



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Procesos de Decisión de Markov Descontados
Sensibles al Riesgo

Tesis presentada al

Posgrado en Matemáticas

como requisito para la obtención del grado de

MAESTRO EN CIENCIAS MATEMÁTICAS

por

Ezequiel Hernández García

Asesorado por

Dr. Hugo Adán Cruz Suárez y Dr. Rubén Blancas Rivera

Puebla Pue.
Septiembre 2025



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

Procesos de Decisión de Markov Descontados
Sensibles al Riesgo

Tesis presentada al

Posgrado en Matemáticas

como requisito para la obtención del grado de

MAESTRO EN CIENCIAS MATEMÁTICAS

por

Ezequiel Hernández García

Asesorado por

Dr. Hugo Adán Cruz Suárez y Dr. Rubén Blancas Rivera

Puebla Pue.
Septiembre 2025

Título: Procesos de Decisión de Markov Descontados
Sensibles al Riesgo

Estudiante: EZEQUIEL HERNÁNDEZ GARCÍA
COMITÉ

Dr. Francisco Solano Tajonar Sanabria
Presidente

Dr. Rei Israel Ortega Gutiérrez
Secretario

Dr. Carlos Camilo Garay
Vocal

Dr. Ruy Alberto López Ríos
Vocal

Dr. Hugo Adán Cruz Suárez
Asesor

Dr. Rubén Blancas Rivera
Asesor

Dedicatoria

Con cariño para los seres queridos que físicamente dejaron de estar con nosotros.

*"Significaste tanto para todos nosotros
Eras especial, no miento
Iluminaste los días más oscuros
Y los cielos más nublados.*

*Tu sonrisa alegraba nuestros corazones
Tu risa era como escuchar música
Daría absolutamente cualquier cosa
Por tenerte bien y aquí cerca.*

*No pasa ni un segundo
Que no estés en nuestras mentes
Tu amor nunca olvidaremos
Del dolor con el tiempo sanaremos."*

Del poema *Significaste tanto*, de C. Mitchell.

Agradecimientos

El presente trabajo no es solo el resultado de horas de estudio y escritura, sino también de todo el amor, el apoyo y la compañía que recibí a lo largo de este camino.

A mis padres, Rosario y Jorge, el trayecto fue largo, pero lo logramos juntos. Gracias infinitas por darme la vida y por estar ahí aún con todos los problemas que nos tocó enfrentar, gracias por apoyarme incluso sin entender del todo lo que pasaba. Gracias por su fe inquebrantable en mí, por su amor que no pide explicaciones, y por abrazarme en silencio cuando ya no tenía palabras. Esta tesis también es suya, porque sin ustedes yo no estaría aquí.

Siempre tendré presente y estaré agradecido con mi familia, no hay palabras que alcancen para expresar mis agradecimientos para todos, a mis hermanos Israel y Jorge, a mis hermanas Anabel y Jarumy, a mis sobrinas Karen, Valeria, Paola, Natalia, Julieta y Victoria, y a mi sobrino Vladimir, para todos ustedes, mis sobrinas y sobrino, gracias por siempre sacarme una sonrisa y hacer de la casa un lugar alegre en cada momento, nunca olviden que siempre que pueda estaré ahí para lo que necesiten; Agradezco también a mis cuñadas Kenya y Yuriko y a mis cuñados Haroldo y José Miguel, y a todos los que me falta mencionar. Gracias por ser mi hogar en cada paso, por su amor incondicional y por sostenerme cuando sentí que ya no podía. Gracias por sus abrazos silenciosos, por sus gestos cotidianos llenos de cariño y por recordarme siempre quién soy, aún me falta mucho por aprender.

Agradezco con el corazón a mis asesores, al Dr. Rubén Blancas Rivera y al Dr. Hugo Adán Cruz Suárez, por haber confiado en mí, por guiarme con paciencia, por creer en mí incluso cuando todo se puso feo y difícil, incluso cuando yo dudaba y asomaban en mí intenciones de abandonar el proyecto; les agradezco por motivarme y enseñarme que el conocimiento va de la mano con la humildad.

A todos los profesores que he conocido en la Facultad de Ciencias Físico

Matemáticas y que me han motivado a seguir adelante, incluyendo a los profesores que inesperadamente adelantaron su camino y dejaron de estar físicamente, nuestro más grande reconocimiento, todo su esfuerzo es invaluable. Igualmente agradezco a los profesores que conocí gracias al deporte, al profesor Carmelo y profesor Juan, de los entrenamientos de fútbol del COMDE, gracias por enseñarme que para tener una vida balanceada uno siempre debe de considerar el deporte.

Gracias a quienes me acompañaron en el alma, incluso desde la distancia o desde otros planos. José Miguel, la pequeña Ana Victoria, mi querido Gunter, pensé en ustedes más veces de las que podría contar. Y en cada momento de debilidad, sentí su fuerza y cariño, nunca los olvidaremos.

A Ximena, no tienes idea de lo agradecido que estoy contigo por la compañía, el cariño, el apoyo y las risas, sabes que tienes y siempre tendrás un lugar especial en mi vida y en mi corazón, te amo, amorcito.

Quiero agradecer a todos mis amigos. A los amigos que hice durante el tiempo que llevo en la BUAP, A Belén, Jorge, Zaira, Andrés, Lou, Yess, Alexis, Óscar, y muchos más. A Adriana, gracias por permitirme forjar una bonita amistad contigo a pesar de nuestras diferencias. También agradezco a las amistades que hice gracias al deporte, a Pedro, Sergio, Kennethn, Osvaldo, y todos los demás amigos con los que juego al fútbol en Tlaxcala, también a los amigos que conocí en los entrenamientos con el profe Carmelo y el profe Juan: a Victor, Charly, Filos, Yamileth, Lalo, Pili, Gio, Andrey, Luis, y todos los que me falta mencionar, gracias por los momentos divertidos, mis mejores deseos siempre.

A mis sinodales, Dr. Francisco Solano Tajonar Sanabria, Dr. Rei Israel Ortega Gutiérrez, Dr. Carlos Camilo Garay y Dr. Ruy Alberto López Ríos mi más sincero agradecimiento pues varios de ustedes me acompañaron desde la licenciatura hasta este importante paso de mi formación académica. Gracias por su tiempo y por la dedicación con la que revisaron mi trabajo, todo su apoyo contribuyó significativamente a enriquecer este proyecto.

Al Consejo de Ciencia y Tecnología (CONACyT) por el apoyo económico brindado para la culminación de mis estudios de maestría y elaboración de la presente tesis. También agradezco a la Benemérita Universidad Autónoma de Puebla (BUAP) y a la Vicerrectoría de Investigación y Estudios de Posgrado (VIEP) por los apoyos otorgados.

Índice general

Introducción	1
1 Preliminares	5
1.1 El Modelo de Control de Markov	5
1.1.1 Políticas	6
1.1.2 Construcción del Proceso Estocástico	8
1.1.3 Modelo de Ecuaciones en Diferencias	9
1.1.4 Sensibilidad al Riesgo	9
1.1.5 Criterio de Rendimiento Descontado: Horizonte Finito e Infinito	12
1.1.6 El Problema de Control Óptimo (PCO)	13
2 PDMs Descontados Sensibles al Riesgo: Horizonte Finito con Criterio de Costos	15
2.1 Resultados teóricos	15
2.1.1 Condiciones de Optimalidad en el caso de horizonte finito . . .	16
2.1.2 Caso Neutral al Riesgo	17
2.1.3 Caso Sensible al Riesgo	17
2.2 Ejemplo: Problema de Reemplazamiento Óptimo de n máquinas . . .	18
2.2.1 Planteamiento	18
2.2.2 Elementos de un Modelo Numérico	21
2.2.3 Resultados Numéricos Obtenidos	22
2.2.4 Conclusiones	23
3 PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas	25
3.1 Caso Neutral al Riesgo	25
3.1.1 Condiciones de Optimalidad	26
3.1.2 Existencia de una Política Óptima	28
3.1.3 Ejemplo: Problema de Consumo-Inversión	30

<i>ÍNDICE GENERAL</i>	XI
3.2 Caso Sensible al Riesgo	32
3.2.1 Condiciones de Optimalidad	33
3.2.2 Existencia de una Política Óptima	40
3.2.3 Ejemplo	48
Conclusiones	60
A Kérneles Estocásticos e Integración	63
B Esperanza Condicional	65
C Algunos Resultados de Convergencia	67
D Notación de Landau y propiedades de $o(\cdot)$	69
E Códigos elaborados en Python	71
E.1 Códigos para el Problema de Reemplazamiento Óptimo de n Máquinas	71
E.2 Código para la aproximación del producto P	76
Bibliografía	77

Introducción

La toma de decisiones en entornos inciertos ha sido un área central de estudio en diversas disciplinas, como la matemática aplicada, la inteligencia artificial, la economía y la investigación operativa. En este contexto, los Procesos de Decisión de Markov (PDM) han emergido como una herramienta fundamental para modelar sistemas estocásticos donde las decisiones influyen en la evolución futura del sistema.

Tradicionalmente, los PDM se han abordado bajo el enfoque neutral al riesgo ([20, 15] y [12]), por ejemplo, suponiendo que el objetivo del agente decisor es maximizar el valor esperado de la recompensa total descontada [10]. Sin embargo, en muchos contextos reales —como la gestión financiera, el control robótico, la planificación médica o la logística bajo incertidumbre— es fundamental incorporar la aversión o sensibilidad al riesgo en el proceso de decisión. Un ejemplo de modelos que sugieren incorporar sensibilidad al riesgo puede consultarse en [5]. Esta necesidad ha dado lugar al estudio de Procesos de Decisión de Markov Sensibles al Riesgo, donde el objetivo no es solamente maximizar la expectativa, sino también controlar la variabilidad o el impacto de eventos adversos. Un primer acercamiento a los PDM puede hacerse mediante el estudio de Cadenas de Markov con Costos asociados no controladas como se hace en [4] y [13], trabajos en los que se hace un análisis de la relación entre criterios de rendimiento asociados a dichas cadenas.

La presente tesis se relaciona con el estudio de los PDM sensibles al riesgo que evolucionan a tiempo discreto, ya sea hasta un horizonte N o bien con horizonte infinito, en un espacio de estados de Borel y con respecto a optimizar en torno al costo total descontado, o bien, en torno a la recompensa total descontada. Los principales objetivos de este trabajo son:

1. Analizar, comprender y reportar los avances que se tienen en torno al estudio de los PDM descontados neutrales y sensibles al riesgo, tanto en horizonte finito como en horizonte infinito, tomando como referencias principales [15, 12, 10, 8, 7] y [11].

2. Aportar ejemplos aplicados que resulten novedosos en el sentido de que contribuyan a la comprensión de la teoría sensible al riesgo, como en [14].
3. Con base en los puntos anteriores, proponer problemas de investigación futuros y posibles aportaciones en torno a la teoría existente.

Este trabajo se organiza en tres capítulos. El primer capítulo expone los preliminares necesarios para el entendimiento y formulación de los resultados de estudio. En particular, se define el concepto de Modelo de Control de Markov a tiempo discreto, se enuncia la construcción del Proceso de Decisión de Markov (PDM) como Proceso Estocástico, y se estudia la noción de Sensibilidad al Riesgo para después relacionarla con los PDM mediante los respectivos criterios de rendimiento (también conocidos como funciones objetivo). En esta tesis dichos criterios corresponden al costo o recompensa total descontados; a estos PDM con criterios sensibles al riesgo se les denomina Procesos de Decisión de Markov Descontados Sensibles al Riesgo. El capítulo concluye con la definición del Problema de Control Óptimo (PCO), que consiste en optimizar tales criterios.

El segundo capítulo presenta el estudio de los PDM Descontados Sensibles al Riesgo con horizonte finito, exponiendo algunos de los resultados principales abordados en [8] y [15], para posteriormente aplicar dichos resultados al Problema de Reemplazamiento Óptimo de n Máquinas. En la solución numérica de este problema inicia la aportación de la tesis, pues en la literatura actual no se cuenta con una solución para el caso sensible al riesgo con criterio descontado.

Finalmente, el tercer capítulo analiza la solución del PCO para los PDM Descontados Sensibles al Riesgo con horizonte infinito, dividiendo el capítulo en dos secciones: la primera dedicada al caso Neutral al Riesgo, ilustrando la solución y un ejemplo presentados en [10], y la segunda al caso Sensible al Riesgo, donde se presentan y analizan los avances más recientes, particularmente los de [11]. Es en esta parte donde se realiza otra aportación, al proponer un modelo basado en el ejemplo formulado para el caso neutral al riesgo, constituyendo un modelo novedoso que, entre otras cosas, permite ilustrar el procedimiento de iteración de valores para la solución del PCO de manera explícita.

Resumen de Notación y Terminología

Símbolos y abreviaciones

\mathbb{N}	Conjunto de enteros positivos
\mathbb{N}_0	Conjunto de enteros no negativos
\mathbb{R}	Conjunto de números reales
$:=$	Igualdad por definición
S	Espacio de Estados
$I_D(\cdot)$	Función indicadora del conjunto D $I_D(x) := 1$ si $x \in D$ y $I_D(x) := 0$ si $x \notin D$
\mathbb{K}	Espacio de pares estado-acción admisibles
Q	Ley de Transición
N	Horizonte del proceso
\mathbb{H}_t	Conjunto de historias del proceso hasta el tiempo t
$h_t \in \mathbb{H}_t$	Historia del proceso hasta el tiempo t
Π	Conjunto de Políticas
$\pi \in \Pi$	Política
Ω	Espacio muestral
\mathcal{F}	σ -álgebra
P_x^π	Medida de Probabilidad del proceso
E_x^π	Operador Esperanza del proceso
v.a.	Variable Aleatoria
v.a.i.i.d.	Variables Aleatorias independientes e idénticamente distribuidas.

Notación

Se utiliza la siguiente notación:

A	Conjunto de acciones
$a \in A$	Acción
$A(x)$	Conjunto de acciones admisibles en el estado $x \in S$
H	Función de respuesta
R	Función de recompensa
C	Función de costo
a_t	Acción en el tiempo t
Π	Conjunto de Políticas
$\pi \in \Pi$	Política
\mathbb{F}	Conjunto de Funciones medibles f tal que $f(x) \in A(x)$
Π_{RM}	Políticas de Markov Aleatorizadas
Π_{RS}	Políticas Aleatorizadas Estacionarias
Π_D	Políticas Deterministas
Π_{DM}	Políticas Deterministas de Markov
Π_{DS}	Políticas Estacionarias Deterministas

Respecto a los criterios de rendimiento se tiene:

Criterio

$\mathcal{J}_N(\lambda, \pi, x, \alpha)$ Criterio λ -sensible del costo descontado con horizonte finito.

$\mathcal{J}(\lambda, \pi, x, \alpha)$ Criterio λ -sensible de la recompensa descontada con horizonte infinito.

Para un espacio de Borel X , se utiliza la siguiente notación (véase [15]):

$\mathcal{B}(X)$	σ -álgebra de Borel en X
$P(X)$	Espacio de medidas de probabilidad en X
$P(X Y)$	Familia de kérneos estocásticos en X dado Y donde Y es un espacio de Borel.

Capítulo 1

Preliminares

Este capítulo establece los fundamentos teóricos necesarios para el desarrollo del presente trabajo. Según el perfil del lector, parte del contenido puede resultar familiar; sin embargo, su inclusión es esencial para garantizar una comprensión sólida de los temas que se abordarán en los capítulos posteriores. La información aquí presentada servirá como referencia y será empleada en las siguientes secciones, facilitando tanto el análisis como la aplicación de los conceptos clave que sustentan el enfoque de este estudio.

La teoría expuesta en este capítulo se basa principalmente en [12, 15, 20]. Para el caso sensible al riesgo, las referencias centrales son [8] y [11], aunque también pueden encontrarse otros enfoques en [7].

1.1. El Modelo de Control de Markov

Consideremos

$$\mathcal{M} = (S, A, \{A(x) : x \in S\}, Q, H)$$

un Modelo de Control de Markov (**MCM**) estacionario a tiempo discreto, donde:

- S es el espacio de estados, que se asume es un espacio de Borel.
- A es el espacio de acciones, también un espacio de Borel.
- $A(x)$ denota el conjunto de acciones admisibles cuando el sistema se encuentra en el estado $x \in S$.

- $\mathbb{K} := \{(x, a) : x \in S, a \in A(x)\}$ es el conjunto de pares ordenados estado–acción admisibles.
- Q es un *kérnel estocástico* definido en S dado \mathbb{K} , llamado *ley de transición*. Es decir, para cada $(x, a) \in \mathbb{K}$, $Q(\cdot|x, a)$ es una medida de probabilidad en S , y para cada $B \in \mathcal{B}(S)$, $Q(B|\cdot)$ es una función medible (donde $\mathcal{B}(S)$ denota la σ -álgebra de Borel de S).
- $H : \mathbb{K} \rightarrow \mathbb{R}$ es la función de respuesta, que puede representar una recompensa (denotada $H = R$) o una función de costo (denotada $H = C$). Esta función es medible en \mathbb{K} y representa la recompensa (o costo) asociada a tomar la acción a en el estado x .

Durante cada tiempo $t \in \mathbb{N}$, el sistema dinámico se encuentra en un estado actual, denotado como $X_t = x \in S$, y se elige una acción $A_t = a \in A(x)$. Posteriormente, se genera una recompensa (o costo) $H(x, a)$ y el sistema se traslada al siguiente estado $X_{t+1} = y \in S$ de acuerdo con la ley de transición $Q(\cdot|x, a)$.

Suposición 1.1.1. Se supondrá que \mathbb{K} contiene la gráfica de una función medible de S a A , es decir, existe $f : S \rightarrow A$ medible tal que $f(x) \in A(x)$ para toda $x \in S$ (ver [15]). El conjunto de todas estas funciones se denota por \mathbb{F} y sus elementos se llaman *selectores* de la multifunción $x \mapsto A(x)$.

1.1.1. Políticas

Definimos \mathbb{H}_t , el espacio de las historias observadas del proceso de control hasta el tiempo t , como

$$\begin{aligned}\mathbb{H}_0 &= S, \\ \mathbb{H}_t &= \mathbb{K} \times \mathbb{H}_{t-1},\end{aligned}$$

para $t = 1, 2, \dots$. Un elemento h_t de \mathbb{H}_t llamado t - historia es un vector de la forma

$$(x_0, a_0, x_1, a_1, \dots, x_{t-1}, a_{t-1}, x_t),$$

donde $(x_i, a_i) \in \mathbb{K}$, para $i = 0, 1, \dots, t - 1$ y $x_t \in X$.

Definición 1.1.2. Una política aleatorizada o simplemente política es una sucesión $\pi = \{\pi_t : t = 0, 1, 2, \dots\}$ de kérneles estocásticos definidos sobre A dado la historia del proceso \mathbb{H}_t y satisface que: $\pi_t(A(x_t)|h_t) = 1$ para toda $h_t \in \mathbb{H}_t$ y $t = 0, 1, 2, \dots$.

Notación 1.1.3. Al conjunto de todas las políticas lo denotamos por Π .

De acuerdo con la Definición 1.1.2, una política $\pi = \{\pi_t\}$ puede interpretarse como una sucesión $\{a_t\}$ de variables aleatorias sobre A , tales que, para cada t -historia y $t = 0, 1, 2, \dots$, la distribución de a_t es $\pi_t(\cdot|h_t)$, la cual está concentrada en el conjunto de acciones admisibles $A(x_t)$. En otras palabras, cuando usamos una política arbitraria, la acción en cualquier tiempo t es una variable aleatoria y depende de todas las t -historias.

Notación 1.1.4. Se denota a la familia de k ernes estoc sticos sobre A dado S , como $P(A|S)$. Y se denota por Φ al conjunto de todos los k ernes estoc sticos φ en $P(A|S)$ tales que para toda $x \in S$ se tiene que $\varphi(A(x)|x) = 1$.

