \max_{x \in X} f(x, \hat{y})$$

de donde

$$\min_{y \in Y} f(\hat{x}, y) \leq \min_{y \in Y} \max_{x \in X} f(x, y)$$

y por ende

$$\max_{y \in Y} \min_{x \in X} f(x, y) \leq \min_{x \in X} \max_{y \in Y} f(x, y). \tag{B.B.0.5}$$

Finalmente, juntando B.B.0.4 y B.B.0.5 se obtiene

$$\max_{y \in Y} \min_{x \in X} f(x, y) = \min_{x \in X} \max_{y \in Y} f(x, y)$$

■

Apéndice C

Auxiliares

Teorema C.0.1. (Teorema de Scheffé) *Suponga que $\{F, F_n, n \geq 1\}$ son distribuciones de probabilidad con densidades $\{f, f_n, n \geq 1\}$. Si $f_n(x) \rightarrow f(x)$ casi en todas partes (esto es, para toda x excepto en un conjunto de Lebesgue con medida 0), entonces*

$$\int |f_n(x) - f(x)| dx \rightarrow 0.$$

Demostración. Véase [12], página 253 ■

Teorema C.0.2. (Teorema de la Convergencia Monótona) *Sean X una variable aleatoria y $\{X_n\}$ una sucesión de variables aleatorias. Si*

$$0 \leq X_n \nearrow X,$$

entonces

$$E(X_n) \nearrow E(X),$$

o lo que es equivalente

$$E\left(\lim_{n \rightarrow \infty} X_n\right) \nearrow \lim_{n \rightarrow \infty} E(X_n)$$

Demostración. Véase [12], página 123 ■

Teorema C.0.3. (Teorema de Fubini) *Sea $P = P_1 \times P_2$ una medida producto. Si X es medible y es no-negativa o integrable con respecto a P , entonces*

$$\begin{aligned}\int_{\Omega_1 \times \Omega_2} X dP &= \int_{\Omega_1} \left[\int_{\Omega_2} X_{\omega_1}(\omega_2) P_2(d\omega_2) \right] P_1(d\omega_1) \\ &= \int_{\Omega_2} \left[\int_{\Omega_1} X_{\omega_2}(\omega_1) P_1(d\omega_1) \right] P_2(d\omega_2)\end{aligned}$$

Demostración. Véase [12], página 152 ■

Simbolos y abreviaciones:

- \mathbb{R} : Conjunto de los números reales.
- \mathbb{N} : Conjunto de los números naturales.
- \mathbb{N}_0 : $\mathbb{N} \cup \{0\}$.
- Variables aleatorias i.i.d: Variables aleatorias independientes e idénticamente distribuidas.

Bibliografía

- [1] R.B. Ash. *Real Analysis and Probability*. Academic Press, New York, 1972. 9, 56
- [2] K. Fan. Minimax theorems. *Proceedings of the National Academy of Sciences of the United States of America*, 39, 1953. 59
- [3] O. Hernández-Lerma, J. González-Hernandez, and R. López-Martinez. Constrained average cost markov control processes in borel spaces. Technical Report 262, CINVESTAV-IPN, Departamento de Matemáticas, 2003. 57
- [4] Carmen G. Higuera Chan. Estrategias adaptadas para juegos markovianos de suma cero. Master's thesis, Departamento de Matemáticas, Universidad de Sonora, 2011. 2
- [5] C.J. Himmelberg, T. Parthasarathy, and F.S. Van Vleck. Optimal plans for dynamic programming problems. *Mathematics of Operations Research*, 1, 1976. 57, 58
- [6] K. Hinderer. *Foundations of Non-stationary Dynamic Programming with Discrete Time Parameter*, volume 33 of *Lecture Notes in Operations Research and Mathematical Systems*. Springer, Berlin, 1970. 5, 9
- [7] A. Jáskiewicz and A. Nowak. Zero-sum ergodic stochastic games with feller transition probabilities. *SIAM Journal on Control and Optimization*, 45:773–789, 2006. 17
- [8] J. Adolfo Minjárez-Sosa. *Zero-Sum Discrete-Time Markov Games with Unknown Disturbance Distribution Discounted and Average Criteria*. Springer Briefs in Probability and Mathematical Statistics. Springer, 2020. 2
- [9] A.S. Nowak. Measurable selection theorems for minimax stochastic optimization problems. *SIAM Journal on Control and Optimization*, 23, 1985. 7
- [10] K.R. Parthasarathy. *Probability Measures on Metric Spaces*. Academic Press, New York, 1967. 5
- [11] Estephania Pivac Alcaraz. Zero-sum markov games on borel spaces with non-constant discount factors and unbounded payoff. Master's thesis, Departamento de Matemáticas, Universidad de Sonora, 2017. 2
- [12] Sidney I. Resnick. *A Probability Path*. Birkhäuser, 2005. 62, 63