iThenticate Página 4 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.



Universidad Juárez Autónoma de Tabasco

División Académica de Ciencias Básicas

PROCESOS DE DECISIÓN DE MARKOV CON APLICACIONES A FINANZAS

Tesis

Para obtener el título de:

MAESTRO EN CIENCIAS EN MATEMÁTICAS APLICADAS

Presenta:

RUBÉN ZAVALA CAMARGO

Director de tesis:

DR. HELIODORO DANIEL CRUZ SUÁREZ

Cunduacán, Tabasco , 13 de junio de 2016.

iThenticate Página 4 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 5 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

13 de junio de 2016

Lic. Rubén Zavala Camargo
Pasante de la Maestría en Ciencias
en Matemáticas Aplicadas
P r e s e n t e.

Por medio del presente y de la manera más cordial, me dirijo a Usted para hacer de su conocimiento que proceda a la impresión del trabajo titulado "Procesos de Decisión de Markov con Aplicaciones a Finanzas", en virtud de que reúne los requisitos para el EXAMEN PROFESIONAL para obtener el grado de Maestro en Ciencias en Matemáticas Aplicadas.

Sin otro particular, reciba un cordial saludo.

A t e n t a m e n t e.

Dr. Gerardo Delgadillo Piñón
Director

DIVISIÓN ACADÉMICA DE CIENCIAS BÁSICAS
DIRECCIÓN.

C.c.p.- Archivo
Dr'GDP/Dr'JLSC/emt

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 5 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

CARTA DE AUTORIZACIÓN

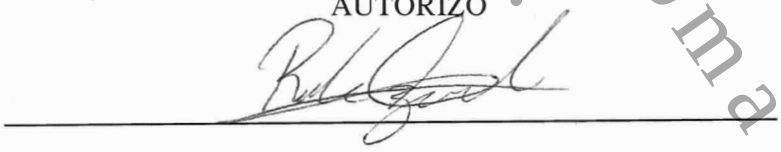
El que suscribe, autoriza por medio del presente escrito a la Universidad Juárez Autónoma de Tabasco para que utilice tanto física como digitalmente la tesis de grado denominada "PROCESOS DE DECISIÓN DE MARKOV CON APLICACIONES A FINANZAS", de la cual soy autor y titular de los Derechos de Autor.

La finalidad del uso por parte de la Universidad Juárez Autónoma de Tabasco de la tesis antes mencionada, será única y exclusivamente para difusión, educación y sin fines de lucro; autorización que se hace de manera enunciativa más no limitativa para subirla a la Red Abierta de Bibliotecas Digitales (RABID) y a cualquier otra red académica con las que la Universidad tenga relación institucional.

Por lo antes manifestado, libero a la Universidad Juárez Autónoma de Tabasco de cualquier reclamación legal que pudiera ejercer respecto del uso y manipulación de la tesis mencionada y para los fines estipulados en éste documento.

Se afirma la presente autorización en la ciudad de Villahermosa, Tabasco a los 13 días del mes de junio del año 2016.

AUTORIZO



Rubén Zavala Camargo
Matricula: 122A15002

iThenticate Página 7 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Agradecimientos

Agradezco a Dios por todo su amor y por estar siempre conmigo, por darme la fuerza y sabiduría necesaria para terminar con éxito este trabajo.

Quiero agradecer enormemente a mis padres y mis hermanos por todo su apoyo a lo largo de mi vida.

A mi esposa Mónica y mis hijos Nathanael, Vanessa y Caleb, por todo su amor, toda su confianza y su paciencia. ¡Los amo!

Al Dr. Heliodoro Daniel Cruz Suárez, por su excelente dirección de este trabajo, dedicación y amabilidad. ¡Muchas gracias!

Agradezco al Dr. Aroldo Pérez, Dr. Justino Alavez y Dr. Gamaliel Blé, quienes son parte de mi formación. También, agradezco a la Dra. Addy Bolívar, Dr. Aroldo Pérez, Dr. Lucas López y al Dr. Edilberto Nájera, por su tiempo dedicado para leer este trabajo y por sus excelentes correcciones que ayudaron a mejorar la versión final de esta tesis.