Clasificaci n de pol ticas. Una pol tica $\pi \in \Pi$ es:

- **Markoviana Aleatorizada** (Π_{RM}). Si existe una sucesi n $\{\varphi_t\}$ de k ernes estoc sticos con $\varphi_t \in \Phi$ (definidas sobre A dado X), tales que, $\pi_t(\cdot|h_t) = \varphi_t(\cdot|x_t)$ para todo $h_t \in \mathbb{H}_t$ y $t = 0, 1, 2, \dots$
- **Markoviana Aleatorizada Estacionaria** (Π_{RS}). Si existe $\varphi \in \Phi$ k ernel estoc stico, tal que: $\pi_t(\cdot|h_t) = \varphi(\cdot|x_t)$ para toda $h_t \in \mathbb{H}_t$ y $t = 0, 1, 2, \dots$
- **Determinista** (Π_D). Si existe una sucesi n $\{g_t\}$ de funciones medibles con $g_t : \mathbb{H}_t \rightarrow A$, tales que, para cada $h_t \in \mathbb{H}_t$ y $t = 0, 1, 2, \dots$, se tiene que $g_t(h_t) \in A(x_t)$ y $\pi_t(\cdot|h_t)$ est  concentrada en $g_t(h_t)$.
- **Determinista Markoviana** (Π_{DM}). Si existe una sucesi n $\{f_t\}$ de funciones medibles con $f_t : X \rightarrow A$ (o $f_t \in \mathbb{F}$), tales que $f_t(x_t) \in A(x_t)$ y $\pi_t(\cdot|h_t)$ est  concentrada en $f_t(x_t)$ para cada $h_t \in \mathbb{H}_t$ y $t = 0, 1, 2, \dots$
- **Determinista Markoviana Estacionaria** (Π_{DS}). Si existe una funci n medible $f : X \rightarrow A$ (o $f \in \mathbb{F}$), tal que $f(x_t) \in A(x_t)$ y $\pi_t(\cdot|h_t)$ est  concentrada en $f(x_t)$ para cada $h_t \in \mathbb{H}_t$ y $t = 0, 1, 2, \dots$

Nota 1.1.5. a) Se dice que $\pi(\cdot|h)$ est  concentrada en $g(h)$, si, $\pi(C|h) = I_C(g(h))$ para cada $C \in \mathcal{B}(A)$. Donde I_C es la funci n indicadora del conjunto C .

b) Notar que

$$\Pi_{RS} \subset \Pi_{RM} \subset \Pi \quad \text{y} \quad \Pi_{DS} \subset \Pi_{DM} \subset \Pi_D \subset \Pi.$$

1.1.2. Construcción del Proceso Estocástico

Sea (Ω, \mathcal{F}) un espacio medible, donde el espacio muestral canónico se define como

$$\Omega := \overline{H}_\infty = (S \times A)^\infty,$$

y \mathcal{F} es la correspondiente σ -álgebra producto. Los elementos de Ω son de la forma

$$\omega = (x_0, a_0, x_1, a_1, \dots),$$

con $x_t \in S$ y $a_t \in A$ para todo $t = 0, 1, \dots$. Las proyecciones x_t y a_t de Ω sobre S y A se denominan, respectivamente, *estado* y *acción*.

Observemos que $H_\infty := \mathbb{K}^\infty \subset \Omega$ es el conjunto de parejas estado-acción admisibles. Sea $\pi \in \Pi$ una política arbitraria y $x_0 = x \in S$. Entonces, por el **Teorema de Ionescu–Tulcea** (también conocido como **Teorema de la medida producto**, ver Teorema A.0.2 en el Apéndice A), existe una única medida de probabilidad P_x^π sobre (Ω, \mathcal{F}) .

Además, para cada $C \in \mathcal{B}(A)$, $B \in \mathcal{B}(S)$, $h_t \in \mathbb{H}_t$ y $t = 0, 1, 2, \dots$, se cumple que

$$P_x^\pi(A_t \in C \mid h_t) = \pi_t(C \mid h_t), \quad (1.1)$$

$$P_x^\pi(X_{t+1} \in B \mid h_t, A_t) = Q(B \mid X_t, A_t). \quad (1.2)$$

Nota 1.1.6. De manera general, en lugar de dar $x_0 \in S$ puede darse una medida de probabilidad μ sobre S , referida como distribución inicial, de manera que

$$P_\mu^\pi(x_0 \in B) = \mu(B),$$

para cada $B \in \mathcal{B}(S)$.

El proceso estocástico $((\Omega, \mathcal{F}, P_x^\pi), \{X_t\})$ es llamado **Proceso de control de Markov a tiempo discreto** o **Proceso de Decisión de Markov (PDM)**.

La esperanza con respecto a P_x^π será denotada por E_x^π .

Nota 1.1.7. El PDM descrito aquí, es llamado **estacionario** pues sus componentes no dependen del parámetro tiempo t , en caso de que esto suceda se dice ser un modelo **no estacionario**, es decir, un modelo de la forma $(S_t, A_t, \{A_t(x) \mid x \in S_t\}, Q_t, R_t)$ para $t = 0, 1, 2, \dots$

La suposición 1.1.1 asegura que \mathbb{F} es no vacío y por lo tanto Π también. Esto se debe a que cada $f \in \mathbb{F}$ puede ser identificada por un kernel estocástico φ , de la siguiente manera: $\varphi(B \mid x) = I_B(f(x))$ para todo $B \in \mathcal{B}(A)$ y $x \in S$.

1.1.3. Modelo de Ecuaciones en Diferencias

En diversos contextos aplicados, la ley de transición Q es inducida por una ecuación en diferencias estocásticas de la forma

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

con $x_0 = x$ conocido. Aquí, $\{\xi_t\}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas (v.a. i.i.d.), que toman valores en un espacio X , con distribución común μ e independientes del estado inicial x_0 . Además, $F : \mathbb{K} \times X \rightarrow S$ es una función medible conocida.

En este caso, la ley de transición Q viene dada por:

$$\begin{aligned} Q(B \mid x, a) &= P(x_{t+1} \in B \mid x_t = x, a_t = a) \\ &= P(F(x_t, a_t, \xi_t) \in B \mid x_t = x, a_t = a) \\ &= P(F(x, a, \xi_t) \in B) \\ &= \int_S I_B(F(x, a, s)) \mu(ds) \\ &= E[I_B(F(x, a, \xi))], \end{aligned} \tag{1.3}$$

para todo $B \in \mathcal{B}(S)$ y $(x, a) \in \mathbb{K}$, donde I_B es la función indicadora del conjunto B y E denota la esperanza respecto a la variable aleatoria ξ .

En el caso determinista, es decir, cuando la dinámica está dada por

$$x_{t+1} = G(x_t, a_t), \quad G : \mathbb{K} \rightarrow S,$$

la ley de transición se reduce a

$$Q(B \mid x, a) = I_B(G(x, a)), \quad B \in \mathcal{B}(S).$$

1.1.4. Sensibilidad al Riesgo

Históricamente, los Procesos de Markov con sensibilidad al riesgo surgen a partir del enfoque de la *Teoría de la Utilidad* en Economía, la cual establece que, bajo ciertas condiciones, es posible modelar las preferencias de un consumidor mediante una función numérica denominada *función de utilidad*. Para definirla, consideremos lo siguiente.

Notación 1.1.8. Sea W un conjunto no vacío que representa las alternativas del consumidor, y sean $x, y \in W$. Usaremos la notación $x \succcurlyeq y$ para indicar que x representa una alternativa al menos tan buena para el consumidor como y .

Definición 1.1.9. Una función $u : W \rightarrow \mathbb{R}$ se denomina **función de utilidad** (o también **orden de preferencia**) definida sobre el conjunto de alternativas W si, para cada $x, y \in W$, se cumple que

$$x \succ y \Leftrightarrow u(x) \geq u(y).$$

En otras palabras, la función u representa las preferencias del controlador si y sólo si, al comparar dos alternativas, asigna un número real mayor a aquella que resulta preferida.

Sea ahora Δ el conjunto de distribuciones de probabilidad definidas sobre W . Un resultado fundamental debido a [18] establece que existe una función $U : \Delta \rightarrow \mathbb{R}$ que modela un orden de preferencias (\succ) en Δ , es decir,

$$\forall P, Q \in \Delta : P \succ Q \Leftrightarrow U(P) \geq U(Q).$$

Este resultado muestra que también pueden representarse preferencias entre distribuciones de probabilidad mediante funciones de utilidad, denominadas **funciones de utilidad esperada**.

Dependiendo de la forma de la función de utilidad, es posible interpretar la actitud del controlador frente al riesgo. En la literatura se introduce la medida de aversión al riesgo, llamada **coeficiente de aversión al riesgo**, definida por

$$\lambda(Y) = \frac{u''(Y)}{u'(Y)} = \frac{d}{dY} \ln(u'(Y)),$$

donde se supone que u es dos veces diferenciable y Y representa una recompensa o un costo. En general, este cociente puede depender de Y . Esta medida genera funciones de utilidad conocidas como de *Aversión Absoluta al Riesgo* (**ARA**, por sus siglas en inglés). Cuando el cociente es constante, se obtiene una medida de tipo **CARA** (*Constant Absolute Risk Aversion*).

En particular:

- Si $\lambda(\cdot) > 0$, el controlador es **averso** al riesgo.
 - Si $\lambda(\cdot) < 0$, el controlador es **propenso** al riesgo.
 - Si $\lambda(\cdot) = 0$, lo que implica que u es lineal, el controlador es **neutral** al riesgo.
-

De esta manera, disponemos de una forma de medir la sensibilidad al riesgo del controlador. Cuando λ no depende de la riqueza actual, sino que es constante, la función de utilidad del controlador debe adoptar la forma

$$u_\lambda(Y) := \begin{cases} e^{\lambda Y}, & \lambda \neq 0, \\ Y, & \lambda = 0, \end{cases} \quad (1.4)$$

conocida como **función de utilidad exponencial**.

Otra variante, utilizada por ejemplo en [8], está dada por

$$u_\lambda(Y) := \begin{cases} \frac{1}{\lambda} e^{\lambda Y}, & \lambda \neq 0, \\ Y, & \lambda = 0, \end{cases} \quad (1.5)$$

la cual también se emplea frecuentemente en aplicaciones.

Definición 1.1.10. Definimos la **certeza equivalente** (\mathcal{Q}) como la función

$$\mathcal{Q}(Y) = u^{-1}(E[u(Y)]).$$

En el caso de la expresión para $\lambda(Y)$ mencionada en (1.4), esto se traduce en:

$$\mathcal{Q}_\lambda(Y) := \begin{cases} \frac{1}{\lambda} \ln(E[e^{\lambda Y}]), & \lambda \neq 0, \\ E[Y], & \lambda = 0, \end{cases} \quad (1.6)$$

la cual es conocida como la **medida entrópica de riesgo**.

Nos referiremos a un controlador del sistema como:

- **Neutral al riesgo** si $\mathcal{Q}(Y) = E[Y]$ siempre se cumple.
- **Averso al riesgo** si $\mathcal{Q}(Y) > E[Y]$ para algún Y no constante.
- **Propenso al riesgo** si $\mathcal{Q}(Y) < E[Y]$ para algún Y no constante.

donde Y es una recompensa o un costo aleatorio.

Observación 1.1.11. De (1.4) y (1.6) se sigue que

$$u_\lambda(\mathcal{Q}_\lambda(Y)) = E[u_\lambda(Y)],$$

lo cual equivale a

$$\mathcal{Q}_\lambda(Y) = u_\lambda^{-1}(E[u_\lambda(Y)]),$$

donde Y nuevamente es una recompensa o un costo aleatorio. Esto muestra que la función de utilidad y la certeza equivalente están fuertemente relacionadas, pues implican que el controlador puede intercambiar la oportunidad de obtener la recompensa aleatoria (o pagar un costo aleatorio) Y por la respectiva **certeza equivalente** (o **equivalente seguro**) $\mathcal{Q}_\lambda(Y)$.

Además, puede demostrarse (ver [16]) que

$$\mathcal{Q}_\lambda(Y) \approx E[Y] - \frac{\lambda}{2} \text{Var}[Y].$$

Esta aproximación establece un vínculo directo con la varianza, y permite interpretar que $\lambda > 0$ corresponde al caso **averso al riesgo**, $\lambda < 0$ al caso **propenso al riesgo** y $\lambda = 0$ al caso **neutral al riesgo**.

1.1.5. Criterio de Rendimiento Descontado: Horizonte Finito e Infinito

Cada PDM estará asociado a una función real llamada **criterio de rendimiento**, la cual mide la calidad de cada política a través de la sucesión de costos generados. Los criterios de rendimiento que se utilizarán en este trabajo se basan en la literatura existente [8, 11, 12, 15, 20].

Consideremos un MCM estacionario y un conjunto de políticas Π . Supondremos que el controlador tiene un coeficiente de sensibilidad al riesgo constante $\lambda \in \mathbb{R}$ y utiliza la función de utilidad exponencial definida en (1.4).

Horizonte Finito

Para el caso de horizonte finito, se estudia el enfoque de costos ($H = C$), considerando el siguiente criterio de rendimiento:

Definición 1.1.12. Sean $x_0 = x \in S$, $\pi \in \Pi$ y $N \in \mathbb{N}$. Definimos el **equivalente seguro del costo total descontado hasta el tiempo (u horizonte) N** , dado por $\sum_{t=0}^{N-1} \alpha^t C(X_t, A_t)$, donde $\alpha \in (0, 1)$ es el factor de descuento, como

$$J_N(\lambda, \pi, x, \alpha) = \begin{cases} \frac{1}{\lambda} \ln \left(E_x^\pi \left[e^{\lambda \sum_{t=0}^{N-1} \alpha^t C(X_t, A_t)} \right] \right), & \lambda \neq 0, \\ E_x^\pi \left[\sum_{t=0}^{N-1} \alpha^t C(X_t, A_t) \right], & \lambda = 0. \end{cases} \quad (1.7)$$

Nota 1.1.13. En la literatura, el costo total descontado hasta el horizonte N suele incluir un término adicional, de la forma

$$\sum_{t=0}^{N-1} \alpha^t C(X_t, A_t) + C_N(X_N),$$

donde C_N es una **función de costo terminal**. Sin embargo, en este trabajo, por simplicidad, se considerará que dicha función es nula, es decir, $C_N(\cdot) \equiv 0$.

Horizonte Infinito

Para el caso de horizonte infinito, se estudia el enfoque de recompensas ($H = R$), considerando el siguiente criterio de rendimiento:

Definición 1.1.14. Sea $x_0 = x \in S$, $\pi \in \Pi$. Definimos el **equivalente seguro de la recompensa total descontada**, dada por $\sum_{t=0}^{\infty} \alpha^t R(X_t, A_t)$, donde $\alpha \in (0, 1)$ es el factor de descuento, como

$$J(\lambda, \pi, x, \alpha) = \begin{cases} \frac{1}{\lambda} \ln \left(E_x^\pi \left[e^{\lambda \sum_{t=0}^{\infty} \alpha^t R(X_t, A_t)} \right] \right), & \lambda \neq 0, \\ E_x^\pi \left[\sum_{t=0}^{\infty} \alpha^t R(X_t, A_t) \right], & \lambda = 0. \end{cases} \quad (1.8)$$

1.1.6. El Problema de Control Óptimo (PCO)

El **problema de control óptimo (PCO)** consiste en encontrar una política óptima, es decir, minimizar la función $\pi \mapsto J_N(\lambda, \pi, x, \alpha)$ sobre Π , o bien maximizar la función $\pi \mapsto J(\lambda, \pi, x, \alpha)$ sobre Π , para toda $x \in S$.

De esta manera, el **PCO** correspondiente a cada enfoque se define como sigue:

Definición 1.1.15. El PCO en torno a cada enfoque consiste en:

- hallar una política $\pi^* \in \Pi$ tal que

$$J_N(\lambda, \pi^*, x, \alpha) = \inf_{\pi \in \Pi} J_N(\lambda, \pi, x, \alpha) := J_N^*(\lambda, x, \alpha),$$

para el **enfoque de costos**;

- hallar una política $\pi^* \in \Pi$ tal que

$$J(\lambda, \pi^*, x, \alpha) = \sup_{\pi \in \Pi} J(\lambda, \pi, x, \alpha) := J^*(\lambda, x, \alpha),$$

para el **enfoque de recompensas**.

En cualquiera de los casos, a J_N^* (o bien a J^*) se le denomina **λ -función de valor óptimo**, y a la política π^* tal que $J_N(\lambda, \pi^*, x, \alpha) = J_N^*(\lambda, x, \alpha)$ (o bien $J(\lambda, \pi^*, x, \alpha) = J^*(\lambda, x, \alpha)$) se la conoce como **política óptima α -descontada**.

Nota 1.1.16. En la literatura se consideran también otros criterios de rendimiento basados en diferentes funciones de costo. Por ejemplo, el **equivalente seguro del costo total hasta el horizonte N** , dado por $\sum_{t=0}^{N-1} C(X_t, A_t)$, definido como

$$\mathcal{V}_N(\lambda, \pi, x) := \begin{cases} \frac{1}{\lambda} \ln \left(E_x^\pi \left[e^{\lambda \sum_{t=0}^{N-1} C(X_t, A_t)} \right] \right), & \lambda \neq 0, \\ E_x^\pi \left[\sum_{t=0}^{N-1} C(X_t, A_t) \right], & \lambda = 0. \end{cases} \quad (1.9)$$

Entre muchos otros criterios, como los estudiados en [4] y [14].

En este trabajo se incluyen tanto el enfoque de costos como el de recompensas con el fin de analizar una gama más amplia de problemas y explorar resultados más diversos en la literatura existente. Cabe destacar que, a pesar de los avances obtenidos en el estudio de los PDM sensibles al riesgo y en particular en los criterios descontados, aún existen variantes que no han sido exploradas. Por ejemplo, resolver el PCO en el caso del enfoque de costos no acotados con horizonte infinito. Estos problemas abiertos constituyen propuestas naturales para futuros estudios e investigaciones.

Capítulo 2

PDMs Descontados Sensibles al Riesgo: Horizonte Finito con Criterio de Costos

En este capítulo se presenta la teoría correspondiente a la solución del PCO en el caso neutral y sensible al riesgo, para horizonte finito considerando costos. Las referencias principales en torno a los resultados teóricos son [8] y [15].

Asimismo, se estudia el **Problema de Reemplazamiento Óptimo de n máquinas**. Un primer acercamiento a este problema en el caso neutral al riesgo se encuentra en [2], donde se analiza una única máquina que sigue un proceso estocástico de deterioro, al cual se le asocian costos por estado y acción en cada etapa. El caso de una sola máquina bajo sensibilidad al riesgo ha sido tratado en [9]. El modelo que se considerará en este trabajo está basado en [17].

Cabe destacar que el presente estudio se apoya también en [14], trabajo estrechamente relacionado en el sentido de que el criterio utilizado allí es el de **costo total** definido en (1.9).

2.1. Resultados teóricos

En esta sección se enuncian los resultados teóricos que se emplearán en el análisis de horizonte finito. En primer lugar, se presentan las hipótesis fundamentales que deben considerarse.

2.1.1. Condiciones de Optimalidad en el caso de horizonte finito

Debido a que las condiciones que garantizan la solución del PCO en ambos casos son condiciones estándar, se formulará un único conjunto de hipótesis válido tanto para el caso neutral como para el caso sensible al riesgo.

Hipótesis 2.1.1. Consideraremos lo siguiente:

- $A(x)$ es compacto para cada $x \in S$.
- La correspondencia $x \mapsto A(x)$ es superiormente semicontinua (u.s.c., por sus siglas en inglés).
- La función $(x, a) \mapsto C(x, a)$ es inferiormente semicontinua (l.s.c., por sus siglas en inglés).
- Q es débilmente continuo, es decir, para toda función $w : S \rightarrow \mathbb{R}$ acotada y continua,

$$(x, a) \mapsto \int_S w(y) Q(dy \mid x, a)$$

es continua.

Suposición 2.1.2. Además de las condiciones anteriores, supondremos que la función de costo C es acotada; es decir, existe $c > 0$ tal que $C : \mathbb{K} \rightarrow [-c, c]$. Esta condición será particularmente útil para aplicar los resultados de [8].

A continuación se presentan los resultados teóricos en torno a la solución del PCO, tanto para el caso neutral como para el caso sensible al riesgo, en un horizonte finito $N \in \mathbb{N}$. Siguiendo el planteamiento de [8], si consideramos $\lambda > 0$ (aversión al riesgo), entonces, debido a la monotonía de la función $\ln(\cdot)$, puede estudiarse la siguiente modificación de (1.7).

Observación 2.1.3. Sean $x_0 = x \in S$, $\pi \in \Pi$ y $N \in \mathbb{N}$. El criterio de rendimiento basado en el costo total descontado hasta el tiempo (u horizonte) N , dado por $\sum_{t=0}^{N-1} \alpha^t C(X_t, A_t)$ con $\alpha \in (0, 1)$ factor de descuento, se define como

$$\mathcal{J}_N(\lambda, \pi, x, \alpha) = \begin{cases} E_x^\pi \left[\frac{1}{\lambda} e^{\lambda \sum_{t=0}^{N-1} \alpha^t C(X_t, A_t)} \right], & \lambda > 0, \\ E_x^\pi \left[\sum_{t=0}^{N-1} \alpha^t C(X_t, A_t) \right], & \lambda = 0. \end{cases} \quad (2.1)$$

2.1.2. Caso Neutral al Riesgo

Para el caso neutral al riesgo ($\lambda = 0$ en (2.1)), la técnica empleada es la de *Programación Dinámica* y el resultado teórico que la respalda es el siguiente (ver [15]):

Teorema 2.1.4. Sean u_0, u_1, \dots, u_N funciones de S a \mathbb{R} , definidas por $u_0(x) := 0$, y para cada $t = 1, 2, \dots, N$,

$$u_t(x) := \min_{a \in A(x)} \left\{ C(x, a) + \alpha \int_S u_{t-1}(y) Q(dy | x, a) \right\}. \quad (2.2)$$

Supóngase que dichas funciones son medibles y que para cada $t = 1, 2, \dots, N$ existe un selector $f_t \in \mathbb{F}$ con $f_t(x) \in A(x)$ tal que

$$u_t(x) = C(x, f_t(x)) + \int_S u_{t-1}(y) Q(dy | x, f_t(x)).$$

Entonces, la política determinista de Markov $\pi^* \in \Pi_{DM}$ dada por $\pi^* = \{f_N, f_{N-1}, \dots, f_1\}$ es óptima, y la función de valor óptimo \mathcal{J}_N es u_N . Es decir, para cada $x \in S$ se cumple que

$$\mathcal{J}_N^*(0, x, \alpha) = \mathcal{J}_N(0, \pi^*, x, \alpha) = u_N(x).$$

Al procedimiento de determinar las funciones u_0, u_1, \dots, u_N se le denomina **proceso de iteración de valores**. La relación (2.2) se conoce como la **Ecuación de Programación Dinámica (EPD) de Bellman**, o simplemente **Ecuación de Bellman**, junto con su condición inicial u_0 . Su nombre se debe a **Richard Bellman**.

2.1.3. Caso Sensible al Riesgo

Para el caso sensible al riesgo ($\lambda > 0$ en (2.1)), se tiene el siguiente resultado (ver [8]):

Teorema 2.1.5. Sean U_0, U_1, \dots, U_N funciones de S a \mathbb{R} , definidas por $U_0(x) := \frac{1}{\lambda}$, y para cada $t = 1, 2, \dots, N$,

$$U_t(x) := \min_{a \in A(x)} \left\{ e^{\alpha^{N-t} \lambda C(x, a)} \int_S U_{t-1}(y) Q(dy | x, a) \right\}. \quad (2.3)$$

2.2 Ejemplo: Problema de Reemplazamiento Óptimo de n máquinas

Dichas funciones son medibles, inferiormente semicontinuas y, por tanto, para cada $t = 1, 2, \dots, N$ existe un selector $g_t \in \mathbb{F}$ con $g_t(x) \in A(x)$ tal que

$$U_t(x) = e^{\alpha^{N-t}\lambda C(x, g_t(x))} \int_S U_{t-1}(y) Q(dy | x, g_t(x)).$$

Entonces, la política determinista de Markov $\pi^* \in \Pi_{DM}$ dada por $\pi^* = \{g_N, g_{N-1}, \dots, g_1\}$ es óptima y

$$\mathcal{J}_N^*(\lambda, x, \alpha) = \mathcal{J}_N(\lambda, \pi^*, x, \alpha) = U_N(x).$$

En este caso, el proceso iterativo definido en (2.3) con condición inicial U_0 es de carácter **multiplicativo**. A dicho proceso se le denomina **Ecuación Multiplicativa de Bellman**.

De esta manera, ya se han establecido los preliminares teóricos necesarios para el ejemplo en esta sección de horizonte finito. Se presenta ahora el modelo de n máquinas a estudiar. .

2.2. Ejemplo: Problema de Reemplazamiento Óptimo de n máquinas

En esta sección se planteará el modelo de **Reemplazamiento Óptimo de n Máquinas**, que será el objeto de estudio para la parte correspondiente a horizonte finito en PDMs descontados sensibles al riesgo. El planteamiento está basado en [17], mientras que el ejemplo numérico a trabajar toma como referencia [14].

2.2.1. Planteamiento

Se considera un sistema compuesto por n máquinas similares que operan de manera independiente. Los posibles niveles de deterioro de cada máquina se denotan por $1, 2, \dots, D$, donde $D \in \mathbb{N}$. El nivel 1 indica que la máquina está en condiciones óptimas, y en general se asume que el nivel i es mejor que el nivel $i + 1$, para $i = 1, 2, \dots, D - 1$.

Al inicio de cada período de observación, el estado del sistema puede representarse mediante el vector (d_1, d_2, \dots, d_n) , donde d_k , $k = 1, 2, \dots, n$, indica el nivel de deterioro de la k -ésima máquina. De este modo, el espacio de estados queda determinado por

$$S = \{(d_1, d_2, \dots, d_n) \mid d_k \in \{1, 2, \dots, D\}, k = 1, 2, \dots, n\}.$$

PDMs Descontados Sensibles al Riesgo: Horizonte Finito con Criterio de Costos

2.2 Ejemplo: Problema de Reemplazamiento Óptimo de n máquinas

En cuanto al espacio de acciones, se supone que al inicio de cada etapa pueden tomarse las siguientes decisiones:

- a) Permitir que la k -ésima máquina ($k = 1, 2, \dots, n$) continúe operando en su nivel de deterioro actual, o bien,
- b) Reemplazarla por una unidad nueva con un costo fijo $C_R > 0$.

Para formalizar el espacio de acciones, se considera el vector (a_1, a_2, \dots, a_n) con $a_k \in \{0, 1\}$, donde $a_k = 0$ significa que la k -ésima máquina continúa en el nivel d_k , y $a_k = 1$ significa que dicha máquina es reemplazada. Por lo tanto,

$$A = A(x) = \{(a_1, a_2, \dots, a_n) \mid a_k \in \{0, 1\}, k = 1, 2, \dots, n\}.$$

Para definir los costos, sea $g : \{1, 2, \dots, D\} \rightarrow \mathbb{R}$ una función conocida, donde $g(i)$, $i = 1, 2, \dots, D$, representa el costo de operación de una máquina en el i -ésimo nivel de deterioro. Se asume que g es no decreciente, es decir, $g(1) \leq g(2) \leq \dots \leq g(D)$.

Así, el costo en cualquier período de tiempo t se define como

$$C(x_t, a_t) := \sum_{k=1}^n \gamma(x_{k,t}, a_{k,t}), \quad (2.4)$$

donde

$$\gamma(x_{k,t}, a_{k,t}) := \begin{cases} g(x_{k,t}), & \text{si } a_{k,t} = 0, \\ g(1) + C_R, & \text{si } a_{k,t} = 1, \end{cases}$$

con $x_{k,t}$ representando el estado de la máquina k en el instante t y $a_{k,t}$ la acción aplicada sobre dicha máquina en ese mismo momento.

Obsérvese que, para decidir si una máquina debe ser reemplazada, únicamente se analiza el estado de dicha máquina, sin importar el estado de las demás. Luego, para cualquier máquina k se define la matriz de transición del proceso de deterioro cuando la máquina no es reemplazada:

$$P^0 := (p_{i,j}^0)_{D \times D},$$

donde $p_{i,j}^0$ representa la probabilidad de que una máquina pase del nivel i al nivel j . Como una máquina no puede pasar a un mejor nivel de deterioro, se cumple que $p_{i,j}^0 = 0$ si $j < i$.

PDMs Descontados Sensibles al Riesgo: Horizonte Finito con Criterio de Costos

2.2 Ejemplo: Problema de Reemplazamiento Óptimo de n máquinas

De manera análoga, se define la matriz de transición cuando la máquina es reemplazada por

$$P^1 := (p_{i,j}^1)_{D \times D},$$

donde $p_{i,j}^1 = 1$ si $j = 1$ y $p_{i,j}^1 = 0$ en otro caso. Esto significa que, con certeza, se regresa al nivel 1 de deterioro cuando la k -ésima máquina es reemplazada.

Una vez establecido lo anterior, se define

$$Q^a := (q_{i,j}^a)_{D^n \times D^n}$$

como la matriz de transición del sistema que describe el paso del estado i al estado j cuando se toma la acción $a \in A$, con $i, j \in S$, $a = (a_1, a_2, \dots, a_n)$, $i = (i_1, i_2, \dots, i_n)$ y $j = (j_1, j_2, \dots, j_n)$. En este caso,

$$q_{i,j}^a := \prod_{k=1}^n p_{i_k, j_k}^{a_k}.$$

Finalmente, en el caso sensible al riesgo se considera $\lambda > 0$, lo cual corresponde a aversión al riesgo. El Problema de Reemplazamiento Óptimo consiste entonces en minimizar la función objetivo dada por $\mathcal{J}_N(\lambda, \pi, x, \alpha)$ en (2.1), tomando como función de costo C la definida en (2.4).

Dado que este modelo satisface las hipótesis de continuidad–compacidad (ver [17]), es posible aplicar los resultados teóricos previos para determinar soluciones numéricas en ejemplos concretos, en particular el expuesto en [14].

2.2.2. Elementos de un Modelo Numérico

El ejemplo numérico trabajado considera los siguientes parámetros:

- $n = 2$ máquinas,
- $D = 3$ niveles de deterioro,
- matriz del proceso de deterioro P^0 dada por

$$P^0 = \begin{pmatrix} 0,4 & 0,3 & 0,3 \\ 0 & 0,3 & 0,7 \\ 0 & 0 & 1 \end{pmatrix},$$

- función de costo de operación de cada máquina $g(i) = \frac{3}{2}i^2 - \frac{7}{2}i + 4$, $i = 1, 2, 3$,
- costo de reemplazo $C_R = 1$,
- factor de descuento $\alpha = 0,3$,
- dos coeficientes de sensibilidad al riesgo: $\lambda_1 = 0,1$ y $\lambda_2 = 10$, y
- horizonte $N = 5$.

Los valores de λ fueron seleccionados con el propósito de comparar de manera clara la variación en la función de valor óptimo y en las políticas óptimas entre el caso neutral y el sensible al riesgo.

En la Tabla 2.1 se presentan los resultados obtenidos numéricamente mediante la implementación realizada en Python (ver Apéndice E).

PDMs Descontados Sensibles al Riesgo: Horizonte Finito con Criterio de Costos

2.2 Ejemplo: Problema de Reemplazamiento Óptimo de n máquinas

2.2.3. Resultados Numéricos Obtenidos

$\lambda = 0$									
Estado Etapa	(1, 1)	(1, 2)	(1, 3)	(2, 1)	(2, 2)	(2, 3)	(3, 1)	(3, 2)	(3, 3)
1	(0, 0)	(0, 0)	(0, 1)	(0, 0)	(0, 0)	(0, 1)	(1, 0)	(1, 0)	(1, 1)
2	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
3	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
4	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
5	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
Valor de \mathcal{J}_5^*	6.13	6.98	6.98	6.98	7.82	7.82	6.98	7.82	7.82
$\lambda = 0.1$									
1	(0, 0)	(0, 0)	(0, 1)	(0, 0)	(0, 0)	(0, 1)	(1, 0)	(1, 0)	(1, 1)
2	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
3	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
4	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
5	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
Valor de \mathcal{J}_5^*	18.46	20.10	20.10	20.10	21.87	21.87	20.10	21.87	21.87
$\lambda = 10$									
1	(0, 0)	(0, 0)	(0, 1)	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 0)	(1, 1)
2	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
3	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
4	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
5	(0, 0)	(0, 1)	(0, 1)	(1, 0)	(1, 1)	(1, 1)	(1, 0)	(1, 1)	(1, 1)
Valor de \mathcal{J}_5^*	13.45 $\times 10^{25}$	39.23 $\times 10^{28}$	39.23 $\times 10^{28}$	39.23 $\times 10^{28}$	14.46 $\times 10^{32}$	14.46 $\times 10^{32}$	39.23 $\times 10^{28}$	14.46 $\times 10^{32}$	14.46 $\times 10^{32}$

Tabla 2.1: Resultados obtenidos numéricamente por estado inicial, etapa y valor de \mathcal{J}_5^* para distintos valores de λ .

2.2.4. Conclusiones

Puede observarse en la Tabla 2.1 que, además de que el valor de la función de valor óptimo \mathcal{J}_5^* difiere entre el caso neutral y el caso sensible al riesgo, también lo hacen las políticas óptimas.

En particular, en la etapa inicial para el caso neutral al riesgo, el estado $(2, 2)$ — que representa a las dos máquinas en el segundo nivel de deterioro — conduce a la decisión de continuar operando ambas máquinas. En contraste, en el caso sensible al riesgo con $\lambda = 10$, la política óptima consiste en reemplazar únicamente la segunda máquina.

De manera general, de la Tabla 2.1 se desprende que, bajo una actitud aversa al riesgo con $\lambda = 10$, las decisiones óptimas son:

- No reemplazar máquinas en el primer nivel de deterioro, independientemente de la etapa.
- Reemplazar las máquinas si alcanzan el segundo nivel de deterioro en cualquier etapa.
- En la etapa final, reemplazar a partir del segundo nivel de deterioro.

Estos resultados muestran que un enfoque sensible al riesgo puede conducir a decisiones distintas frente al enfoque neutral al riesgo. Sin embargo, es importante destacar que dichas decisiones dependen tanto del modelo específico considerado como de los parámetros seleccionados.

Capítulo 3

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

En este capítulo se abordará el PCO con horizonte infinito tanto en el caso neutral como en el caso sensible al riesgo. La referencia principal para el caso neutral al riesgo es [12], mientras que para el caso sensible al riesgo es [11]. Además, se presenta un ejemplo aplicado en el caso neutral al riesgo estudiado en [10], mientras que en el caso sensible al riesgo se propone un modelo cuyas componentes permiten estudiar la iteración de valores y exhibir tanto políticas óptimas como la función de valor óptimo.

3.1. Caso Neutral al Riesgo

Sea $\mathcal{M} = (S, A, A(x), Q, R)$ un MCM estacionario, y consideremos $\lambda = 0$ en (1.8) junto con su correspondiente función de valor óptimo J^* . En esta sección se presentan las condiciones para la existencia de una solución al PCO en el caso neutral al riesgo, y de igual manera se establecen los resultados que permiten obtener soluciones de dicho problema.

Las demostraciones de los resultados enunciados en esta sección se encuentran en [12] y [10].

3.1.1. Condiciones de Optimalidad

Se presentan a continuación las condiciones necesarias para la existencia de una política óptima estacionaria.

Definición 3.1.1. La ecuación de valor óptimo α -descontada se define mediante la siguiente expresión funcional:

$$J^*(0, x, \alpha) = \max_{a \in A(x)} \left\{ R(x, a) + \alpha \int_S J^*(0, y, \alpha) Q(dy | x, a) \right\}, \quad x \in S. \quad (3.1)$$

Para el estudio del criterio descontado (ver (1.8)) se consideran tres conjuntos de hipótesis. La primera, Hipótesis 3.1.2, corresponde a la condición de continuidad–compacidad para modelos de control de Markov. La segunda, Hipótesis 3.1.7, introduce una función de peso w que regula el crecimiento de la función de costo. Finalmente, la Hipótesis 3.1.9 es una condición de continuidad adicional.

Hipótesis 3.1.2. (Condición de continuidad–compacidad) Para cada estado $x \in S$:

- (a) $A(x)$ es compacto.
- (b) $R(x, \cdot)$ es superiormente semicontinua (u.s.c.) en $A(x)$.
- (c) La función $\mu'(x, a) := \int_S \mu(y) Q(dy | x, a)$ es continua en $A(x)$ para cada función $\mu \in \mathbb{B}(S)$, donde $\mathbb{B}(S)$ denota el espacio de Banach de funciones medibles, continuas y acotadas en S , con la norma del supremo

$$\|\mu\|_\infty := \sup_{x \in S} |\mu(x)|.$$

Teorema 3.1.3. La Hipótesis 3.1.2 (c) se satisface si y sólo si $\int_S v(y) Q(dy | x, a)$ es u.s.c. en $A(x)$ para cada función no negativa $v \in \mathbb{B}(S)$.

Para establecer los resultados principales de esta sección introducimos la siguiente notación.

Notación 3.1.4. Sea $w : S \rightarrow [1, \infty)$ una función medible, denominada **función de peso** (o función de ponderación).

Definición 3.1.5. Si u es una función real definida en S , su w -norma se define como

$$\|u\|_w := \left\| \frac{u}{w} \right\|_\infty = \sup_{x \in S} \frac{|u(x)|}{w(x)}. \quad (3.2)$$

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.1 Caso Neutral al Riesgo

De acuerdo con esta definición, se dice que u es **acotada** si $\|u\| < \infty$ y w -acotada si $\|u\|_w < \infty$.

Observación 3.1.6. Sea $\mathbb{B}_w(S)$ el espacio lineal normado de funciones medibles w -acotadas en S . Este espacio es **completo**, por lo que $\mathbb{B}_w(S)$ es un espacio de Banach que contiene a $\mathbb{B}(S)$.

Hipótesis 3.1.7. (Condición de crecimiento) Existen constantes no negativas M y β , con $1 \leq \beta \leq \frac{1}{\alpha}$, y una función de peso $w \geq 1$ sobre S tal que, para cada estado $x \in S$:

- (a) $\sup_{a \in A(x)} |R(x, a)| \leq Mw(x)$,
- (b) $\sup_{a \in A(x)} \int_S w(y)Q(dy | x, a) \leq \beta w(x)$.

Observación 3.1.8. Si la función de recompensa R es acotada, entonces la Hipótesis 3.1.7 (a) se satisface de manera inmediata.

Hipótesis 3.1.9. (Condición de continuidad) Para cada estado $x \in S$, la función

$$w'(x, a) := \int_S w(y)Q(dy | x, a)$$

es continua en $A(x)$.

Proposición 3.1.10. Si la Hipótesis 3.1.7 se cumple, entonces

$$r(x) := \sum_{t=0}^{\infty} \alpha^t r_t(x) < \infty, \tag{3.3}$$

donde $r_0(x) := \sup_{a \in A(x)} |C(x, a)|$ y

$$r_t(x) := \sup_{a \in A(x)} \left\{ \int_S r_{t-1}(y)Q(dy | x, a) \right\}, \quad t \geq 1. \tag{3.4}$$

Recíprocamente, si

- (a) $r(x) \geq 1$,
- (b) se satisface la desigualdad en (3.3) para algún α_0 con $0 < \alpha_0 < \alpha$,

entonces la Hipótesis 3.1.7 se cumple con

$$M := 1, \quad w(x) := r(x), \quad \beta := \alpha_0^{-1}.$$

3.1.2. Existencia de una Política Óptima

Ahora se presenta el resultado principal de este capítulo. Primero consideremos lo siguiente.

Definición 3.1.11. Se define el n -ésimo costo descontado como

$$J_n(0, \pi, x, \alpha) := E_x^\pi \left[\sum_{t=0}^{n-1} \alpha^t R(X_t, A_t) \right].$$

En base a la Definición 3.1.11, puede observarse que, del Teorema de la Convergencia Monótona (ver Apéndice C),

$$J(0, \pi, x, \alpha) = \lim_{n \rightarrow \infty} J_n(0, \pi, x, \alpha).$$

Definición 3.1.12. El **algoritmo de iteración de valores óptimos α -descontado** se define como

$$j_n(x) = \max_{a \in A(x)} \left\{ R(x, a) + \alpha \int_S j_{n-1}(y) Q(dy | x, a) \right\}, \quad (3.5)$$

para cada $n \geq 1$ y $x \in S$, con $j_0(\cdot) \equiv 0$.

A las funciones $\{j_n\}$ se les conoce como **funciones de iteración de valores (IV)**.

Definición 3.1.13. Una función medible $v : S \rightarrow \mathbb{R}$ es una **solución de la Ecuación Óptima de Costo Descontado (α -EOCD)** si cumple que

$$v(x) = \max_{a \in A(x)} \left\{ R(x, a) + \alpha \int_S v(y) Q(dy | x, a) \right\}, \quad \forall x \in S.$$

Definición 3.1.14. Se define el operador $T_\alpha : \mathbb{B}_w(S) \rightarrow \mathbb{B}_w(S)$ de programación dinámica como

$$T_\alpha v(x) := \max_{a \in A(x)} \left\{ R(x, a) + \alpha \int_S v(y) Q(dy | x, a) \right\}.$$

Un resultado clave es el siguiente.