A mis amigos y compañeros, Domingo, Jonathan, Lino y Luis, por su apoyo dentro y fuera de clases.

A mi amigo y compañero de trabajo, al M.C. Miguel Ángel Sánchez Ovando por su apoyo y sus comentarios en este trabajo.

Agradezco al Ing. Carlos A. Hernández Rodríguez, director general de SIA Software, por su apoyo y confianza.

A la Universidad Juárez Autónoma de Tabasco (UJAT), en particular a la División Académica de Ciencias Básicas (DACB), por proporcionarme una excelente preparación académica.

Por último, agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) la beca que me fue otorgada durante los estudios de maestría, la cual fue apoyo indispensable para destinar tiempo a los estudios que este proyecto exige.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 7 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

Índice general

1. Teoría de Procesos de Decisión de Markov de Horizonte Finito **1**
1.1. Modelos de Decisión de Markov 1
1.2. Modelos de Decisión de Markov de Horizonte Finito 4
1.3. La Ecuación de Bellman 6
1.4. Modelos de Decisión de Markov Estructurados 12
1.4.1. Modelos de Decisión de Markov Semicontinuos 14
1.4.2. Modelos de Decisión de Markov Continuos 17
1.4.3. Modelos de Decisión de Markov Medibles 18
1.5. Modelos de Decisión de Markov Estacionarios 20

2. Teoría de Procesos de Decisión de Markov de Horizonte Infinito **25**
2.1. Modelos de Decisión de Markov de Horizonte Infinito 25
2.2. Modelos de Decisión de Markov Semicontinuos 33
2.3. Modelos de Decisión de Markov Restringidos 36

3. Los Mercados Financieros **41**
3.1. Dinámica de Activos y Estrategias de Portafolios 41
3.2. Funciones de Utilidad y Utilidad Esperada 47

4. Aplicaciones a Finanzas **49**
4.1. Problemas de Optimización Financiera 49
4.1.1. El Problema de Optimización de un Periodo 49
4.1.2. El Problema de Consumo-Inversión 52
4.2. Problemas de Paro con Horizonte Finito 56
4.2.1. Precio de Opciones Americanas 59

Conclusiones **66**

A. Herramientas de Matemática Financiera **69**
A.1. Teoría de Precios bajo No Arbitraje 69

B. Códigos de Programación **71**

Universidad Autónoma de Nuevo México

VI

Introducción

Esta tesis trata con Procesos de Decisión de Markov (PDMs) con aplicaciones en finanzas tales como problemas de optimización de un periodo, riqueza terminal y el precio de una opción Americana.

Los PDMs, dependen de una sucesión de decisiones (llamada política) aplicadas en cada etapa de tiempo. Para evaluar la calidad de las políticas se cuenta con un criterio de rendimiento, definido en términos de una recompensa (o costo) por etapa, aquí el criterio que se considera es la recompensa total esperada descontada. Entonces el problema básico de los PDMs (llamado problema de control óptimo), consiste en optimizar el criterio de rendimiento sobre el conjunto de las políticas. A la política que optimiza el criterio de rendimiento se le llama política óptima (ver [3], [17], [18] y [22]).

Los modelos en matemáticas financieras con frecuencia se definen en tiempo continuo. Así, los problemas de optimización conducen a problemas de control estocástico en tiempo continuo. Sin embargo, sólo unos cuantos de estos problemas se pueden resolver explícitamente. Así, se requiere del uso de métodos para modelarlo como un proceso a tiempo discreto. El problema resultante es un Problema de Decisión de Markov (ver [1], [9], [10] y [11]).

Las finanzas son las actividades relacionadas con los flujos de capital y dinero entre los individuos, empresas o estados. Por extensión, también se denomina finanzas al análisis de estas actividades como especialidad de la economía y la administración, es decir, una rama de la economía que se dedica de forma general al estudio del dinero y particularmente está relacionado con las transacciones y la administración del dinero en cuyo marco se estudia la obtención y gestión, por parte de una compañía, un individuo o del propio estado, de los fondos que necesita para cumplir sus objetivos, y de los criterios con que se dispone de sus activos (ver [1], [12], [13] y [14]).