Teorema 3.1.15. Sea $0 < \alpha < 1$. Supóngase que las Hipótesis 3.1.2, 3.1.7 y 3.1.9 se cumplen, y sea T_α el operador definido en 3.1.14. Entonces:

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.1 Caso Neutral al Riesgo

- (a) T_α es un operador de contracción en $\mathbb{B}_w(S)$, con módulo $\gamma := \alpha\beta$. En particular, T_α mapea $\mathbb{B}_w(S)$ en sí mismo y

$$\|T_\alpha u - T_\alpha u'\|_w \leq \gamma \|u - u'\|_w \quad \forall u, u' \in \mathbb{B}_w(S).$$

- (b) Para cada función $u \in \mathbb{B}_w(S)$ existe un selector $f \equiv f_u \in \mathbb{F}$ tal que

$$T_\alpha u(x) = R(x, f(x)) + \alpha \int_S u(y) Q(dy | x, f(x)), \quad \forall x \in S.$$

La parte (b) se cumple para cualquier $0 \leq \alpha \leq 1$.

Con esto, ya estamos en posición de enunciar el resultado principal de esta sección.

Teorema 3.1.16. Supóngase que las Hipótesis 3.1.2, 3.1.7 y 3.1.9 se cumplen. Sea β la constante de la Hipótesis 3.1.7, y definamos $\gamma := \alpha\beta$. Entonces:

- (a) La función de valor α -descontada J^* es la única solución de la α -EOCD 3.1.13 en el espacio $\mathbb{B}_w(S)$. Además,

$$\|j_n - J^*\|_w \leq \frac{M\gamma^n}{1 - \gamma}, \quad n = 1, 2, \dots,$$

donde M es la constante de la Hipótesis 3.1.7 (a).

- (b) Existe un selector $f_* \in \mathbb{F}$ tal que $f_*(x) \in A(x)$ alcanza el máximo en 3.1.13 para cada estado x . Es decir,

$$J^*(0, x, \alpha) = R(x, f_*(x)) + \alpha \int_S J^*(0, y, \alpha) Q(dy | x, f_*(x)), \quad x \in S, \quad (3.6)$$

y la política determinista estacionaria $f_*^\infty \in \Pi_{DS}$ es óptima α -descontada. Recíprocamente, si $f_*^\infty \in \Pi_{DS}$ es óptima α -descontada, entonces satisface (3.6).

- (c) Una política π^* es óptima α -descontada si y sólo si la función correspondiente $J(0, \pi^*, \cdot, \alpha)$ satisface la α -EOCD.
- (d) Si existe una política óptima α -descontada, entonces existe una política determinista estacionaria que también es óptima α -descontada.

En resumen, los resultados expuestos muestran la importancia de las suposiciones a considerar en el modelo. Muchas de estas suposiciones son consideradas estándar en esta teoría y son ampliamente utilizadas. Existen además otros enfoques para la solución del PCO en el caso neutral al riesgo, algunos de los cuales se presentan en [15].

3.1.3. Ejemplo: Problema de Consumo-Inversión

En ésta sección se presenta un ejemplo conocido como el **Problema de consumo-inversión** que se suele estudiar en el enfoque neutral al riesgo, para ver los detalles sobre los resultados que se enunciarán, se puede consultar [10].

Planteamiento

Un inversor desea distribuir su riqueza actual x_t entre su inversión a_t y su consumo $x_t - a_t$, para cada tiempo $t = 0, 1, \dots$. Se asume que no se permite pedir prestado ni endeudarse. La dinámica del modelo está dada por

$$x_{t+1} = a_t \xi_t,$$

de manera que la riqueza al tiempo $t+1$ es proporcional al monto invertido al tiempo t , donde se supone que el ruido ξ_t consiste de variables aleatorias independientes e idénticamente distribuidas, independientes del capital inicial x_0 .

Se considerará que

- $S = A = [0, \infty)$,
- $A(x) = [0, x]$, para cada $x \in S$.

Aplicando (1.3) se obtiene

$$Q(B \mid x, a) = \begin{cases} \int_B \frac{1}{a} f\left(\frac{z}{a}\right) dz, & \text{si } a \neq 0, \\ I_B(0), & \text{si } a = 0, \end{cases}$$

donde $(x, a) \in \mathbb{K}$, $B \in \mathcal{B}(S)$ y f es la densidad común de ξ_t (ver Suposición 3.1.17).

La función de recompensa $R(x, a)$ está dada por

$$R(x, a) := \begin{cases} \frac{b}{\gamma} (x - a)^\gamma, & \text{si } x \neq a, \\ 0, & \text{si } x = a. \end{cases}$$

Suposición 3.1.17. Supongamos que:

- a) $\{\xi_t\}$ son variables aleatorias no negativas con densidad común f ,

b) $m := E[\xi_0]$ y $m_\gamma := E[\xi_0^\gamma]$ son finitos, $m > 1$, y

$$0 < \alpha m_\gamma < 1,$$

donde α es el factor de descuento y γ la constante en la función de recompensa.

Proposición 3.1.18. Si la Suposición 3.1.17 se satisface, entonces también se cumple la Hipótesis 3.1.2. Más aún, se cumplen las Hipótesis 3.1.7 y 3.1.9.

Una de las razones por las que este modelo resulta interesante proviene del hecho de que la función de peso $w(x)$ puede obtenerse explícitamente mediante la Proposición 3.1.10. Dicha función está dada por

$$w(x) = \frac{b}{\gamma} x^\gamma \sum_{t=0}^{\infty} (\alpha m_\gamma)^t,$$

donde la serie $\sum_{t=0}^{\infty} (\alpha m_\gamma)^t$ converge gracias a la Suposición 3.1.17.

Solución

Se establecen ahora los resultados principales en torno a este ejemplo. Las demostraciones correspondientes pueden consultarse en [10].

Teorema 3.1.19. Definamos $\delta := (\alpha m_\gamma)^{\frac{1}{1-\gamma}}$ y supongamos que la Suposición 3.1.17 se cumple. Entonces las funciones de IV son

$$j_0(x) = 0, \quad \forall x \in S,$$

$$j_n(x) = \begin{cases} \left(\delta^{n-1} \frac{1-\delta}{1-\delta^n} \right)^{\gamma-1} \frac{b}{\gamma} x^\gamma, & x > 0, \\ 0, & x = 0, \end{cases} \quad n = 1, 2, \dots$$

Además, las funciones de decisión de IV están dadas por

$$f_n(x) = \frac{1-\delta^{n-1}}{1-\delta^n} x, \quad \forall x \in S, \quad n = 1, 2, \dots$$

Teorema 3.1.20. Asumiendo que la Suposición 3.1.17 se cumple. Entonces la función de valor óptimo J^* está dada por

$$J^*(0, x, \alpha) = \exp\left[(\gamma - 1) \ln\left(\frac{\delta-1}{\delta}\right)\right] \frac{b}{\gamma} x^\gamma, \quad x > 0,$$

y $J^*(0, 0, \alpha) = 0$. Más aún, la política óptima estacionaria está dada por

$$f^*(x) = \frac{x}{\delta}, \quad \forall x \geq 0.$$

Una de las motivaciones principales de presentar este ejemplo proviene de que una modificación del problema de consumo–inversión se estudiará en el siguiente capítulo, en el contexto sensible al riesgo.

Además, este ejemplo ofrece una conexión entre el caso neutral y el caso sensible al riesgo en relación con la construcción de una función de peso (como en la Proposición 3.1.10), ya que, como se verá más adelante, una condición fundamental para resolver un PDM descontado y sensible al riesgo consiste en exhibir una función de peso que acote el crecimiento del logaritmo de la función de recompensa.

3.2. Caso Sensible al Riesgo

En este capítulo se estudia la solución del PCO en el caso sensible al riesgo ($\lambda \neq 0$). El objetivo es resolver el problema de encontrar una política $\pi^* \in \Pi$ tal que

$$J(\lambda, \pi^*, x, \alpha) = \sup_{\pi \in \Pi} J(\lambda, \pi, x, \alpha) =: J^*(\lambda, x, \alpha), \quad \forall x \in S,$$

donde

$$J(\lambda, \pi, x, \alpha) = \frac{1}{\lambda} \ln \left(E_x^\pi \left[e^{\lambda \sum_{t=0}^{\infty} \alpha^t R(X_t, A_t)} \right] \right), \quad \lambda \neq 0. \quad (3.7)$$

En lo que sigue, consideraremos $\lambda \in (0, 1]$ y fijaremos $\alpha \in (0, 1)$.

Observación 3.2.1. Como la función $\ln(\cdot)$ es creciente y continua, el PCO anterior es equivalente a encontrar una política $\pi^* \in \Pi$ tal que

$$\mathcal{J}(\lambda, \pi^*, x) = \sup_{\pi \in \Pi} \mathcal{J}(\lambda, \pi, x) =: \mathcal{J}^*(\lambda, x), \quad \forall x \in S,$$

donde

$$\mathcal{J}(\lambda, \pi, x) := E_x^\pi \left[e^{\lambda \sum_{t=0}^{\infty} \alpha^t R(X_t, A_t)} \right]. \quad (3.8)$$

Así, el problema de control óptimo sensible al riesgo consiste en establecer la existencia de una política óptima y caracterizar la función de valor óptimo $\mathcal{J}^*(\lambda, x)$, la cual jugará un papel análogo al de la α -EOCD en el caso neutral al riesgo (ver Sección 3.1). En lo que sigue, se presentarán los supuestos técnicos y resultados teóricos necesarios para la solución de este problema.

3.2.1. Condiciones de Optimalidad

En analogía con lo estudiado en la Sección 3.1, en esta sección nos interesa establecer la existencia de solución para la siguiente **ecuación de optimalidad (EO)**:

$$V(x, \lambda) = \sup_{a \in A(x)} \left\{ e^{\lambda R(x,a)} \int_S V(y, \alpha\lambda) Q(dy | x, a) \right\}, \quad x \in S, \lambda \in (0, 1]. \quad (3.9)$$

Para garantizar la finitud de $\mathcal{J}(\lambda, \pi, x)$ y poder establecer la existencia de una solución $V(x, \lambda)$ de (3.9) con recompensa (posiblemente no acotada) $R(x, a)$, introducimos a continuación condiciones suficientes.

Nota 3.2.2. Si f es una función, c una constante y $x \in \text{Dom}(f)$, usaremos ocasionalmente la notación abreviada $f^c(x) := (f(x))^c$.

Hipótesis 3.2.3. Existen constantes $K \geq 0$, $L \geq 0$ y $m > 1$, así como una función $W \in \mathbb{M}(S)$ ($\mathbb{M}(S)$ denota el espacio de funciones medibles en S), con $W \geq 1$ en S , tales que:

- (i) $\int_S W^m(y) Q(dy | x, a) \leq K W^m(x)$ para todo $(x, a) \in \mathbb{K}$;
- (ii) $|R(x, a)| \leq \ln W(x) + L$ para todo $(x, a) \in \mathbb{K}$.

Una forma alternativa (y equivalente) de verificar 3.2.3 es la siguiente.

Proposición 3.2.4. La Hipótesis 3.2.3 es equivalente a: existen constantes $K \geq 0$, $L \geq 0$ y $\beta \in (0, 1)$, así como una función $U \in \mathbb{M}(S)$ con $U \geq 1$ en S , tales que:

- (a) $\int_S U(y) Q(dy | x, a) \leq K U(x)$ para todo $(x, a) \in \mathbb{K}$;
- (b) $|R(x, a)| \leq \beta \ln U(x) + L$ para todo $(x, a) \in \mathbb{K}$.

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

Demostración. Supongamos primero que la Hipótesis 3.2.3 es cierta. Definiendo $U(x) := W^m(x)$ para todo $x \in S$, se obtiene

$$\int_S U(y)Q(dy | x, a) \leq K U(x) \quad \text{y} \quad |R(x, a)| \leq \frac{1}{m} \ln U(x) + L, \quad \forall (x, a) \in \mathbb{K}.$$

Esto implica (a) y (b) con $\beta = \frac{1}{m}$.

Recíprocamente, supongamos que se cumplen (a) y (b). Definamos ahora $W(x) := U^\beta(x)$ para cada $x \in S$ y sea $m := 1/\beta$. Entonces, para todo $(x, a) \in \mathbb{K}$,

$$\int_S W^m(y)Q(dy | x, a) = \int_S U(y)Q(dy | x, a) \leq K U(x) = K W^m(x),$$

y además

$$|R(x, a)| \leq \beta \ln U(x) + L = \frac{1}{m} \ln W^m(x) + L = \ln W(x) + L.$$

De este modo se verifica la Hipótesis 3.2.3. ■

Con el fin de establecer la finitud de $\mathcal{J}(\pi, x, \lambda)$, se presenta a continuación el siguiente lema.

Lema 3.2.5. Supóngase que la Hipótesis 3.2.3 se cumple y que $\lambda \in (0, 1 - \alpha]$. Entonces, para toda política $\pi \in \Pi$ y todo $x \in S$ se tiene:

(a) Para $n \geq 1$ y $\gamma \in (0, 1]$,

$$E_x^\pi \left[e^{\gamma |R(X_n, A_n)|} \mid X_0, A_0, \dots, X_{n-1}, A_{n-1} \right] \leq e^{L\gamma} \left(K \frac{1}{m} W(X_{n-1}) \right)^\gamma.$$

(b) Para $n \geq 2$,

$$E_x^\pi \left[e^{\lambda \sum_{t=0}^{n-1} \alpha^t |R(X_t, A_t)|} \right] \leq e^{\lambda L \sum_{t=0}^{n-1} \alpha^t} K \frac{\lambda}{m} \sum_{t=1}^{n-1} \alpha^t (W(x))^{\lambda \sum_{t=0}^{n-1} \alpha^t}.$$

Demostración. (a) Para cualquier $n \geq 1$, de la Hipótesis 3.2.3 (i) y (a,c) de la Proposición B.0.1 en el Apéndice B, tenemos

$$\begin{aligned}
 E_x^\pi [W^\gamma(X_n) | X_0, A_0, \dots, X_{n-1}, A_{n-1}] &= \int_S W^\gamma(x_n) Q(dx_n | X_{n-1}, A_{n-1}) \\
 &\leq \left(\int_S W(x_n) Q(dx_n | X_{n-1}, A_{n-1}) \right)^\gamma \\
 &\leq \left(\left(\int_S W^m(x_n) Q(dx_n | X_{n-1}, A_{n-1}) \right)^{\frac{1}{m}} \right)^\gamma \\
 &\leq (KW^m(X_{n-1}))^{\frac{\gamma}{m}} \\
 &= \left(K^{\frac{1}{m}} W(W_{n-1}) \right)^\gamma.
 \end{aligned} \tag{3.10}$$

Luego, como $e^{|R(x,a)|} \leq e^L W(x)$ para $(x, a) \in \mathbb{K}$, para $n \geq 1$ se obtiene que

$$\begin{aligned}
 E_x^\pi [e^{\gamma|R(X_n, A_n)} | X_0, A_0, \dots, X_{n-1}, A_{n-1}] &\leq e^{\gamma L} \\
 &\quad \times E_x^\pi [(W(X_n))^\gamma | X_0, A_0, \dots, X_{n-1}, A_{n-1}] \\
 &\leq e^{\gamma L} \left(K^{\frac{1}{m}} W(W_{n-1}) \right)^\gamma,
 \end{aligned}$$

demostrando la parte (a).

Para la parte (b), consideremos cualquier $n \geq 2$ y tomando $\gamma = \lambda\alpha^n, \lambda\alpha^{n-2}, \frac{\lambda\alpha^n}{1-\alpha}$ (notar que en todos los casos se cumple que $\gamma \in (0, 1)$ pues $\lambda \in (0, 1 - \alpha)$), entonces de (a) y (3.10), así como también por (b) de la Proposición B.0.1 en el Apéndice B, vemos que

$$\begin{aligned}
& E_x^\pi \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, A_t)} \right] \\
&= E_x^\pi \left[e^{\lambda \sum_{t=0}^{n-1} \alpha^t R(X_t, A_t)} E_x^\pi \left[e^{\lambda \alpha^n R(X_n, A_n)} | X_0, A_0, \dots, X_{n-1}, A_{n-1} \right] \right] \\
&\leq E_x^\pi \left[e^{\lambda \sum_{t=0}^{n-1} \alpha^t R(X_t, A_t)} e^{\lambda \alpha^n L} \left(K \frac{1}{m} W(X_{n-1}) \right)^{\lambda \alpha^n} \right] \\
&= e^{\lambda \alpha^n L} K \frac{\lambda \alpha^n}{m} \\
&\times E_x^\pi \left[e^{\lambda \sum_{t=0}^{n-2} \alpha^t R(X_t, A_t)} E_x^\pi \left[e^{\lambda \alpha^{n-1} R(X_{n-1}, A_{n-1})} (W(X_{n-1}))^{\lambda \alpha^n} | X_0, \dots, X_{n-2}, A_{n-2} \right] \right] \\
&\leq e^{\lambda \alpha^n L} K \frac{\lambda \alpha^n}{m} E_x^\pi \left[e^{\lambda \sum_{t=0}^{n-2} \alpha^t R(X_t, A_t)} \left(E_x^\pi \left[e^{\lambda \alpha^{n-2} R(X_{n-1}, A_{n-1})} | X_0, \dots, A_{n-2} \right] \right)^\alpha \right] \\
&\leq e^{\lambda \alpha^n L} K \frac{\lambda \alpha^n}{m} \\
&\times E_x^\pi \left[e^{\lambda \sum_{t=0}^{n-2} \alpha^t R(X_t, A_t)} \left(e^{\lambda \alpha^{n-2} L} \left(K \frac{1}{m} W(X_{n-2}) \right)^{\lambda \alpha^{n-2}} \right)^\alpha \left(K \frac{1}{m} W(X_{n-2}) \right)^{\lambda \alpha^n} \right] \\
&= e^{\lambda \alpha^n L (\alpha^n + \alpha^{n-1})} K^{\frac{2\lambda \alpha^n + \lambda \alpha^{n-1}}{m}} E_x^\pi \left[e^{\lambda \sum_{t=0}^{n-2} \alpha^t R(X_t, A_t)} (W(X_{n-2}))^{\lambda \alpha^{n-1} + \lambda \alpha^n} \right] \\
&\leq e^{\lambda \alpha^n L (\alpha^n + \alpha^{n-1} + \alpha^{n-2})} K^{\frac{2\lambda \alpha^n + \lambda \alpha^{n-1}}{m}} \\
&\times E_x^\pi \left[e^{\lambda \sum_{t=0}^{n-3} \alpha^t R(X_t, A_t)} \left(K \frac{1}{m} W(X_{n-3}) \right)^{\lambda \alpha^{n-2} + \lambda \alpha^{n-1} + \lambda \alpha^n} \right] \\
&= e^{\lambda \alpha^n L (\alpha^n + \alpha^{n-1} + \alpha^{n-2})} K^{\frac{3\lambda \alpha^n + 2\lambda \alpha^{n-1} + \lambda \alpha^{n-2}}{m}} \\
&\times E_x^\pi \left[e^{\lambda \sum_{t=0}^{n-3} \alpha^t R(X_t, A_t)} (W(X_{n-3}))^{\lambda \alpha^{n-2} + \lambda \alpha^{n-1} + \lambda \alpha^n} \right] \\
&\leq e^{\lambda L (\alpha^n + \alpha^{n-1} + \dots + \alpha^2)} K^{\frac{(n-1)\lambda \alpha^n + \dots + \lambda \alpha^2}{m}} E_x^\pi \left[e^{\lambda \sum_{t=0}^1 \alpha^t R(X_t, A_t)} (W(X_1))^{\lambda (\alpha^n + \dots + \alpha^2)} \right] \\
&= e^{\lambda L (\alpha^n + \alpha^{n-1} + \dots + \alpha^2)} K^{\frac{(n-1)\lambda \alpha^n + \dots + \lambda \alpha^2}{m}} E_x^\pi \left[e^{\lambda R(x, a)} e^{\lambda \alpha R(X_1, A_1)} (W(X_1))^{\lambda (\alpha^n + \dots + \alpha^2)} \right] \\
&\leq e^{\lambda L (\alpha^n + \alpha^{n-1} + \dots + \alpha^2 + \alpha + 1)} K^{\frac{\lambda ((n-1)\alpha^n + \dots + \alpha^2)}{m}} E_x^\pi \left[W^\lambda(x) (W(X_1))^{\lambda (\alpha^n + \alpha^{n-1} + \dots + \alpha^2 + \alpha)} \right] \\
&\leq e^{\lambda L (1 + \alpha + \alpha^2 + \dots + \alpha^n)} K^{\frac{\lambda ((n-1)\alpha^n + (n-2)\alpha^{n-1} + \dots + \alpha^2)}{m}} W^\lambda(x) \left(K \frac{1}{m} W(x) \right)^{\lambda (\alpha^n + \alpha^{n-1} + \dots + \alpha)} \\
&= e^{\lambda L \sum_{t=0}^n \alpha^t} K^{\frac{\lambda}{m} \sum_{t=1}^n t \alpha^t} (W(x))^{\lambda \sum_{t=0}^n \alpha^t}.
\end{aligned}$$

Concluyendo (b). ■

Ahora podemos establecer la finitud de $\mathcal{J}(\pi, x, \lambda)$:

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

Teorema 3.2.6. Supóngase que la Hipótesis 3.2.3 se cumple y que $\lambda \in (0, 1 - \alpha)$. Entonces, para toda política $\pi \in \Pi$ y todo estado $x \in S$, se verifica que

$$|\mathcal{J}(\pi, x, \lambda)| \leq e^{\frac{L\lambda}{1-\alpha}} K^{\frac{\lambda}{m}} \left(\frac{\alpha}{(1-\alpha)^2} \right) (W(x))^{\frac{\lambda}{1-\alpha}}.$$

Demostración. Del Lema 3.2.5 (b) se tiene que

$$E_x^\pi \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, A_t)} \right] \leq e^{\lambda L \sum_{t=0}^n \alpha^t} K^{\frac{\lambda}{m} \sum_{t=1}^n t \alpha^t} (W(x))^{\lambda \sum_{t=0}^n \alpha^t}. \quad (3.11)$$

Al hacer $n \rightarrow \infty$ en (3.11), y notando que las sumas geométricas convergen por la hipótesis sobre λ , se obtiene el resultado. ■

A continuación se establece la existencia de una solución a la EO para el caso en que las funciones de recompensa son no negativas, mediante una técnica de iteración análoga a la utilizada en el escenario neutral al riesgo.

Teorema 3.2.7. Supongamos que la hipótesis 3.2.3 se cumple y que $R \geq 0$ sobre \mathbb{K} . Definamos $v_0(x, \lambda) \equiv 1$, y para $n \geq 0$ definamos $v_{n+1}(x, \lambda)$ mediante

$$v_{n+1}(x, \lambda) := \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x,a)} v_n(y, \alpha \lambda) Q(dy|x, a) \right\}, \quad (3.12)$$

con $x \in S$, $\lambda \in (0, 1 - \alpha)$. Entonces, para todo $n \geq 0$, $x \in S$ y $\lambda \in (0, 1 - \alpha)$, se cumple que:

- (a) $v_{n+1}(x, \lambda) \geq v_n(x, \lambda)$.
- (b) $v_{n+1}(x, \lambda) \leq e^{L\lambda \sum_{t=0}^n \alpha^t} K^{\frac{\lambda}{m} \sum_{t=0}^n t \alpha^t} (W(x))^{\lambda \sum_{t=0}^n \alpha^t}$.
- (c) El límite $v^*(x, \lambda) := \lim_{n \rightarrow \infty} v_n(x, \lambda)$ satisface la siguiente ecuación de optimalidad (3.13):

$$\begin{cases} v^*(x, \lambda) = \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x,a)} v^*(y, \alpha \lambda) Q(dy|x, a) \right\}, \\ 1 \leq v^*(x, \lambda) \leq e^{\frac{L\lambda}{1-\alpha}} K^{\frac{\lambda}{m}} \left(\frac{\alpha}{(1-\alpha)^2} \right) (W(x))^{\frac{\lambda}{1-\alpha}}. \end{cases} \quad (3.13)$$

Demostración. (a) Demostraremos por inducción que $v_{n+1}(x, \lambda) \geq v_n(x, \lambda)$ para cada

$n \geq 0$. Para $n = 0$, se tiene

$$\begin{aligned} v_1(x, \lambda) &= \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x,a)} v_0(y, \alpha\lambda) Q(dy|x, a) \right\} \\ &= \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x,a)} Q(dy|x, a) \right\} \\ &= \sup_{a \in A(x)} \left\{ e^{\lambda R(x,a)} \right\} \\ &\geq 1 = v_0(x, \lambda). \end{aligned}$$

Definamos ahora el operador T como

$$Tu(x, \lambda) := \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x,a)} u(y, \alpha\lambda) Q(dy|x, a) \right\}.$$

Notemos que T es monótono: si $u_1 \leq u_2$, entonces $Tu_1 \leq Tu_2$. Suponiendo que $v_{n+1} \geq v_n$ para algún $n \geq 1$, se obtiene

$$v_{n+2}(x, \lambda) = Tv_{n+1}(x, \lambda) \geq Tv_n(x, \lambda) = v_{n+1}(x, \lambda),$$

lo cual demuestra (a).

(b) Probaremos por inducción que, para $n \geq 0$,

$$v_{n+1}(x, \lambda) \leq e^{L\lambda \sum_{t=0}^n \alpha^t} K^{\frac{\lambda}{m} \sum_{t=0}^n t\alpha^t} (W(x))^{\lambda \sum_{t=0}^n \alpha^t}. \quad (3.14)$$

De la Hipótesis 3.2.3 se tiene $e^{\lambda|R(x,a)|} \leq e^{\lambda L} W(x)^\lambda$ para todo $(x, a) \in \mathbb{K}$. Por (3.12) y $v_0 \equiv 1$ se obtiene

$$v_1(x, \lambda) = \sup_{a \in A(x)} \left\{ e^{\lambda R(x,a)} \right\} \leq e^{\lambda L} W(x)^\lambda, \quad x \in S,$$

lo que verifica (3.14) para $n = 0$.

Supongamos ahora que (3.14) es cierto para algún $n \geq 0$. Entonces

$$\begin{aligned} v_{n+2}(x, \lambda) &= \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x,a)} v_{n+1}(y, \alpha\lambda) Q(dy|x, a) \right\} \\ &\leq \sup_{a \in A(x)} \left\{ e^{\lambda L} W(x)^\lambda \int_S v_{n+1}(y, \alpha\lambda) Q(dy|x, a) \right\} \\ &\leq e^{L\lambda \sum_{t=0}^{n+1} \alpha^t} K^{\frac{\lambda}{m} \sum_{t=0}^{n+1} t\alpha^t} (W(x))^{\lambda \sum_{t=0}^{n+1} \alpha^t}, \end{aligned}$$

lo que prueba (b).

(c) De (a) y (b) se sigue que el límite

$$v^*(x, \lambda) := \lim_{n \rightarrow \infty} v_n(x, \lambda)$$

existe. Además, de (3.12) y la monotonía de v_n , se cumple para todo $a \in A(x)$,

$$\int_S e^{\lambda R(x,a)} v_n(y, \alpha \lambda) Q(dy|x, a) \leq v_{n+1}(x, \lambda),$$

y al tomar el límite $n \rightarrow \infty$ se obtiene

$$\int_S e^{\lambda R(x,a)} v^*(y, \alpha \lambda) Q(dy|x, a) \leq v^*(x, \lambda).$$

Así,

$$v^*(x, \lambda) = \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x,a)} v^*(y, \alpha \lambda) Q(dy|x, a) \right\},$$

es decir, v^* satisface la ecuación de optimalidad (3.13), lo cual concluye la demostración. ■

Procedemos ahora a caracterizar la sucesión v_n introducida en el Teorema 3.2.7.

Teorema 3.2.8. Supongamos que la Hipótesis 3.2.3 se cumple y que $R \geq 0$ sobre \mathbb{K} . Entonces, para las funciones v_n y v^* definidas en el Teorema 3.2.7, se cumplen las siguientes afirmaciones:

(a) $v_n(x, \lambda) \geq E_x^\pi \left[e^{\lambda \sum_{t=0}^{n-1} \alpha^t R(X_t, A_t)} \right]$ para cada $\pi \in \Pi$, $n \geq 1$, $x \in S$, $\lambda \in (0, 1 - \alpha)$;

(b) $v^*(x, \lambda) \geq \sup_{\pi \in \Pi} \mathcal{J}(\pi, \lambda, x) = \mathcal{J}^*(x, \lambda)$ para todo $\lambda \in (0, 1 - \alpha)$ y $x \in S$.

Demostración. Dado que (b) se sigue directamente de (a), basta con demostrar (a) por inducción. Consideremos una política arbitraria $\pi = (\pi_0, \pi_1, \dots, \pi_n, \dots)$.

Para $n = 1$, por la definición de v_1 se tiene

$$\begin{aligned} v_1(x, \lambda) &= \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x,a)} Q(dy|x, a) \right\} \\ &= \sup_{a \in A(x)} E_x^\pi \left[e^{\lambda R(x,a)} \right] \\ &\geq E_x^\pi \left[e^{\lambda R(X_0, A_0)} \right], \quad x \in S, \lambda \in (0, 1 - \alpha). \end{aligned}$$

Ahora, supongamos que (a) es válida para algún $n \geq 1$, es decir,

$$v_n(x, \lambda) \geq E_x^\pi \left[e^{\lambda \sum_{t=0}^{n-1} \alpha^t R(X_t, A_t)} \right], \quad x \in S, \pi \in \Pi, \lambda \in (0, 1 - \alpha).$$

Sea $x \in S$ y $a \in A(x)$. Definamos $\pi^{x,a} := (\hat{\pi}_0, \hat{\pi}_1, \dots, \hat{\pi}_n, \dots)$ mediante

$$\hat{\pi}_k(\cdot | h_k) := \pi_{k+1}(\cdot | x, a, h_k),$$

donde $h_k = (y, a_0, \dots, x_{k-1}, a_{k-1}, x_k) \in \mathbb{H}_k$. Entonces, aplicando la hipótesis de inducción se obtiene

$$\begin{aligned} v_{n+1}(x, \lambda) &= \sup_{\tilde{a} \in A(x)} \left\{ \int_S e^{\lambda R(x, \tilde{a})} v_n(y, \alpha \lambda) Q(dy | x, \tilde{a}) \right\} \\ &\geq \int_S e^{\lambda R(x, a)} E_y^{\pi^{x,a}} \left[e^{\alpha \lambda \sum_{t=0}^n \alpha^t R(X_t, A_t)} \right] Q(dy | x, a) \pi_0(da | x) \\ &= E_x^\pi \left[e^{\lambda R(X_0, A_0) + \lambda \sum_{t=1}^n \alpha^t R(X_t, A_t)} \right] \\ &= E_x^\pi \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, A_t)} \right]. \end{aligned}$$

Con lo cual queda demostrada (a). ■

3.2.2. Existencia de una Política Óptima

En esta sección se demostrará la existencia de una política óptima. Para ello, se requieren las siguientes condiciones de continuidad-compacidad, las cuales son bien conocidas en la literatura.

Hipótesis 3.2.9. (i) Para cada $x \in S$, el conjunto $A(x)$ es compacto.

(ii) Para cada $x \in S$, tanto $R(x, a)$ como $\int_S u(y) Q(dy | x, a)$ son continuas en $a \in A(x)$ para toda función $u \in \mathbb{M}_b(S)$, donde $\mathbb{M}_b(S)$ denota el espacio de funciones medibles y acotadas en S .

(iii) Para cada $x \in S$, la función $\int_S W(y) Q(dy | x, a)$ es continua en $a \in A(x)$.

Ahora, para poder establecer uno de los resultados principales de esta sección, haremos uso del siguiente lema.

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

Lema 3.2.10. Bajo las Hipótesis 3.2.3 y 3.2.9, para cualquier $v(\cdot, \lambda) \in \mathbb{B}_W(S)$ con $\lambda \in (0, 1]$, la aplicación

$$a \mapsto \int_S e^{\lambda R(x,a)} v(y, \lambda) Q(dy|x, a)$$

es continua en $a \in A(x)$, para cada $x \in S$ fijo.

Demostración. Como $e^{\lambda R(x,a)}$ es continua en $a \in A(x)$, el resultado se sigue directamente del Lema 8.3.7 en [15]. ■

Podemos enunciar ahora el resultado principal de esta sección en el caso de funciones de recompensa no negativas.

Teorema 3.2.11. Bajo las Hipótesis 3.2.3 y 3.2.9, sea $v^*(\cdot, \lambda)$ la función definida en el Teorema 3.2.7. Si $R \geq 0$ sobre \mathbb{K} , entonces se cumplen las siguientes afirmaciones:

- (a) Para cualquier $\lambda \in (0, 1 - \alpha)$ existe un mapeo A -valuado $g^*(\cdot, \lambda)$ sobre S (dependiente de λ), tal que $g^*(x, \lambda) \in A(x)$ para todo $x \in S$, y

$$v^*(x, \lambda) = \int_S e^{\lambda R(x, g^*(x, \lambda))} v^*(y, \alpha \lambda) Q(dy|x, g^*(x, \lambda)), \quad x \in S.$$

- (b) Para cualquier $\lambda \in (0, 1 - \alpha)$, sea $\pi^* := (f_0^*, f_1^*, \dots, f_n^*, \dots)$, con $f_k^*(x) := g^*(x, \alpha^k \lambda)$ para $k \geq 0$ y $x \in S$. Entonces π^* es óptima y

$$\mathcal{J}(\pi^*, x, \lambda) = \mathcal{J}^*(x, \lambda) = v^*(x, \lambda), \quad x \in S.$$

Demostración. (a) Bajo la Hipótesis 3.2.3 y del Teorema 3.2.6, tenemos que

$$e^{\lambda R(x,a)} \leq e^{\lambda L} W(x),$$

y además

$$1 \leq v^*(x, \lambda) \leq e^{\frac{\lambda L}{1-\alpha}} K^{\frac{\lambda}{m}} \left(\frac{\alpha}{(1-\alpha)^2} \right) W(x)^{\frac{\lambda}{1-\alpha}} \leq e^{\frac{\lambda L}{1-\alpha}} K^{\frac{\lambda}{m}} \left(\frac{\alpha}{(1-\alpha)^2} \right) W(x),$$

para cada $(x, a) \in \mathbb{K}$ y $\lambda \in (0, 1 - \alpha)$. De aquí se sigue que

$$v^*(\cdot, \alpha \lambda) \in \mathbb{B}_W(S), \quad \lambda \in (0, 1 - \alpha).$$

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

Así, el Teorema de selección medible en [19], junto con el Lema 3.2.10, la Hipótesis 3.2.9 y la ecuación (3.13), garantiza la existencia de un mapeo A -valuado $g^*(\cdot, \lambda)$ sobre S (dependiente del valor de λ), tal que $g^*(x, \lambda) \in A(x)$ para todo $x \in S$ y

$$v^*(x, \lambda) = \int_S e^{\lambda R(x, g^*(x, \lambda))} v^*(y, \alpha \lambda) Q(dy | x, g^*(x, \lambda)). \quad (3.15)$$

Esto demuestra (a).

(b) A partir de (3.15), observamos que

$$v^*(x_n, \lambda) = \int_S e^{\lambda R(x_n, g^*(x_n, \lambda))} v^*(x_{n+1}, \alpha \lambda) Q(dx_{n+1} | x_n, g^*(x_n, \lambda)),$$

para $x_n \in S$, $n \geq 1$ y $\lambda \in (0, 1 - \alpha)$. Aplicando este argumento inductivamente y usando la definición de π^* , obtenemos

$$v^*(x, \lambda) = E_x^{\pi^*} \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, f_t^*(X_t))} v^*(X_{n+1}, \alpha^{n+1} \lambda) \right], \quad \forall n \geq 1. \quad (3.16)$$

Sea ahora $1 < p_1 < \frac{1-\alpha}{\lambda}$ y p_2 tal que $\frac{1}{p_1} + \frac{1}{p_2} = 1$. Elijamos N tal que $p_2 \alpha^{n+1} \lambda \leq 1 - \alpha$ para todo $n \geq N$. Entonces, de (3.16) y aplicando la desigualdad de Hölder, para todo $n \geq N$ se tiene

$$\begin{aligned} v^*(x, \lambda) &= E_x^{\pi^*} \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, f_t^*(X_t))} v^*(X_{n+1}, \alpha^{n+1} \lambda) \right] \\ &\leq \left(E_x^{\pi^*} \left[e^{p_1 \lambda \sum_{t=0}^n \alpha^t R(X_t, f_t^*(X_t))} \right] \right)^{\frac{1}{p_1}} \left(E_x^{\pi^*} \left[v^*(X_{n+1}, \alpha^{n+1} \lambda)^{p_2} \right] \right)^{\frac{1}{p_2}} \\ &\leq e^{\frac{L \alpha^{n+1} \lambda}{1-\alpha}} K^{\frac{\lambda}{m} \alpha^{n+1} \left(\frac{\alpha}{(1-\alpha)^2} \right)} \left(E_x^{\pi^*} \left[e^{p_1 \lambda \sum_{t=0}^n \alpha^t R(X_t, f_t^*(X_t))} \right] \right)^{\frac{1}{p_1}} \\ &\quad \times \left(E_x^{\pi^*} \left[W(X_{n+1})^{\frac{p_2 \alpha^{n+1} \lambda}{1-\alpha}} \right] \right)^{\frac{1}{p_2}}. \end{aligned}$$

Usando que $E_x^{\pi^*} [W(X_{n+1})] \leq K^{\frac{n+1}{m}} W(x)$ para todo $n \geq 1$ (ver (3.10)), se obtiene

$$v^*(x, \lambda) \leq e^{\frac{\alpha^{n+1} \lambda L}{1-\alpha}} K^{\frac{\alpha^{n+1} \lambda}{m} \left(\frac{\alpha}{(1-\alpha)^2} + \frac{n+1}{1-\alpha} \right)} W(x)^{\frac{\alpha^{n+1} \lambda}{1-\alpha}} \left(E_x^{\pi^*} \left[e^{p_1 \lambda \sum_{t=0}^n \alpha^t R(X_t, f_t^*(X_t))} \right] \right)^{\frac{1}{p_1}}. \quad (3.17)$$

Ahora, como $\sum_{n=0}^{\infty} \left(\frac{\alpha}{(1-\alpha)^2} + \frac{n+1}{1-\alpha} \right) \alpha^{n+1} < \infty$ y $\alpha \in (0, 1)$, se tiene que $\lim_{n \rightarrow \infty} \alpha^{n+1} = 0$ y el exponente asociado también tiende a cero. Por tanto, al tomar $n \rightarrow \infty$ en (3.17) resulta

$$v^*(x, \lambda) \leq \left(E_x^{\pi^*} \left[e^{p_1 \lambda \sum_{t=0}^{\infty} \alpha^t R(X_t, f_t^*(X_t))} \right] \right)^{\frac{1}{p_1}},$$

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

para todo p_1 tal que $1 < p_1 < \frac{1-\alpha}{\lambda}$. Finalmente, al hacer $p_1 \rightarrow 1$ y usando (3) de la Proposición B.0.1 en el Apéndice B, se concluye que

$$v^*(x, \lambda) \leq E_x^{\pi^*} \left[e^{\lambda \sum_{t=0}^{\infty} \alpha^t R(X_t, f_t^*(X_t))} \right] = \mathcal{J}(x, \pi^*, \lambda) = \mathcal{J}^*(x, \lambda). \quad (3.18)$$

Junto con el Teorema 3.2.8, se obtiene que $v^*(x, \lambda) = \mathcal{J}^*(x, \lambda)$. ■

Como segundo resultado principal, consideramos ahora el caso general de funciones de recompensa no acotadas ni superior ni inferiormente.