Las finanzas tratan, por lo tanto, de las condiciones y las oportunidades en que se consigue el capital, de los usos de éste, y de los pagos e intereses que se encargan de las transacciones en dinero. Por lo tanto, el optimizar tales pagos e intereses justifica un estudio con herramientas matemáticas tales como los Procesos de Decisión de Markov.

Los PDMs modelan ciertos fenómenos aleatorios que evolucionan con el tiempo en los cuales se puede influir en su comportamiento. Se dará brevemente la descripción de tales procesos; dado un sistema observado en cada época del tiempo $n = 0, 1, \ldots$, el estado x_n del sistema proporciona al decisor toda la información necesaria para elegir una acción a_n en un conjunto de acciones $D_n(x)$ disponibles en cada estado. Como un resultado de la elección de la acción a_n, ocurren dos cosas: 1) el decisor incurre en un costo o recompensa y 2) el sistema evoluciona de acuerdo a una ley de probabilidad a un nuevo estado x_{n+1}, un nuevo control es elegido y el proceso se repite, en otras palabras, en cualquier tiempo

iThenticate Página 10 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

dado, el costo y la ley de transición dependen sólo del estado actual del sistema y la acción presente (ver [1], [3], [18], [19] y [22]). En conclusión, un modelo de decisión de Markov se define como:

$$(E, A, \{D(x) : x \in E\}, Q, r)$$

con el siguiente significado:

- E denota el *espacio de estados* del sistema. Un estado $x \in E$ es la información que está disponible para el controlador. Dada esta información, una acción tiene que ser elegida.
- A denota el *espacio de acciones*. Dado un estado específico $x \in E$, cierta subclase de acciones $D(x) \subset A$ sólo pueden ser admisibles.
- $\{D(x) : x \in E\}$ es una familia de subconjuntos no vacíos $D(x)$ de A.
- $Q(B|x,a)$ es un kernel de transición estocástico el cual da la probabilidad de que el siguiente estado esté en el conjunto B si el estado actual es x y se tomó la acción a.
- $r(x,a)$ da la *recompensa en una etapa* del sistema, si el estado actual es x y se tomó la acción a.

Los Procesos de Decisión de Markov se distinguen, entre otras cosas, por el horizonte de tiempo finito y el horizonte de tiempo infinito.

En este trabajo se estudiarán problemas de finanzas que pueden ser modelados por PDMs.

El matemático Richard Bellman propuso la programación dinámica en 1953 que se utiliza para optimizar problemas complejos que pueden ser discretizados. Actualmente existen métodos de solución para problemas de Control Estocástico por medio de programación dinámica.

El control estocástico se enriqueció por el trabajo de Robert C. Merton de 1973, "Theory of Rational Option Pricing", en el que se hacía referencia a un método matemático desarrollado por Fisher Black y Myron Scholes. A este modelo lo denominó Black-Scholes y fue empleado para estudiar el valor actual de una opción europea para la compra (Call), o venta (Put), de acciones en una fecha futura. Posteriormente el modelo se amplió para opciones sobre acciones que producen dividendos, y luego se adoptó para opciones europeas, americanas, y mercado monetario. Este trabajo seminal originó un importante desarrollo matemático en finanzas (ver [25], [26], [27] y [28]).

Los PDMs han resuelto satisfactoriamente problemas de finanzas, el interés de este trabajo es hacer una revisión de esta aplicación: Procesos de Decisión de Markov con Aplicaciones a Finanzas.

Este trabajo de tesis se encuentra organizado de la siguiente manera:

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 10 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

En el Capítulo 1 se establece la teoría de Procesos de Decisión de Markov con horizonte de tiempo finito. Se proporciona el modelo matemático, una suposición de integrabilidad suficiente y la definición de políticas. Bajo suposiciones de estructura en el modelo, se mostrará que el Problema de Decisión de Markov se puede resolver recursivamente mediante la *Ecuación de Bellman*. Es investigado el caso de Procesos de Decisión de Markov estacionarios. La noción de "estacionario" indica que los datos del modelo no dependen del indice del tiempo.