Teorema 3.2.12. Bajo las Hipótesis 3.2.3 y 3.2.9, supóngase que $m > 1$ y $K^{\frac{1}{m}}\alpha < 1$. Entonces se cumple:

- (a) La EO (3.13) tiene una única solución $v^*(\cdot, \lambda)$ (dependiente de $\lambda \in (0, 1 - \alpha)$), tal que $v^*(\cdot, \lambda) \in \mathbb{B}_W(S)$ y

$$\begin{cases} v^*(x, \lambda) = \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x,a)} v^*(y, \alpha\lambda) Q(dy | x, a) \right\}, \\ 1 - \frac{\lambda}{1 - \alpha K^{\frac{1}{m}}} W(x) - \frac{\lambda L}{1 - \alpha} \leq v^*(x, \lambda) \leq e^{\frac{\lambda L}{1-\alpha}} K^{\frac{\lambda}{m} \left(\frac{\alpha}{(1-\alpha)^2} \right)} (W(x))^{\frac{\lambda}{1-\alpha}}, \quad x \in S. \end{cases} \quad (3.19)$$

- (b) Si además $\lambda \in (0, (1 - \frac{1}{m})(1 - \alpha))$, entonces existe una política óptima $\pi^* \in \Pi_{DM}$ y

$$\mathcal{J}(\pi^*, x, \lambda) = \mathcal{J}^*(x, \lambda) = v^*(x, \lambda), \quad x \in S.$$

Demostración. (a) Para cualquier función real valuada $h(x, a)$ definida en \mathbb{K} , consideremos

$$V_h(x, \lambda) := \sup_{\pi \in \Pi} E_x^\pi \left[e^{\lambda \sum_{t=0}^{\infty} \alpha^t h(X_t, A_t)} \right], \quad x \in S, \lambda \in (0, 1 - \alpha),$$

siempre que $V_h(x, \lambda)$ esté bien definida. Para $n \geq 1$, definamos

$$R_n(x, a) := \max\{-n, R(x, a)\}, \quad G_n(x, a) := R_n(x, a) + n, \quad (x, a) \in \mathbb{K}.$$

De las hipótesis, $0 \leq G_n \leq |R| + n \leq \ln W + L + n$ en \mathbb{K} . Por el Teorema 3.2.11,

$$\begin{aligned} V_{G_n}(x, \lambda) &= \sup_{a \in A(x)} \left\{ \int_S e^{\lambda G_n(x,a)} V_{G_n}(y, \alpha\lambda) Q(dy | x, a) \right\} \\ &= \sup_{a \in A(x)} \left\{ \int_S e^{\lambda (R_n(x,a)+n)} V_{G_n}(y, \alpha\lambda) Q(dy | x, a) \right\}, \quad x \in S, \lambda \in (0, 1 - \alpha). \end{aligned}$$

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

Por otro lado, por definición de V_h ,

$$\begin{aligned} V_{G_n}(x, \lambda) &= V_{R_{n+n}}(x, \lambda) = \sup_{\pi \in \Pi} E_x^\pi \left[e^{\lambda \sum_{t=0}^{\infty} \alpha^t (R_{n+n})(X_t, A_t)} \right] \\ &= e^{\frac{\lambda n}{1-\alpha}} \sup_{\pi \in \Pi} E_x^\pi \left[e^{\lambda \sum_{t=0}^{\infty} \alpha^t R_n(X_t, A_t)} \right] = e^{\frac{\lambda n}{1-\alpha}} V_{R_n}(x, \lambda). \end{aligned}$$

Por tanto,

$$\begin{aligned} e^{\frac{\lambda n}{1-\alpha}} V_{R_n}(x, \lambda) &= \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R_n(x,a) + \lambda n} e^{\frac{\alpha \lambda n}{1-\alpha}} V_{R_n}(y, \alpha \lambda) Q(dy | x, a) \right\} \\ &= \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R_n(x,a)} e^{\frac{\lambda n}{1-\alpha}} V_{R_n}(y, \alpha \lambda) Q(dy | x, a) \right\}, \end{aligned}$$

lo que implica

$$V_{R_n}(x, \lambda) = \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R_n(x,a)} V_{R_n}(y, \alpha \lambda) Q(dy | x, a) \right\}, \quad x \in S, \lambda \in (0, 1 - \alpha). \quad (3.20)$$

Obsérvese que R_n es decreciente en $n \geq 1$, pues

$$R_n(x, a) = \max\{-n, R(x, a)\} \geq \max\{-(n+1), R(x, a)\} = R_{n+1}(x, a).$$

Podemos suponer que $V_{R_n}(\cdot, \lambda) \downarrow v^*(\cdot, \lambda)$ puntualmente (para cierta función medible v^* que depende de λ) cuando $n \rightarrow \infty$. De (3.20), para todo $a \in A(x)$,

$$V_{R_n}(x, \lambda) \geq \int_S e^{\lambda R_n(x,a)} V_{R_n}(y, \alpha \lambda) Q(dy | x, a) \geq \int_S e^{\lambda R_n(x,a)} v^*(y, \alpha \lambda) Q(dy | x, a),$$

y pasando al límite $n \rightarrow \infty$,

$$v^*(x, \lambda) \geq \int_S e^{\lambda R(x,a)} v^*(y, \alpha \lambda) Q(dy | x, a), \quad a \in A(x).$$

Por consiguiente,

$$v^*(x, \lambda) \geq \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x,a)} v^*(y, \alpha \lambda) Q(dy | x, a) \right\}, \quad x \in S, \lambda \in (0, 1 - \alpha). \quad (3.21)$$

Para la desigualdad opuesta, nótese que $V_{G_n} = e^{\frac{n\lambda}{1-\alpha}} V_{R_n}$ y $0 \leq G_n(x, a) \leq \ln W(x) + (L+n)$ en \mathbb{K} . Aplicando el Teorema 3.2.6,

$$V_{G_n}(x, \lambda) \leq e^{\frac{(L+n)\lambda}{1-\alpha}} K^{\frac{\lambda}{m}} \left(\frac{\alpha}{(1-\alpha)^2} \right) (W(x))^{\frac{\lambda}{1-\alpha}},$$

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

de donde

$$|V_{R_n}(x, \lambda)| \leq e^{\frac{\lambda L}{1-\alpha}} K^{\frac{\lambda}{m} \left(\frac{\alpha}{(1-\alpha)^2} \right)} (W(x))^{\frac{\lambda}{1-\alpha}} \leq e^{\frac{\lambda L}{1-\alpha}} K^{\frac{\lambda}{m} \left(\frac{\alpha}{(1-\alpha)^2} \right)} W(x), \quad (3.22)$$

para $n \geq 1$, $x \in S$, $\lambda \in (0, 1 - \alpha)$.

Así, fijados $x \in S$, $\lambda \in (0, 1 - \alpha)$, para cada $n \geq 1$, de (3.20) existe $a_n(x) \in A(x)$ tal que

$$V_{R_n}(x, \lambda) \leq \int_S e^{\lambda R_n(x, a_n(x))} V_{R_n}(y, \alpha \lambda) Q(dy | x, a_n(x)) + \frac{1}{n}. \quad (3.23)$$

Como $A(x)$ es compacto, existe una subsucesión $a_{n_s}(x) \rightarrow a^*(x) \in A(x)$. Entonces, por (3.22)–(3.23) y el Lema 8.3.7 de [15], al pasar al límite obtenemos

$$\begin{aligned} v^*(x, \lambda) &\leq \int_S e^{\lambda R(x, a^*(x))} v^*(y, \alpha \lambda) Q(dy | x, a^*(x)) \\ &\leq \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x, a)} v^*(y, \alpha \lambda) Q(dy | x, a) \right\}. \end{aligned}$$

Junto con (3.21) y (3.22), se concluye

$$\begin{aligned} v^*(x, \lambda) &= \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x, a)} v^*(y, \alpha \lambda) Q(dy | x, a) \right\}, \text{ y} \\ v^*(x, \lambda) &\leq e^{\frac{\lambda L}{1-\alpha}} K^{\frac{\lambda}{m} \left(\frac{\alpha}{(1-\alpha)^2} \right)} (W(x))^{\frac{\lambda}{1-\alpha}}, \end{aligned} \quad (3.24)$$

para todo $x \in S$ y $\lambda \in (0, 1 - \alpha)$.

Además, para todo $a \in A(x)$,

$$v^*(x, \lambda) \geq \int_S e^{\lambda R(x, a)} v^*(y, \alpha \lambda) Q(dy | x, a), \quad x \in S, \lambda \in (0, 1 - \alpha).$$

De aquí, para cualquier $\pi \in \Pi$ y $n \geq 0$,

$$\begin{aligned} v^*(X_n, \lambda) &\geq E_x^\pi \left[e^{\lambda R(X_n, A_n)} v^*(X_{n+1}, \alpha \lambda) \mid X_0, A_0, \dots, X_n, A_n \right], \text{ y} \\ &e^{\lambda \sum_{t=0}^{n-1} \alpha^t R(X_t, A_t)} v^*(X_n, \alpha^n \lambda) \\ &\geq E_x^\pi \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, A_t)} v^*(X_{n+1}, \alpha^{n+1} \lambda) \mid X_0, A_0, \dots, X_n, A_n \right]. \end{aligned}$$

Tomando esperanzas e iterando, se obtiene

$$v^*(x, \lambda) \geq E_x^\pi \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, A_t)} v^*(X_{n+1}, \alpha^{n+1} \lambda) \right], \quad n \geq 0, \pi \in \Pi, \lambda \in (0, 1 - \alpha). \quad (3.25)$$

Por otro lado,

$$\begin{aligned} V_{R_n}(x, \lambda) &= \sup_{\pi \in \Pi} E_x^\pi \left[e^{\lambda \sum_{t=0}^{\infty} \alpha^t R_n(X_t, A_t)} \right] \\ &\geq \sup_{\pi \in \Pi} E_x^\pi \left[e^{\lambda \sum_{t=0}^{\infty} \alpha^t R(X_t, A_t)} \right] \geq \sup_{\pi \in \Pi} \exp \left(\lambda E_x^\pi \left[\sum_{t=0}^{\infty} \alpha^t R(X_t, A_t) \right] \right), \end{aligned}$$

y, por (3.10) y la Hipótesis 3.2.3,

$$E_x^\pi \left[\sum_{t=0}^{\infty} \alpha^t |R(X_t, A_t)| \right] \leq \sum_{t=0}^{\infty} (\alpha K^{\frac{1}{m}})^t W(x) + \frac{L}{1-\alpha} = \frac{1}{1-\alpha K^{\frac{1}{m}}} W(x) + \frac{L}{1-\alpha}.$$

Así,

$$V_{R_n}(x, \lambda) \geq e^{-\lambda \left(\frac{1}{1-\alpha K^{\frac{1}{m}}} W(x) + \frac{L}{1-\alpha} \right)} \geq 1 - \frac{\lambda}{1-\alpha K^{\frac{1}{m}}} W(x) - \frac{L\lambda}{1-\alpha},$$

y en consecuencia,

$$v^*(x, \lambda) \geq 1 - \frac{\lambda}{1-\alpha K^{\frac{1}{m}}} W(x) - \frac{L\lambda}{1-\alpha}, \quad x \in S, \lambda \in (0, 1-\alpha),$$

lo cual, junto con (3.24), concluye (a).

(b) Para cualquier $\pi \in \Pi$, $n \geq 0$, $x \in S$, de (3.25),

$$\begin{aligned} v^*(x, \lambda) &\geq E_x^\pi \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, A_t)} \left(1 - \frac{\lambda \alpha^{n+1}}{1-\alpha K^{\frac{1}{m}}} W(X_{n+1}) - \frac{L\lambda \alpha^{n+1}}{1-\alpha} \right) \right] \\ &= E_x^\pi \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, A_t)} \right] \\ &\quad - E_x^\pi \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, A_t)} \left(\frac{\lambda \alpha^{n+1}}{1-\alpha K^{\frac{1}{m}}} W(X_{n+1}) + \frac{L\lambda \alpha^{n+1}}{1-\alpha} \right) \right]. \end{aligned}$$

Así, para $x \in S$, $\pi \in \Pi$, $\lambda \in (0, (1 - \frac{1}{m})(1 - \alpha))$, usando el Lema de Fatou obtenemos

$$v^*(x, \lambda) + \liminf_{n \rightarrow \infty} E_x^\pi \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, A_t)} \left(\frac{1}{1-\alpha K^{\frac{1}{m}}} + \frac{L}{1-\alpha} \right) W(X_{n+1}) \right] \geq \mathcal{J}(\pi, x, \lambda). \quad (3.26)$$

Se mostrará que

$$b_n := E_x^\pi \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, A_t)} \left(\frac{\lambda \alpha^{n+1}}{1-\alpha K^{\frac{1}{m}}} + \frac{L\lambda \alpha^{n+1}}{1-\alpha} \right) W(X_{n+1}) \right] \xrightarrow{n \rightarrow \infty} 0.$$

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

Como $\lambda \in (0, (1 - \frac{1}{m})(1 - \alpha))$, entonces $\frac{m\lambda}{m-1} \in (0, 1 - \alpha)$. Por 3.2.3 y el Teorema 3.2.6,

$$\begin{aligned} b_n &\leq \left(\frac{\lambda\alpha^{n+1}}{1 - \alpha K^{\frac{1}{m}}} + \frac{L\lambda\alpha^{n+1}}{1 - \alpha} \right) \left(E_x^\pi \left[e^{\frac{m\lambda}{m-1} \sum_{t=0}^n \alpha^t |R(X_t, A_t)|} \right] \right)^{\frac{m-1}{m}} \left(E_x^\pi [W^m(X_{n+1})] \right)^{\frac{1}{m}} \\ &\leq \left(\frac{\lambda\alpha^{n+1}}{1 - \alpha K^{\frac{1}{m}}} + \frac{L\lambda\alpha^{n+1}}{1 - \alpha} \right) \left(e^{\frac{\lambda L m}{(1-\alpha)(m-1)}} K^{\left(\frac{\lambda m}{m-1}\right)\left(\frac{\alpha}{(1-\alpha)^2}\right)} (W(x))^{\frac{\lambda m}{(1-\alpha)(m-1)}} \right)^{\frac{m-1}{m}} K^{\frac{n+1}{m}} W(x), \end{aligned}$$

donde usamos que $E_x^\pi [W^m(X_{n+1})] \leq K^{n+1} W(x)$ para $n \geq 1$ (Hip. 3.2.3). Por tanto,

$$0 \leq b_n \leq e^{\frac{\lambda L}{1-\alpha}} K^{\lambda\left(\frac{\alpha}{(1-\alpha)^2}\right)} (W(x))^{\frac{\lambda}{1-\alpha}} \left(\frac{\lambda}{1 - \alpha K^{\frac{1}{m}}} + \frac{L\lambda}{1 - \alpha} \right) (\alpha K^{\frac{1}{m}})^{n+1} W(x) \rightarrow 0,$$

cuando $n \rightarrow \infty$. Con (3.26), se concluye

$$v^*(x, \lambda) \geq \mathcal{J}(\pi, x, \lambda), \quad \pi \in \Pi, \quad x \in S. \quad (3.27)$$

Por el Teorema de selección medible [19], existe un mapeo medible A -valuado $f^*(\cdot, \lambda)$ (dependiente de λ) sobre S tal que $f^*(x, \lambda) \in A(x)$ y

$$v^*(x, \lambda) = \int_S e^{\lambda R(x, f^*(x, \lambda))} v^*(y, \alpha\lambda) Q(dy | x, f^*(x, \lambda)), \quad x \in S.$$

Sea $\pi^* = (f_0^*, f_1^*, \dots)$ con $f_n^*(x) := f^*(x, \alpha^n \lambda)$, $n \geq 0$. Como en (3.16), para $x \in S$ y $\lambda \in (0, (1 - \frac{1}{m})(1 - \alpha))$,

$$v^*(x, \lambda) = E_x^{\pi^*} \left[e^{\lambda \sum_{t=0}^n \alpha^t R(X_t, f_t^*(X_t))} v^*(X_{n+1}, \alpha^{n+1} \lambda) \right], \quad n \geq 0. \quad (3.28)$$

Repitiendo el argumento final de (3.18), se obtiene

$$v^*(x, \lambda) \leq \mathcal{J}(\pi^*, x, \lambda).$$

Con (3.27), concluimos

$$v^*(x, \lambda) = \mathcal{J}(\pi^*, x, \lambda) = \sup_{\pi \in \Pi} \mathcal{J}(\pi, x, \lambda),$$

por lo que π^* es óptima.

Finalmente, si $\mathcal{L}(\cdot, \lambda)$ es otra función que satisface (3.19), repitiendo el argumento anterior se deduce que $\mathcal{L}(x, \lambda) = \mathcal{J}(x, \lambda)$ para todo $x \in S$, lo que prueba la unicidad. ■

3.2.3. Ejemplo

A continuación, se presenta un ejemplo que ilustra los resultados obtenidos en la solución del PCO descontado sensible al riesgo. El modelo a considerar es una variante del **modelo de consumo–inversión** estudiado en [15].

Planteamiento

Dada la siguiente modificación del modelo de consumo–inversión inducida por la ecuación en diferencias

$$x_{n+1} = \delta \xi_n (x_n + a_n), \quad n = 0, 1, \dots, \quad (3.29)$$

se considerará que:

- $S = A = [1, \infty)$.
- $A(x) = [1, x]$ para cada $x \in S$.

Las perturbaciones ξ_n en (3.29) se asumen como variables aleatorias continuas, independientes e idénticamente distribuidas (v.a.i.i.d.) con recorrido mayor o igual que 1 y con $E[\xi^2] < \infty$, además de que $\delta \geq \frac{1}{2}$.

Si f es la función de densidad común de las v.a. ξ_n , al usar (1.3) se obtiene

$$Q(C|x, a) = \int_S I_C[\delta y(x + a)] f(y) dy, \quad \forall x \in S, a \in A(x), C \in \mathcal{B}(S).$$

Así, cuando el estado del sistema es $x \in S$, el controlador selecciona una acción $a \in A(x)$, y el par (x, a) genera una recompensa $R(x, a)$ antes de que el proceso transite a otro estado mediante Q .

Estudio del modelo

Se propone la siguiente función de recompensa:

$$R(x, a) := \ln \left(\sqrt{\frac{x^2 + a^2}{2}} \right), \quad x \in S, a \in A(x).$$

Como se señaló en la sección 3.1 (caso neutral al riesgo) y en las referencias allí citadas, el estudio del problema de consumo–inversión resulta fundamental para

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

comprender cómo los agentes económicos toman decisiones que afectan su bienestar presente y futuro. En esta sección se analiza una modificación de dicho modelo, la cual permite obtener resultados adicionales y, en particular, ilustrar la aplicación de la técnica de iteración de valores en el caso sensible al riesgo.

Proposición 3.2.13. La modificación del modelo de consumo–inversión considerada en esta sección satisface las **Hipótesis** 3.2.3 y 3.2.9.

Demostración. Consideremos como función de peso a la función W dada por

$$W(x) := x \quad \text{con} \quad x \in S.$$

Notar que es claro que $W \geq 1$ sobre S , notar también que, si tomamos $x \in S$ y $a \in A(x) = [1, x]$, entonces

$$\begin{aligned} |R(x, a)| &= \left| \ln \left(\sqrt{\frac{x^2 + a^2}{2}} \right) \right| \\ &\leq |\ln(|x|)| \quad \text{pues} \quad 1 \leq a \leq x \\ &= \ln(x) \quad \text{pues} \quad x > 1 \\ &\leq \ln W(x) + L, \end{aligned}$$

donde $L > 0$ es una constante arbitraria. Obteniendo que se cumple 3.2.3 a).

Para 3.2.3 b), observar que si tomamos $m = 2$, entonces

$$\begin{aligned} \int_S W^m(y) Q(dy|x, a) &= E [(\delta\xi(x + a))^2] \\ &= \delta^2(x + a)^2 E [\xi^2] \\ &\leq 4\delta^2 E [\xi^2] x^2 \\ &= KW^m(x) \end{aligned}$$

es decir, se cumple 1 (b) con $m = 2$ y $K = 4\delta^2 E [\xi^2] > 0$.

Ahora, en torno al segundo juego de hipótesis (**Hipótesis 3.2.9**), vemos que 3.2.9 i) claramente se cumple pues $A(x) = [1, x]$, el cual es compacto para cada $x \in S$.

Finalmente, las condiciones ii) y iii) se siguen como consecuencia del hecho de que la función de densidad de la distribución común de las v.a.i.i.d ξ_n es continua y del **Teorema de la Convergencia Dominada** (ver Apéndice C).

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

Por lo tanto, se concluye que éste modelo cumple las condiciones de las Hipótesis 3.2.3 y 3.2.9. ■

Como consecuencia de la proposición 3.2.13 obtenemos inmediatamente el siguiente resultado.

Corolario 3.2.14. Para cualquier factor de descuento α y factor de sensibilidad al riesgo $\lambda \in (0, 1 - \alpha)$, existe una política óptima $\pi^* \in \Pi_{DM}$ para el modelo de consumo-inversión descrito en ésta sección.

Demostración. Se sigue del Teorema 3.2.11 pues el modelo en cuestión satisface las Hipótesis 3.2.3 y 3.2.9. ■

Ejemplo iteración de valores: caso sensible al riesgo

En esta subsección procederemos a calcular las funciones de iteración de valores para el ejemplo de consumo-inversión considerado.

Recordemos que dichas funciones se definen como $v_0(x, \lambda) \equiv 1$, y para $n \geq 0$,

$$v_{n+1}(x, \lambda) := \sup_{a \in A(x)} \left\{ \int_S e^{\lambda R(x,a)} v_n(y, \alpha\lambda) Q(dy|x, a) \right\}, \quad x \in S, \lambda \in (0, 1 - \alpha).$$

Se obtiene el siguiente resultado particular para este ejemplo.

Proposición 3.2.15. Las funciones de iteración de valores están dadas por

$$v_{n+1}(x, \lambda) = (2\delta)^\lambda \sum_{i=1}^n i\alpha^i \prod_{j=0}^{n-1} E \left[\xi^{\lambda \sum_{i=j+1}^n \alpha^i} \right] x^{\lambda \sum_{i=0}^n \alpha^i}, \quad n \in \mathbb{N}.$$

Considerando la convención $\sum_{i=1}^0 a_i := 0$ y $\prod_{j=0}^{-1} a_j := 1$ para cualquier sucesión $\{a_n\}$.