En el Capítulo 2 se establece la teoría de Procesos de Decisión de Markov con horizonte de tiempo infinito donde se asume que la parte positiva de las recompensas son acotadas y convergen. Se muestra que el problema de horizonte infinito puede ser visto como una aproximación de los problemas de horizonte de tiempo finito si se satisface una suposición de estructura. También se da una caracterización de la función de valor de horizonte infinito como una solución más grande de la *ecuación de Bellman*. Se proporcionan algunas condiciones de compactación y continuidad, bajo las cuales la suposición de estructura para el problema de horizonte a tiempo infinito se satisface.

En el Capítulo 3 se introduce el mercado financiero el cual aparece en una aplicación en el Capítulo 4. Se presenta un mercado a tiempo discreto. Sin embargo, en general, no se restringe a espacios de probabilidad finita. Se define una estrategia de portafolio y se caracteriza la ausencia de arbitraje en este mercado.

En el capítulo 4, se da solución a un problema de optimización de un periodo, utilizando la teoría del Capítulo 1 y la solución de un problema de consumo e inversión considerando un horizonte de tiempo infinito. Se introduce el concepto de tiempo de paro para dar solución a un problema de opción de compra Americana.

Universidad Juárez Autónoma de Tabasco. México

Capítulo 1

Teoría de Procesos de Decisión de Markov de Horizonte Finito

En este capítulo se establece la teoría de Procesos de Decisión de Markov (PDMs) con un horizonte de tiempo finito, con espacios de estados y acciones generales (ver [1], [8] y [17]).

1.1. Modelos de Decisión de Markov

A continuación se enuncia la definición de un modelo de decisión de Markov.

Definición 1.1.1. Un *Modelo de Decisión de Markov* no estacionario con horizonte de planeación $N \in \mathbb{N}$ consiste de una sextupla $(E, A, D_n, Q_n, r_n, g_N)$ con el siguiente significado para $n = 0, 1, 2, \cdots, N-1$:

- E es el *espacio de estados*, con una σ-álgebra $\mathfrak{E}$. Los elementos (estados) se denotarán por $x \in E$.
- A es el *espacio de acciones*, con una σ-álgebra $\mathfrak{A}$. Los elementos (acciones) se denotarán por $a \in A$.
- $D_n \subset E \times A$ es un subconjunto medible de $E \times A$ y denota el conjunto de combinaciones posibles de estado-acción en el tiempo n. Suponer que D_n contiene la gráfica de una función medible $f_n : E \to A$, es decir, $(x, f_n(x)) \in D_n$ para todo $x \in E$. Para $x \in E$, el conjunto

$$D_n(x) = \{a \in A | (x, a) \in D_n\}$$

iThenticate Página 13 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

es el conjunto de *acciones admisibles* (o *permitidas*) en el estado x en el tiempo n.

- Q_n es un kernel de transición estocástico de D_n a E, es decir, para cualquier par fijo $(x,a) \in D_n$, la función $B \mapsto Q_n(B|x,a)$ es una medida de probabilidad sobre $\mathfrak{E}$ y $(x,a) \mapsto Q_n(B|x,a)$ es una función medible para todo $B \in \mathfrak{E}$. Q_n describe la *ley de transición*, la cantidad $Q_n(B|x,a)$ proporciona la probabilidad de que el siguiente estado, en el tiempo $n+1$, esté en B si el estado actual es x y se tomó la acción a en el tiempo n.

- $r_n : D_n \to \mathbb{R}$ es una función medible. $r_n(x,a)$ proporciona la *recompensa de una etapa* del sistema en el tiempo n, si el estado actual es x y se tomó la acción a.

- $g_N : E \to \mathbb{R}$ es una función medible. $g_N(x)$ proporciona una *recompensa terminal* del sistema en el tiempo N si el estado es x.

Observación 1.1.1. *a)* En muchas aplicaciones los espacios de estados y acciones son subconjuntos de Borel de espacios Polacos (es decir, espacios métricos, separables y completos) o conjuntos finitos o numerables. Entonces, las σ-álgebras $\mathfrak{E}$ y $\mathfrak{A}$ están dadas por las σ-álgebras $\mathscr{B}(E)$ y $\mathscr{B}(A)$ de todos los subconjuntos de Borel de E y A, respectivamente. En algunas aplicaciones E y A son subconjuntos de $\mathbb{R}^d$ ó $\mathbb{R}^d_+$.

b) Si la función r'_n de recompensa de una etapa también depende del siguiente estado, es decir, $r'_n = r'_n(x,a,x')$, entonces definir

$$r_n(x,a) := \int r'_n(x,a,x') Q_n(dx'|x,a).$$

c) Muchas veces D_n y Q_n son independientes de n y $r_n(x,a) := \beta^n r(x,a)$ y $g_N(x) = \beta^N g(x)$, para un factor (de descuento) $\beta \in (0,1]$. En este caso el modelo de decisión de Markov se llamará *estacionario* (ver la Sección 1.5).

$\Diamond$

La ley de transición estocástica de un modelo de decisión de Markov está muchas veces dada por una función del sistema. Para precisar, suponer que $Z_1, Z_2, \ldots, Z_N$ son variables aleatorias con valores en un espacio medible $(\mathscr{Z}, \mathfrak{Z})$. Estas variables aleatorias se llaman *perturbaciones*. Z_{n+1} influye en la transición del estado del sistema en el tiempo n al estado en el tiempo $n+1$. La distribución Q_n^Z de Z_{n+1} puede depender del estado y la acción actual en el tiempo n tal que $Q_n^Z(\cdot|x,a)$ es un kernel estocástico para $(x,a) \in D_n$. El nuevo estado del sistema, en el tiempo $n+1$, puede describirse por una *función del sistema* $T_n : D_n \times \mathscr{Z} \to E$ tal que

$$x_{n+1} = T_n(x_n, a_n, z_{n+1}).$$

Así, la ley de transición del Modelo de Decisión de Markov está determinada por T_n y Q_n^Z.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 13 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 14 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Teorema 1.1.1. Un modelo de decisión de Markov *se define de manera equivalente por* $(E, A, D_n, \mathscr{Z}, T_n, Q_n^Z, r_n, g_N)$ con el siguiente significado:

- (E, A, D_n, r_n, g_N) como en la definición 1.1.1.
- $\mathscr{Z}$ es el *espacio de perturbaciones*, con la σ-álgebra $\mathfrak{Z}$.
- $Q_n^Z(\cdot|x,a)$ es un *kernel de transición estocástico* para $B \in \mathfrak{Z}$ y $(x,a) \in D_n$ y $Q_n^Z(\cdot|x,a)$ denota la probabilidad que Z_{n+1} esté en B si el estado actual es x y se tomó la acción a.
- $T_n : D_n \times \mathscr{Z} \to E$ es una *función medible* y se llamará *función del sistema* (o *transición*). $T_n(x,a,z)$ proporciona el estado siguiente del sistema en el tiempo $n+1$ si, en el tiempo n, el sistema está en el estado x, se toma la acción a y ocurre la perturbación z en el tiempo $n+1$.

Demostración. Primero suponer que un modelo de decisión de Markov está dado como en la Definición 1.1.1. Definir $\mathscr{Z} := E$, $T_n(x,a,z) = z$ y $Q_n^Z(B|x,a) := Q_n(x,a)$ para $B \in \mathfrak{E}$. Inversamente, si están dados los datos $(E, A, D_n, \mathscr{Z}, T_n, Q_n^Z, r_n, g_N)$, entonces definir

$$Q_n(B|x,a) = Q_n^Z\Big(\{z \in \mathscr{Z} | T_n(x,a,z) \in B\}|x,a\Big), \quad B \in \mathfrak{E},$$

y se obtiene el kernel estocástico del modelo de decisión de Markov.

□

En lo siguiente suponer que existe un horizonte de planeación fijo $N \in \mathbb{N}$, es decir, N denota el número de etapas.