Demostración. El resultado se demuestra por inducción. Para $n = 1$ se tiene

$$v_1(x, \lambda) = \sup_{a \in [1,x]} \{e^{\lambda R(x,a)}\} = \sup_{a \in [1,x]} \left\{ \left(\sqrt{\frac{x^2+a^2}{2}} \right)^\lambda \right\} = x^\lambda,$$

donde el supremo se alcanza en $a = x$, de modo que $f_1(x) = x$ es el selector óptimo.

Para $n = 2$,

$$\begin{aligned}
 v_2(x, \lambda) &= \sup_{a \in [1, x]} \left\{ \left(\sqrt{\frac{x^2 + a^2}{2}} \right)^\lambda E[v_1(\delta\xi(x + a), \alpha\lambda)] \right\} \\
 &= \sup_{a \in [1, x]} \left\{ \left(\sqrt{\frac{x^2 + a^2}{2}} \right)^\lambda E[(\delta\xi(x + a))^{\alpha\lambda}] \right\} \\
 &= \delta^{\alpha\lambda} E[\xi^{\alpha\lambda}] \sup_{a \in [1, x]} \left\{ \left(\sqrt{\frac{x^2 + a^2}{2}} \right)^\lambda (x + a)^{\alpha\lambda} \right\} \\
 &= (2\delta)^{\alpha\lambda} E[\xi^{\alpha\lambda}] x^{\lambda + \alpha\lambda},
 \end{aligned}$$

lo cual coincide con la fórmula de la proposición para $n = 1$. El selector óptimo sigue siendo $f_2(x) = x$.

Suponiendo ahora la hipótesis inductiva para n , se tiene

$$v_n(x, \lambda) = (2\delta)^{\lambda \sum_{i=1}^{n-1} i \alpha^i} \prod_{j=0}^{n-2} E[\xi^{\lambda \sum_{i=j+1}^{n-1} \alpha^i}] x^{\lambda \sum_{i=0}^{n-1} \alpha^i}.$$

Sustituyendo esta expresión en la definición de v_{n+1} y realizando los cálculos, se obtiene finalmente

$$v_{n+1}(x, \lambda) = (2\delta)^{\lambda \sum_{i=1}^n i \alpha^i} \prod_{j=0}^{n-1} E[\xi^{\lambda \sum_{i=j+1}^n \alpha^i}] x^{\lambda \sum_{i=0}^n \alpha^i},$$

con lo que se concluye la demostración. ■

Observación 3.2.16. En el resultado anterior, el selector óptimo siempre coincide con $a = x$, es decir,

$$f_n(x) = x, \quad \forall n \in \mathbb{N}. \tag{3.30}$$

A una política que satisface (3.30) se le denomina **política miope**.

Este hecho puede interpretarse como que, en cada etapa, la estrategia óptima consiste en elegir siempre el mayor nivel posible accesible en ese instante.

Nota 3.2.17. En el resultado siguiente se demuestra la convergencia del límite de los productos mostrados en la proposición 3.2.15, para lograrlo, se utilizaron resultados que tienen que ver sobre expansiones en serie de Taylor y el uso de la notación de Landau ($o(\cdot)$) y sus propiedades, para ver los resultados que se usaron se puede revisar el apéndice D o bien [6].

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

Para continuar, observar que $\sum_{i=j+1}^n \alpha^i = \frac{\alpha^{j+1} - \alpha^{n+1}}{1 - \alpha}$, se tienen ahora los siguientes resultados.

Proposición 3.2.18. Para el caso en el que $\xi \sim Unif(1, M)$ con $M > 1$ una constante, se tiene que el límite cuando $n \rightarrow \infty$ de los productos

$$\prod_{j=0}^{n-1} E \left[\xi^{\lambda \sum_{i=j+1}^n \alpha^i} \right]$$

converge.

Demostración. Para verificarlo, primero notar que si $\xi \sim Unif(1, M)$ con $M > 1$, entonces para cada $j, n \in \mathbb{N}$,

$$\begin{aligned} E \left[\xi^{\lambda \sum_{i=j+1}^n \alpha^i} \right] &= E \left[\xi^{\lambda \left(\frac{\alpha^{j+1} - \alpha^{n+1}}{1 - \alpha} \right)} \right] \\ &= \frac{1}{M - 1} \int_1^M y^{\lambda \left(\frac{\alpha^{j+1} - \alpha^{n+1}}{1 - \alpha} \right)} dy \\ &= \frac{M^{h_{j,n} + 1} - 1}{(h_{j,n} + 1)(M - 1)}, \end{aligned}$$

donde $h_{j,n} := \lambda \left(\frac{\alpha^{j+1} - \alpha^{n+1}}{1 - \alpha} \right)$ Luego, observemos que

$$\begin{aligned} M^{h_{j,n} + 1} - 1 &= M M^{h_{j,n}} - 1 \\ &= M e^{h_{j,n} \ln M} - 1, \end{aligned}$$

entonces, puesto que $M > 1$ es fijo y $h_{j,n} \rightarrow 0$ cuando $j, n \rightarrow \infty$, al usar la expansión en serie de Taylor de primer orden para $e^{h_{j,n} \ln M}$ vemos que

$$\begin{aligned} M^{h_{j,n} + 1} - 1 &= M (1 + h_{j,n} \ln M + o(h_{j,n} \ln M)) - 1 \\ &= M - 1 + M h_{j,n} \ln M + o(h_{j,n}), \end{aligned}$$

luego

$$\begin{aligned}
 \ln(M^{h_{j,n}+1} - 1) &= \ln(M - 1 + Mh_{j,n} \ln M + o(h_{j,n})) \\
 &= \ln\left((M - 1) \left(1 + \frac{Mh_{j,n} \ln M}{M - 1} + \frac{o(h_{j,n})}{M - 1}\right)\right) \\
 &= \ln\left((M - 1) \left(1 + \frac{Mh_{j,n} \ln M}{M - 1} + o(h_{j,n})\right)\right) \\
 &= \ln(M - 1) + \ln\left(1 + \frac{Mh_{j,n} \ln M}{M - 1} + o(h_{j,n})\right),
 \end{aligned}$$

donde la tercera igualdad se da de las propiedades de $o(\cdot)$; de manera que al usar el desarrollo en primer orden de $\ln(1 + x)$ con $x = \frac{Mh_{j,n} \ln M}{M - 1} + o(h_{j,n})$ llegamos a que

$$\begin{aligned}
 \ln(M^{h_{j,n}+1} - 1) &= \ln(M - 1) + \frac{Mh_{j,n} \ln M}{M - 1} + o(h_{j,n}) + o(h_{j,n}) \\
 &= \ln(M - 1) + \frac{Mh_{j,n} \ln M}{M - 1} + o(h_{j,n}).
 \end{aligned}$$

Esto es, obtenemos que

$$\ln(M^{h_{j,n}+1} - 1) = \ln(M - 1) + \frac{Mh_{j,n} \ln M}{M - 1} + o(h_{j,n}). \quad (3.31)$$

Por otro lado, usando una técnica similar vemos que

$$\begin{aligned}
 \ln((h_{j,n} + 1)(M - 1)) &= \ln(M - 1) + \ln(1 + h_{j,n}) \\
 &= \ln(M - 1) + h_{j,n} + o(h_j),
 \end{aligned}$$

es decir, que

$$\ln((h_{j,n} + 1)(M - 1)) = \ln(M - 1) + h_{j,n} + o(h_j). \quad (3.32)$$

Así, puesto que

$$\lim_{n \rightarrow \infty} \prod_{j=0}^{n-1} E[\xi^{h_{j,n}}] = \prod_{j=0}^{\infty} E[\xi^{h_j}],$$

en donde $h_j := \frac{\lambda \alpha^{j+1}}{1 - \alpha}$, y como

$$\prod_{j=0}^{\infty} E [\xi^{h_j}] < \infty \Leftrightarrow \sum_{j=0}^{\infty} \ln (E [\xi^{h_j}]) < \infty,$$

lo anterior pues $\prod_{j=0}^{\infty} E [\xi^{h_j}] = e^{\ln(\prod_{j=0}^{\infty} E[\xi^{h_j}])} = e^{\sum_{j=0}^{\infty} \ln(E[\xi^{h_j}])}$. Entonces analizamos la serie, de manera que, por (3.31) y (3.32),

$$\begin{aligned} \sum_{j=0}^{\infty} \ln (E [\xi^{h_j}]) &= \sum_{j=0}^{\infty} [\ln (M^{h_{j,n}+1} - 1) - \ln ((h_{j,n} + 1)(M - 1))] \\ &= \sum_{j=0}^{\infty} \left\{ \left[\ln (M - 1) + \frac{M h_{j,n} \ln M}{M - 1} + o(h_{j,n}) \right] - [\ln (M - 1) + h_{j,n} + o(h_j)] \right\} \\ &= \sum_{j=0}^{\infty} \left[\left(\frac{M \ln M}{M - 1} - 1 \right) h_j + o(h_j) \right] \\ &= \sum_{j=0}^{\infty} [C h_j + o(h_j)] \quad \text{haciendo } C := \frac{M \ln M}{M - 1} - 1 \\ &= \frac{\lambda C}{1 - \alpha} \sum_{j=0}^{\infty} \alpha^{j+1} + \sum_{j=0}^{\infty} o(h_j) \\ &= \frac{\alpha \lambda C}{(1 - \alpha)^2} + \sum_{j=0}^{\infty} o(h_j). \end{aligned}$$

Por último, sólo basta ver que la serie $\sum_{j=0}^{\infty} o(h_j)$ converge, lo cual se obtiene de la siguiente manera: se tiene que existe una sucesión $\{\varepsilon_j\}_{j \in \mathbb{N}}$ tal que

- i) $\varepsilon_j \rightarrow 0$ y
- ii) $|o(h_j)| \leq \varepsilon_j h_j$,

de manera que

$$\begin{aligned} \sum_{j=0}^{\infty} |o(h_j)| &\leq \sum_{j=0}^{\infty} \varepsilon_j h_j \\ &= \frac{\lambda}{1 - \alpha} \sum_{j=0}^{\infty} \varepsilon_j \alpha^{j+1} \\ &\leq \frac{\lambda}{1 - \alpha} \sum_{j=0}^{\infty} \alpha^{j+1} \quad \text{pues } \varepsilon_j \rightarrow 0. \end{aligned}$$

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

Por tanto, como la serie geométrica $\sum_{j=0}^{\infty} \alpha^{j+1}$ es convergente, se concluye que la serie $\sum_{j=0}^{\infty} o(h_j)$ converge absolutamente, de donde se sigue la convergencia y, en consecuencia, el resultado principal. ■

Función de valor óptimo: obtención y aproximación

Notación 3.2.19. En lo que sigue, consideraremos que $\xi \sim Unif(1, M)$ con $M > 1$, y en vista de la Proposición 3.2.18 denotaremos por

$$P_n := \prod_{j=0}^{n-2} E \left[\xi^{\lambda \sum_{i=j+1}^{n-1} \alpha^i} \right] \quad \text{y} \quad P := \lim_{n \rightarrow \infty} \prod_{j=0}^{n-2} E \left[\xi^{\lambda \sum_{i=j+1}^{n-1} \alpha^i} \right].$$

Se tiene el siguiente resultado principal.

Teorema 3.2.20. La función de valor óptimo λ -sensible al riesgo y α -descontada, con $\lambda \in (0, 1 - \alpha)$, para este modelo está dada por

$$v^*(x, \lambda) = (2\delta)^{\frac{\lambda\alpha}{(1-\alpha)^2}} P x^{\frac{\lambda}{1-\alpha}}, \quad x \in S.$$

Además, el selector óptimo está dado por $f^*(x, \lambda) = x$.

Demostración. En efecto, se tiene

$$\begin{aligned} \lim_{n \rightarrow \infty} v_n(x, \lambda) &= \lim_{n \rightarrow \infty} \left[(2\delta)^{\lambda \sum_{i=1}^{n-1} i \alpha^i} \prod_{j=0}^{n-2} E \left[\xi^{\lambda \sum_{i=j+1}^{n-1} \alpha^i} \right] x^{\lambda \sum_{i=0}^{n-1} \alpha^i} \right] \\ &= \lim_{n \rightarrow \infty} \left[(2\delta)^{\lambda \sum_{i=1}^{n-1} i \alpha^i} P_n x^{\lambda \sum_{i=0}^{n-1} \alpha^i} \right] \\ &= (2\delta)^{\frac{\lambda\alpha}{(1-\alpha)^2}} P x^{\frac{\lambda}{1-\alpha}}. \end{aligned}$$

El hecho de que $f^*(x, \lambda) = x$ se sigue de la Observación 3.2.16, donde se estableció que la política óptima obtenida corresponde a una política miope. ■

La Tabla 3.1 resume los resultados obtenidos de éste ejemplo.

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

	Función de iteración de valores	Función de valor óptimo
Función	$v_n(x, \lambda) = (2\delta)^\lambda \sum_{i=1}^{n-1} i \alpha^i \prod_{j=0}^{n-2} E \left[\xi^{\lambda \sum_{i=j+1}^{n-1} \alpha^i} \right] x^{\lambda \sum_{i=0}^{n-1} \alpha^i}$	$v^*(x, \lambda) = (2\delta)^{\frac{\lambda \alpha}{(1-\alpha)^2}} P x^{\frac{\lambda}{1-\alpha}}$
Política óptima	$f_n(x, \lambda) = x$	$f^*(x, \lambda) = x \text{ y } \pi^* = (x, x, \dots, x, \dots)$

Tabla 3.1: Resultados obtenidos del ejemplo.

Para concluir este capítulo, se incluye una última sección en la que se presentan las gráficas correspondientes a la función de valor óptimo para el ejemplo particular considerado, en el que $\xi \sim Unif(1, M)$ con $M > 1$.

Conclusión del ejemplo: aproximaciones al valor de P

En la Tabla 3.2 se presentan aproximaciones al valor de P , considerando seis decimales de precisión, con $M = 3$ y $\delta = 0,7$, y variando los valores de α y λ .

Al analizar las diferencias $P_{n_{m+1}} - P_{n_m}$, donde $m \in \{1, 2, 3, 4\}$ y P_{n_m} denota el producto truncado en el término n_m -ésimo, se observa que el valor de P_{n_m} a partir del cual dichas diferencias se anulan, independientemente de los valores de α y λ , corresponde a n_3 .

Con el fin de determinar con mayor precisión el valor de n que permite aproximar P bajo un error dado ε , se implementó un código en Python (ver Apéndice E), cuyos resultados se muestran en la Tabla 3.3.

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

α	λ	$n_1 = 10$	$n_2 = 100$	$n_3 = 1000$	$n_4 = 10000$
0.9	0.01	1.166843	1.794910	1.795238	1.795238
	0.05	2.1807976	19.385651	19.403992	19.403992
	0.09	4.128669	222.750000	1.483461	1.483461
0.8	0.01	1.084390	1.138588	1.138588	1.138588
	0.1	2.274612	3.730880	3.730880	3.730880
	0.19	4.869918	12.632599	12.632599	12.632599
0.5	0.01	1.012908	1.013048	1.013048	1.013048
	0.25	1.383161	1.388011	1.388011	1.388011
	0.49	1.902093	1.915277	1.915277	1.915277
0.2	0.01	1.002025	1.002025	1.002025	1.002025
	0.4	1.084892	1.084892	1.084892	1.084892
	0.79	1.175701	1.175701	1.175701	1.175701
0.1	0.01	1.000801	1.000801	1.000801	1.000801
	0.45	1.036775	1.036775	1.036775	1.036775
	0.89	1.074287	1.074287	1.074287	1.074287

Tabla 3.2: Aproximaciones al valor de P .

α_i	$\lambda_{i,j}$	$\varepsilon = 10^{-5}$	$\varepsilon = 10^{-10}$	$\varepsilon = 10^{-15}$	$\varepsilon = 10^{-20}$	$\varepsilon = 10^{-25}$
$\alpha_1 = 0.9$	$\lambda_{1,1} = 0.03$	170	290	400	400	400
	$\lambda_{1,2} = 0.06$	180	300	410	410	410
$\alpha_1 = 0.5$	$\lambda_{2,1} = 0.1$	40	50	70	70	70
	$\lambda_{2,2} = 0.4$	40	60	70	70	70
$\alpha_1 = 0.1$	$\lambda_{3,1} = 0.2$	10	30	30	30	30
	$\lambda_{3,2} = 0.8$	10	30	30	30	30

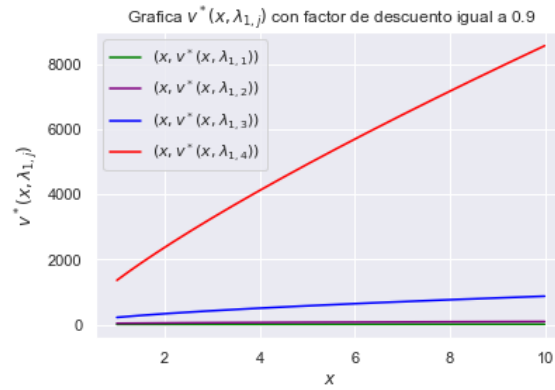
Tabla 3.3: Mejores valores de n para aproximar el valor de P dado un error ε .

Por tanto, con base en los valores obtenidos en las Tablas 3.2 y 3.3, para la representación gráfica de la función de valor óptimo se adopta la aproximación correspondiente a $n = 410$, ya que este es el mayor valor que garantiza una precisión aceptable en la estimación del producto P , considerando un error $\varepsilon = 10^{-20}$.

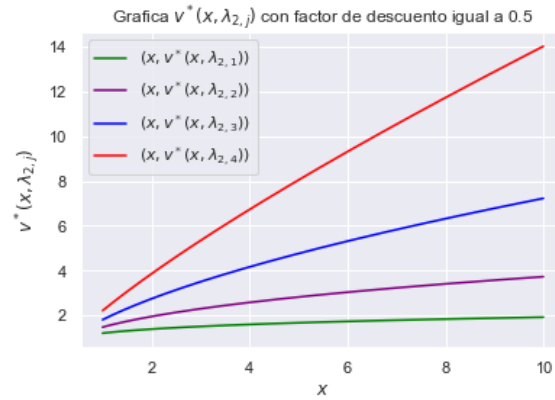
De este modo, se presentan a continuación las gráficas de $v^*(x, \lambda)$ para el caso en que $M = 3$ y $\delta = 0.7$, analizando tres valores del factor de descuento: $\alpha_1 = 0.9$, $\alpha_2 = 0.5$ y $\alpha_3 = 0.1$. Para cada uno de estos valores de α se incluyen, además, cuatro valores correspondientes del parámetro de sensibilidad al riesgo λ .

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

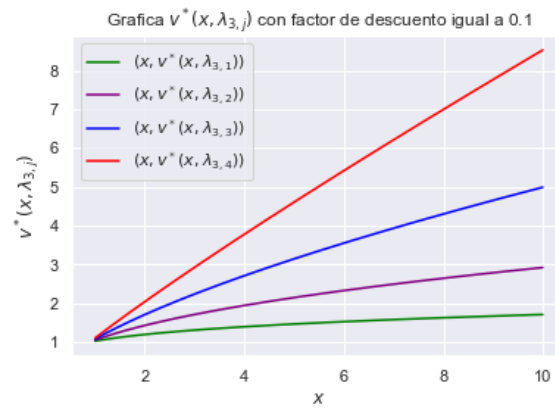
3.2 Caso Sensible al Riesgo



(a) Gráfica α_1 .



(b) Gráfica α_2 .



(c) Gráfica α_3 .

Figura 3.1: Gráficas de $v^*(x, \lambda_{i,j})$.

PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio de Recompensas

3.2 Caso Sensible al Riesgo

Y los valores de α y λ considerados para las gráficas de la figura 3.1 son los que se presentan en la tabla 3.4.

α_i	$\lambda_{i,j}$
$\alpha_1=0.9$	$\lambda_{1,1}=0.02$
	$\lambda_{1,2}=0.04$
	$\lambda_{1,3}=0.06$
	$\lambda_{1,4}=0.08$
$\alpha_2=0.5$	$\lambda_{2,1}=0.1$
	$\lambda_{2,2}=0.2$
	$\lambda_{2,3}=0.3$
	$\lambda_{2,4}=0.4$
$\alpha_3=0.1$	$\lambda_{3,1}=0.2$
	$\lambda_{3,2}=0.4$
	$\lambda_{3,3}=0.6$
	$\lambda_{3,4}=0.8$

Tabla 3.4: Valores de α y λ considerados para gráficas.

Donde $\lambda_{i,j}$ es el factor de sensibilidad al riesgo correspondiente al factor de descuento α_i , con $i \in \{1, 2, 3\}$ y $j \in \{1, 2, 3, 4\}$.