Definición 1.1.2. *a)* Una función medible $f_n : E \to A$, con la propiedad $f_n(x) \in D_n(x)$ para todo $x \in E$, se llamará una *regla de decisión* en el tiempo n. Denotar por F_n al conjunto de todas las reglas de decisión en el tiempo n.

b) Una sucesión de reglas de decisión $\pi = (f_0, f_1, ..., f_{N-1})$, con $f_n \in F_n$, se llamará una *política de N etapas* o *estrategia de N etapas*.

Notar que $F_n \neq \emptyset$ ya que, por la definición 1.1.1, D_n contiene la gráfica de una función medible $f_n : E \to A$.

Observación 1.1.2. Es razonable, en algunos casos, suponer *políticas aleatorizadas* o reglas de decisión aleatorizadas. Una *política de Markov aleatorizada* $\pi = (f_0, f_1, ..., f_{N-1})$ se da si $f_n(B|x)$ es un kernel estocástico con $f_n(D_n(x)|x) = 1$, para toda $x \in E$.

◊

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 14 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México

Un modelo de decisión de Markov es un experimento aleatorio de N etapas. Así, para establecer matemáticamente este experimento se debe definir el espacio de probabilidad. La *construcción canónica* es como sigue. Definir un espacio medible $(\Omega, \mathscr{F})$ por

$$\Omega = E^{N+1}, \quad \mathscr{F} = \mathfrak{E} \otimes \cdots \otimes \mathfrak{E}.$$

Denotar por $\omega = (x_0, x_1, \ldots, x_N) \in \Omega$. Las variables aleatorias $X_0, X_1, \ldots, X_N$ se definen sobre el espacio medible $(\Omega, \mathscr{F})$ por

$$X_n(\omega) = X_n((x_0, x_1, \ldots, x_N)) = x_n,$$

siendo la n-ésima proyección de ω. La variable aleatoria X_n representa el estado del sistema en el tiempo n y (X_n) se llamará *Proceso de Decisión de Markov*. Suponer ahora que $\pi = (f_0, f_1, ..., f_{N-1})$ es una política fija y que $x \in E$ es un estado inicial fijo. Por el teorema de Ionescu-Tulcea existe una medida de probabilidad única $\mathbb{P}_x^\pi$ sobre $(\Omega, \mathscr{F})$ con

(i) $\mathbb{P}_x^\pi(X_0 \in B) = \delta_x(B)$, para todo $B \in \mathfrak{E}$.

(ii) $\mathbb{P}_x^\pi(X_{n+1} \in B | X_1, \ldots, X_n) = \mathbb{P}_x^\pi(X_{n+1} \in B | X_n) = Q_n\big(B | X_n, f_n(X_n)\big)$.

La ecuación (ii) es la llamada *Propiedad de Markov*, es decir, la sucesión de las variables aleatorias $X_0, X_1, \ldots, X_n$ es un proceso de Markov no estacionario con respecto a $\mathbb{P}_x^\pi$. Por $\mathbb{E}_x^\pi$ denotar la esperanza con respecto a $\mathbb{P}_x^\pi$. Además, denotar por $\mathbb{P}_{nx}^\pi$ la probabilidad condicional

$$\mathbb{P}_{nx}^\pi(\cdot) := \mathbb{P}^\pi(\cdot | X_n = x).$$

$\mathbb{E}_{nx}^\pi$ es el correspondiente operador esperanza.

10

1.2. Modelos de Decisión de Markov de Horizonte Finito

Para garantizar que todas las esperanzas que aparecen están bien definidas se hace la siguiente suposición:

Suposición de Integrabilidad (SI_N)**:** Para $n = 0, 1, \cdots, N$

$$\delta_n^N(x) := \sup_\pi \mathbb{E}_{nx}^\pi \left[\sum_{k=n}^{N-1} r_k^+(X_k, f_k(X_k)) + g_N^+(X_N) \right] < \infty, \quad x \in E,$$

donde $x^+ = \max\{0, x\}$ denota la parte positiva de x.

iThenticate Página 16 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Ahora se introduce la recompensa esperada descontada de una política y el problema de optimización de N etapas. Para $n = 0, 1, \ldots, N$ y una política $\pi = (f_0, \ldots, f_{N-1})$ sea $V_{n\pi}(x)$ definida por

$$V_{n\pi}(x) := \mathbb{E}^{\pi}_{nx}\left[\sum_{k=n}^{N-1} r_k(X_k, f_k(X_k)) + g_N(X_N)\right], \quad x \in E.$$

$V_{n\pi}(x)$ es la *recompensa total esperada en el tiempo n sobre las etapas restantes de n a N*, si se usa la política π y se inicia en el estado $x \in E$ en el tiempo n. La *función de valor* V_n se define por

$$V_n(x) := \sup_{\pi} V_{n\pi}(x), \quad x \in E.$$

$V_n(x)$ es la *recompensa total máxima esperada en el tiempo n sobre las etapas restantes de n a N*, si se inicia en el estado $x \in E$ en el tiempo n. Las funciones $V_{n\pi}$ y V_n están bien definidas ya que

$$V_{n\pi}(x) \leq V_n(x) \leq \delta_n^N(x) < \infty, \quad x \in E.$$

Notar que $V_{N\pi}(x) = V_N(x) = g_N(x)$ y que $V_{n\pi}$ sólo depende de $(f_n, \ldots, f_{N-1})$. Además, no es cierto en general que $V_{n\pi}$ es medible. Esto provoca inconveniencias teóricas. Algunas suposiciones más son necesarias para implicar esto (ver la Sección 1.4). Una política $\pi \in F_0 \times \cdots \times F_{N-1}$ se llamará *óptima* para el modelo de decisión de Markov de N etapas si $V_{0\pi}(x) = V_0(x)$, para todo $x \in E$.

La razón por la que se utilizan *políticas* de Markov es la siguiente: Denotar al *conjunto de historias* por

$$H_0 = E$$
$$H_n = H_{n-1} \times A \times E.$$

Un elemento $h_n = (x_0, a_0, x_1, \ldots, x_n) \in H_n$ se llamará *historia hasta el tiempo n*.

Definición 1.2.1. *a)* Una función medible $f_n : H_n \to A$, con la propiedad $f_n(h_n) \in D_n(x_n)$ para toda $h_n \in H_n$, se llamará una *regla de decisión dependiente de la historia* en la etapa n.

b) Un sucesión $\pi = (f_0, f_1, \ldots, f_{N-1})$, donde f_n es una regla de decisión dependiente de la historia en la etapa n, se llamará una *política dependiente de la historia de la etapa N*. Se denota por Π_N al conjunto de todas las políticas dependientes de la historia de la etapa N.

Sea $\pi \in \Pi_N$ una política dependiente de la historia. Entonces $V_{n\pi}(h_n)$ está definida como la esperanza condicional de la recompensa total en $[n, N]$, dada la historia $h_n \in H_n$. El siguiente teorema expresa que las políticas dependientes de la historia no mejoran la máxima recompensa esperada. Para una prueba ver [33].

Universidad Jaárez Autónoma de Tabasco. México.

iThenticate Página 16 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 17 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Teorema 1.2.1. Para $n = 0, \ldots, N$ se cumple que:

$$V_n(x_n) = \sup_{\pi \in \Pi_N} V_{n\pi}(h_n), \quad h_n = (x_0, a_0, x_1, \ldots, x_n).$$

En la manera de calcular $V_0(x)$ también se tiene que determinar la función de valor $V_n(x)$. Esto es una característica estándar de muchas técnicas de optimización multietapas que se explicará con detalle en la siguiente sección.

1.3. La Ecuación de Bellman

Para una política fija $\pi \in F_0 \times \cdots \times F_{N-1}$ es posible calcular la recompensa esperada descontada recursivamente por *iteración de recompensas*. Se denotará por

$$\mathbb{M}(E) := \{v : E \to [-\infty, \infty) | v \text{ es medible}\}.$$

Por las suposiciones se tiene que $V_{n\pi} \in \mathbb{M}(E)$ para todo π y n.