Conclusiones de las gráficas

Para finalizar, a partir de las gráficas de la Figura 3.1 puede observarse que el comportamiento de la función \mathcal{J}^* muestra un crecimiento más pronunciado conforme aumenta el factor de descuento. En términos prácticos, esto se interpreta como que al considerar factores de descuento grandes, la función de valor óptimo tiende a devolver valores de mayor magnitud.

Conclusiones

A lo largo de la presente tesis se estudiaron los PDMs Descontados Sensibles al Riesgo, considerando tanto el enfoque de costos con horizonte finito como el de recompensas con horizonte infinito. Se exploraron los avances existentes en la literatura y se propusieron ejemplos aplicados que ilustran los resultados teóricos.

El análisis realizado permitió mostrar cómo se sustenta el marco teórico de un MCM y de un PDM a tiempo discreto. Al incorporar la Sensibilidad al Riesgo se logró ampliar el alcance de la teoría y de los modelos a los que pueden aplicarse las herramientas estudiadas. Asimismo, al establecer y resolver el PCO se evidenció la complejidad inherente a la determinación de soluciones, además de proporcionar ejemplos concretos que refuerzan dicha teoría.

Las principales contribuciones de esta tesis se encuentran en la formulación y solución de ejemplos aplicados, particularmente el Problema de Reemplazamiento Óptimo de n Máquinas para el caso de horizonte finito con criterio sensible al riesgo, y el modelo presentado en la sección 3.2.3, pues ninguno de ellos había sido previamente explorado en la literatura.

El trabajo desarrollado permitió alcanzar un entendimiento más amplio de los PDMs Descontados Sensibles al Riesgo, abarcando desde los fundamentos teóricos hasta su aplicación en casos específicos. Además, se resaltó la importancia de estos modelos como una herramienta poderosa para la toma de decisiones en contextos complejos e inciertos.

Como líneas de investigación futura en el contexto de los PDMs Descontados Sensibles al Riesgo, se proponen las siguientes:

- Plantear soluciones del PCO en horizonte finito considerando funciones de costo no acotadas, motivados por el trabajo presentado en [11].
- Desarrollar la teoría necesaria para abordar un PCO con un factor de descuento

**PDMs Descontados Sensibles al Riesgo: Horizonte Infinito con Criterio
de Recompensas**

3.2 Caso Sensible al Riesgo

variable (dependiente, por ejemplo, del estado), siguiendo como referencia el caso neutral estudiado en [21].

Apéndice A

Kérneles Estocásticos e Integración

Definición A.0.1. Sean X y Y espacios de Borel. Un **Kérel Estocástico** sobre X dado Y es una función $P(\cdot|\cdot)$ tal que

- (a) $P(\cdot|y)$ es una medida de probabilidad sobre X para cada $y \in Y$ fijo, y
- (b) $P(B|\cdot)$ es una función medible sobre Y para cada $B \in \mathcal{B}(X)$ fijo.

Teorema A.0.2. (Teorema de Ionescu-Tulcea)

Sea X_0, X_1, \dots , una sucesión de espacios de Borel y, para $n = 0, 1, \dots$, definimos $Y_n := X_0 \times X_1 \times \dots \times X_n$ y $Y := \prod_{n=0}^{\infty} X_n$. Sea v una medida de probabilidad arbitraria sobre X_0 y, para cada $n = 0, 1, \dots$, $P_n(dx_{n+1}|y_n)$ es un kérel estocástico sobre X_{n+1} dado Y_n . Entonces existe una única medida de probabilidad P_v sobre Y tal que, para cada rectángulo medible $B_0 \times \dots \times B_n$ en Y_n ,

$$P_v(B_0 \times \dots \times B_n) = \int_{B_0} v(dx_0) \int_{B_1} P_0(dx_1|x_0) \dots \int_{B_n} P_{n-1}(dx_n|x_0, \dots, x_{n-1}).$$

Además, para cualquier función u medible y no negativa sobre Y , la función

$$x \mapsto \int u(y) P_x(dy)$$

es medible en X_0 , donde P_x representa a P_v cuando v es la probabilidad concentrada en $x \in X_0$.

Demostración. Ver [1] y [3]. ■

Apéndice B

Esperanza Condicional

Sean (Ω, \mathcal{F}, P) un espacio de probabilidad, \mathcal{F} una σ -álgebra contenida en \mathcal{F} , y ξ una variable aleatoria. Si ξ es P integrable, entonces definimos a la Esperanza Condicional de ξ dado \mathcal{F} , como la variable aleatoria denotada por $E[\xi|\mathcal{F}]$, y es tal que:

1. $E[\xi|\mathcal{F}]$ es \mathcal{F} -medible.
2. Para todo $A \in \mathcal{F}$, se tiene que $\int_A E[\xi|\mathcal{F}]dP = \int_A \xi dP$.

Tenemos algunas propiedades que son de utilidad:

Proposición B.0.1. Dadas cualesquiera dos variables aleatorias $X, Y \geq 0$ en (Ω, \mathcal{F}, P) , para cualquier sub σ -álgebra $\mathcal{F}_0 \subset \mathcal{F}$ y $\varepsilon \in (0, 1)$, tenemos lo siguiente:

- (1) $E[Y^\varepsilon|\mathcal{F}_0] \leq (E[Y|\mathcal{F}_0])^\varepsilon$ (desigualdad de Jensen);
- (2) $E[XY|\mathcal{F}_0] \leq (E[X^p|\mathcal{F}_0])^{\frac{1}{p}} (E[Y^q|\mathcal{F}_0])^{\frac{1}{q}}$ para $p, q > 1$ con $\frac{1}{p} + \frac{1}{q} = 1$ (desigualdad de Hölder);
- (3) Supóngase que $E[X^{p'}|\mathcal{F}_0] < \infty$ para algún $p' > 1$. Entonces, $E[X^p|\mathcal{F}_0] \leq E[X^{p'}|\mathcal{F}_0]$ para $0 < p \leq p'$, y $(E[X^p|\mathcal{F}_0])^{\frac{1}{p}} \uparrow$ conforme $p \uparrow$, lo cual implica que

$$\lim_{p \rightarrow 1^+} (E[X^p|\mathcal{F}_0])^{\frac{1}{p}} = E[X|\mathcal{F}_0].$$

Apéndice C

Algunos Resultados de Convergencia

Supóngase que (S, \mathcal{F}, μ) es un espacio de medida fijo, y que todas las funciones definidas en S son reales.

Teorema C.0.1. C1 Teorema de la Convergencia Monótona

Sea $\{f_n\}$ una sucesión de funciones medibles no negativas definidas en S , supon-
gamos que convergen a una función medible f , entonces

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu.$$

Nota: la sucesión $\{f_n\}$ puede converger a f casi donde quiera, y el teorema anterior sigue siendo válido.

Teorema C.0.2. C1 Teorema de la Convergencia Dominada

Sea $\{f_n\}$ una sucesión de funciones medibles y g una función medible y μ -
integrable definidas en S , tales que $|f_n| \leq g$, para toda n . Si $f_n \rightarrow f$ casi en todas
partes, entonces, f es μ -integrable y

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu.$$

Apéndice D

Notación de Landau y propiedades de $o(\cdot)$

La Notación de orden de Landau, también conocida como “ o minúscula” y “ O mayúscula”, es fundamental en el análisis asintótico de funciones. Estas notaciones permiten clasificar funciones según su comportamiento cuando la variable independiente tiende a un valor específico, como el cero o el infinito.

Definición D.0.1. Decimos que $f(x) = O(g(x))$ cuando $x \rightarrow 0$ si existe una constante M tal que $|f(x)| \leq M|g(x)|$ en algún entorno agujerado alrededor de 0, esto es, para $x \in (-\delta, \delta) \setminus \{0\}$ para algún valor de δ .

Decimos que $f(x) = o(g(x))$ cuando $x \rightarrow 0$ si $\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = 0$. Esto implica que existe una vecindad agujerada de 0 en la cual g no se desvanece.

El siguiente teorema resume algunas propiedades bien conocidas de los símbolos de orden de Landau:

Teorema D.0.2. 1. Si $f(x) = O(g(x))$, entonces $o(f(x)) + o(g(x)) = o(g(x))$,

2. Si $c \neq 0$, entonces $cO(g(x)) = O(g(x))$ y $co(g(x))$,

3. $o(f(x))O(g(x)) = o(f(x)g(x))$,

4. $x^a = O(x^b)$ para todo $b \leq a$ y $x^a = o(x^b)$ para todo $b < a$,

5. $O(x^a) + O(x^b) = O(x^{\min(a,b)})$, $o(x^a) + o(x^b) = o(x^{\min(a,b)})$, y

$$O(x^a) + o(x^b) = \begin{cases} o(x^b), & b < a, \\ O(x^a), & b \geq a. \end{cases}$$

Apéndice E

Códigos elaborados en Python

Se presentan los códigos de las implementaciones hechas en Python para la solución de los ejemplos numéricos mostrados en la tesis. Además, dichos códigos se encuentran en el siguiente repositorio de GitHub: https://github.com/EzequielHG3105/tesisMCM_EzequielHG.git.

Nota E.0.1. En los siguientes códigos no se agregaron acentos en los comentarios para no interferir en el funcionamiento de los mismos a la hora ejecutarlos.

E.1. Códigos para el Problema de Reemplazamiento Óptimo de n Máquinas

Código para la obtención del espacio de Estados y Acciones:

```
# Funcion para obtener el espacio de estados E
# Entrada: Niveles de deterioro D y numero de maquinas n
# Salida: matriz con estados E
def Estados(D,n):
    E = np.zeros([D**n,n])
    for i in range(n-1,-1,-1):
        cont = 0
        while cont <= D**n-1:
            for j in range(1,D+1):
                for k in range(D**(n-i-1)):
                    E[cont , i] = j
                    cont += 1
```

```
return E
```

```
# Funcion para obtener el espacio de acciones A
# Entrada: Niveles de deterioro D y numero de maquinas n
# Salida: matriz con acciones A
def Acciones(D,n):
    A = np.zeros([2**n,n])
    for i in range(n-1,-1,-1):
        cont = 0
        while cont <= 2**n-1:
            for j in range(2):
                for k in range(2**(n-i-1)):
                    A[cont,i] = j
                    cont += 1
    return A
```

Código para la obtención de la ley de transición Q :

```
# Funcion para obtener la ley de transicion Q(j,k,i),
# que corresponde a pasar del estado j al estado k
# dada la accion i
# Entrada: Niveles de deterioro D, numero de maquinas n,
#          matrices P0 de transiciones entre niveles de
#          deterioro de cada maquina y P1 de transiciones
#          cuando se decide reparar,
#          (ambas matrices son dadas por el usuario)
# Salida: Ley de transicion Q para cada par de estados j,k
#          y accion i
def mat_Q(D,n,P0,P1):
    Q = np.zeros([2**n,D**n,D**n])
    for i in range(2**n):
        for j in range(D**n):
            for k in range(D**n):
                Q[i,j,k] = 1
    for s in range(n):
        if A[i,s] == 0:
            Q[i,j,k] = P0[int(E[j,s])-1,int(E[k,s])-1]*Q[i,j,k]
        else:
```

```

    Q[i , j , k] = P1[int (E[j , s]) - 1 , int (E[k , s]) - 1]*Q[i , j , k]
return Q

```

Nota: Para poder ejecutar de manera correcta el código, realizar la indentación correspondiente.

Código para la obtención la función de valor óptimo \mathcal{J}_N^* y políticas óptimas en el caso neutral al riesgo:

```

# Obtencion la funcion V(i, j) que es la funcion
# de valor en la etapa i-1 del estado j-esimo (E(:, j))
# para el caso neutral al riesgo
# Entrada: Horizonte N, niveles de deterioro D,
#          numero de maquinas n, espacio de estados E,
#          espacio de acciones A, ley de transicion Q,
#          vector de costos c, costo fijo de
#          reparacion R, valores de lambda (labda)
#          y alpha
# Salida: Matrices V de la funcion de valor optimo
#          en cada estado y P de politicas optimas
def fun_Vn(N,D,n,E,A,Q,c,R,labda , alpha):
    V = np.zeros([N+1,D**n]) # valor optimo
    P = np.zeros([N,D**n]) # politicas
    # damos valores a la primera iteracion (V_N=0)
    for j in range(D**n): V[N, j] = 0
    # Ahora hacemos el calculo de la funcion de
    # valor optima y politica optima
    for i in range(N-1,-1,-1): # para cada etapa
        for j in range(D**n): # para cada estado
            vp = np.zeros(2**n) # auxiliar para la esperanza
            C = np.zeros(2**n) # auxiliar para el costo
            v = np.zeros(2**n) # auxiliar para V
            for s in range(2**n): # para cada accion
                vp[s] = 0
                # ahora calculamos el valor esperado de
                # v(y) dado el estado j
                # y la accion s
                for k in range(D**n):
                    vp[s] = vp[s]+V[i+1,k]*Q[s, j , k]

```

```

C[s] = 0
for k in range(n):
    if A[s,k] == 0:
        C[s] = c[int(E[j,k])-1]+C[s]
    else:
        C[s] = c[0]+R+C[s]
v[s] = C[s]+alpha*vp[s]
# ahora buscamos el minimo o accion optima
V[i,j] = v[0]
P[i,j] = 1
# busqueda del minimo y accion optima
for s in range(1,2**n):
    if v[s] < V[i,j]:
        V[i,j] = v[s]
        P[i,j] = s+1
        # numero de accion que es optima
        # para el estado j en la etapa i-1
return V,P

```

Código para la obtención la función de valor óptimo \mathcal{J}_N^* y políticas óptimas en el caso sensible al riesgo:

```

# Obtencion la funcion V(i,j) que es la funcion
# de valor en la etapa i-1 del estado j-esimo (E(:,j))
# para el caso sensible al riesgo
# Entrada: Horizonte N, niveles de deterioro D,
#          numero de maquinas n, espacio de estados E,
#          espacio de acciones A, ley de trancision Q,
#          vector de costos c, costo fijo de
#          reparacion R, valores de lambda (labda)
#          y alpha
# Salida: Matrices V de la funcion de valor optimo
#         en cada estado y P de politicas optimas
def fun_Vn(N,D,n,E,A,Q,c,R,labda,alpha):
    V = np.zeros([N+1,D**n]) # valor optimo
    P = np.zeros([N,D**n]) # politicas
    # damos valores a la primera iteracion (V_N=0)
    for j in range(D**n): V[N,j] = 1/labda

```

```

# Ahora hacemos el calculo de la funcion de
# valor optima y politica optima
for i in range(N-1,-1,-1): # para cada etapa
    for j in range(D**n): # para cada estado
        vp = np.zeros(2**n) # auxiliar para la esperanza
        C = np.zeros(2**n) # auxiliar para el costo
        v = np.zeros(2**n) # auxiliar para V
        for s in range(2**n): # para cada accion
            vp[s] = 0
            # ahora calculamos el valor esperado de
            # v(y) dado el estado j
            # y la accion s
            for k in range(D**n):
                vp[s] = vp[s]+V[i+1,k]*Q[s,j,k]
            C[s] = 0
            for k in range(n):
                if A[s,k] == 0:
                    C[s] = c[int(E[j,k])-1]+C[s]
                else:
                    C[s] = c[0]+R+C[s]
            v[s] = np.exp((alpha**i)*labda*C[s]) alpha*vp[s]
        # ahora buscamos el minimo o accion optima
        V[i,j] = v[0]
        P[i,j] = 1
        # busqueda del minimo y accion optima
        for s in range(1,2**n):
            if v[s] < V[i,j]:
                V[i,j] = v[s]
                P[i,j] = s+1
            # numero de accion que es optima
            # para el estado j en la etapa i-1
return V,P

```

E.2. Código para la aproximación del producto P

```
# funcion que calcula el producto hasta un horizonte n
# Entrada: valores de M, delta, alpha, n y
#           r (numero de decimales a considerar)
# Salida: aproximacion del producto
def prod_n(M,delta , alpha , labda , n , r ):
    mp.dps = r
    producto = mp.mpf(1.0)
    for j in range(n):
        S_j = labda*((alpha**(j+1)-alpha**(n))/(1-alpha))
        term = (M**(S_j+1)-1)/((S_j+1)*(M-1))
        producto *= term
    return producto

# Funcion para determinar n dado un error |epsilon=10^{-r}
# Entrada: valores de M, delta, alpha, n y
#           r (tolerancia para |epsilon)
# Salida: valor de n que permite aproximar mejor a P
def n_aprox(M,delta , alpha , labda , r ):
    mp.dps = r
    n = 10
    producto_1 = prod_n(M,delta , alpha , labda , n , r)
    producto_2 = prod_n(M,delta , alpha , labda , n+10,r)
    while producto_2 - producto_1 != 0:
        producto_1 = prod_n(M,delta , alpha , labda , n , r)
        producto_2 = prod_n(M,delta , alpha , labda , n+10,r)
        n += 10
    return n
```

Bibliografía

- [1] R. B. Ash. *Real analysis and probability: probability and mathematical statistics: a series of monographs and textbooks*. Academic Press, 2014.
- [2] D. Bertsekas. Programming and optimal control. *Athena Scientific, Massachusetts*, 1987.
- [3] D. Bertsekas and S. E. Shreve. *Stochastic optimal control: the discrete-time case*. Athena Scientific, 1996.
- [4] R. Blancas-Rivera, R. Cavazos-Cadena, and H. Cruz-Suárez. Discounted approximations in risk-sensitive average markov cost chains with finite state space. *Mathematical Methods of Operations Research*, 91(2):241–268, 2020.
- [5] D. A. Braun, A. J. Nagengast, and D. M. Wolpert. Risk-sensitivity in sensorimotor control. *Frontiers in human neuroscience*, 5:1, 2011.
- [6] I. N. Bronshtein, K. A. Semendyayev, G. Musiol, and H. Mühlig. Handbook of mathematics. 2007. *Springer-Verlag Berlin Heidelberg*, 10:978–3, 2007.
- [7] N. Bäuerle and A. Jaskiewicz. Markov decision processes with risk-sensitive criteria: an overview. *Mathematical Methods of Operations Research*, 99(1):141–178, 2024.
- [8] N. Bäuerle and U. Rieder. More risk-sensitive markov decision processes. *Mathematics of Operations Research*, 39(1):105–120, 2014.
- [9] S. P. Coraluppi and S. I. Marcus. Risk-sensitive and minimax control of discrete-time, finite-state markov decision processes. *Automatica*, 35(2):301–309, 1999.
- [10] H. D. Cruz-Suárez. A stochastic consumption-investment problem with unbounded utility function. *Morfismos*, 4(1):19–30, 2000.

- [11] X. Guo. Risk-sensitive discounted markov decision processes with unbounded reward functions and borel spaces. *Stochastics*, 96(1):649–666, 2024.
- [12] O. Hernández-Lerma and J. B. Lasserre. *Further topics on discrete-time Markov control processes*, volume 42. Springer Science & Business Media, 2012.
- [13] E. Hernández-García and R. Blancas-Rivera. Simulación de una cadena de markov sensible al riesgo. In *Enseñanza de Probabilidad y Estadística*, pages 133–146. Benemérita Universidad Autónoma de Puebla, Puebla, México, 2023.
- [14] E. Hernández-García, H. Cruz-Suárez, and R. Blancas-Rivera. Análisis del problema de reemplazamiento Óptimo de n máquinas bajo un criterio sensible al riesgo. In *IA y las Matemáticas y Ciencias Naturales en la Vida, el Conocimiento y la Cultura*, pages 6–11. Academia Journals, Puebla, México, 2025.
- [15] O. Hernández-Lerma and J. B. Lasserre. *Discrete-time Markov control processes: basic optimality criteria*. Springer Science & Business Media, 2012.
- [16] R. A. Howard and J. E. Matheson. Risk-sensitive markov decision processes. *Management Science*, 18(7):356–369, 1972.
- [17] R. Ilhuicatzí-Roldán and H. Cruz-Suarez. Optimal replacement in a system of n-machines with random horizon. *Proyecciones (Antofagasta)*, 31(3):219–233, 2012.
- [18] J. Von Neumann and O. Morgenstern. *Theory of games and economic behavior*, 2nd rev. 1947.
- [19] A. S. Nowak. Measurable selection theorems for minimax stochastic optimization problems. *SIAM Journal on Control and Optimization*, 23(3):466–476, 1985.
- [20] M. L. Puterman. Markov decision processes. *Handbooks in operations research and management science*, 2:331–434, 1990.
- [21] Q. Wei and X. Guo. Markov decision processes with state-dependent discount factors and unbounded rewards/costs. *Operations Research Letters*, 39(5):369–374, 2011.