Definición 1.3.1. Se definen los siguientes operadores para $n = 0, 1, \ldots, N-1$.

a) Para $v \in \mathbb{M}(E)$ definir

$$(L_n v)(x,a) := r_n(x,a) + \int v(x') Q_n(dx'|x,a), \quad (x,a) \in D_n$$

cuando la integral existe.

b) Para $v \in \mathbb{M}(E)$ y $f \in F_n$, definir

$$(\mathscr{T}_{nf} v)(x) := (L_n v)(x, f(x)), \quad x \in E.$$

c) Para $v \in \mathbb{M}(E)$ definir

$$(\mathscr{T}_n v)(x) := \sup_{a \in D_n(x)} (L_n v)(x,a), \quad x \in E.$$

$\mathscr{T}_n$ se llamará el operador de recompensa máxima en el tiempo n.

Observación 1.3.1. *a)* Se tiene la siguiente relación entre operadores

$$\mathscr{T}_n v = \sup_{f \in F_n} \mathscr{T}_{nf} v.$$

b) Se cumple que $\mathscr{T}_{nf} v \in \mathbb{M}(E)$ para todo $v \in \mathbb{M}(E)$, pero $\mathscr{T}_n v$ no pertenece a $\mathbb{M}(E)$ en general.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 17 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 18 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

c) Si un modelo de decisión de Markov con perturbaciones (Z_n) está dado como en el Teorema 1.1.1, entonces $L_n v$ puede describirse como

$$(L_n v)(x,a) = r_n(x,a) + \mathbb{E}\left[v\big(T_n(x,a,Z_{n+1})\big)\right].$$

Notación. En lo siguiente se omitirán los paréntesis alrededor de los operadores $(\mathscr{T}_n v)(x)$ y simplemente se escribirá $\mathscr{T}_n v(x)$. Cuando se tenga una sucesión de operadores como $\mathscr{T}_n \mathscr{T}_{n+1} v$, entonces el orden de aplicación está dado por $(\mathscr{T}_n(\mathscr{T}_{n+1} v))$.

Los operadores tiene las siguientes propiedades importantes.

Lema 1.3.1. Los tres operadores son monótonos, es decir, para $v, w \in \mathbb{M}(E)$ con $v(x) \leqslant w(x)$, para todo $x \in E$, se cumple que:

a) $L_n v(x,a) \leqslant L_n w(x,a)$ para todo $(x,a) \in D_n$,

b) $\mathscr{T}_{nf} v(x) \leqslant \mathscr{T}_{nf} w(x)$ para todo $x \in E, f \in F_n$,

c) $\mathscr{T}_n v(x) \leqslant \mathscr{T}_n w(x)$ para todo $x \in E$.

Demostración. Sean $v, w \in \mathbb{M}(E)$ tales que $v(x) \leqslant w(x)$, para todo $x \in E$. Entonces

$$\int v(x') Q_n(dx'|x,a) \leqslant \int w(x') Q_n(dx'|x,a).$$

Así, $L_n v(x,a) \leqslant L_n w(x,a)$ lo que implica *a)* y *b)*. Tomando el supremo sobre todo $a \in D_n(x)$ se tiene *c)*.

□

Ahora, los operadores $\mathscr{T}_{nf}$ pueden usarse para calcular el valor de una política recursivamente.

Teorema 1.3.1. *(Iteración de recompensas).* Sea $\pi = (f_0, \ldots, f_{N-1})$ una política de N-etapas. Para $n = 0, 1, \ldots, N-1$ se cumple que:

a) $V_{N\pi} = g_N$ y $V_{n\pi} = \mathscr{T}_{nf_n} V_{n+1,\pi}$,

b) $V_{n\pi} = \mathscr{T}_{nf_n} \cdots \mathscr{T}_{N-1 f_{N-1}} g_N$.

Demostración. a) Para $x \in E$ se tiene

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 18 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 19 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

$$
\begin{aligned}
V_{n\pi}(x) &= \mathbb{E}^{\pi}_{nx}\left[\sum_{k=n}^{N-1} r_k\big(X_k, f_k(X_k)\big) + g_N(X_N)\right] \\
&= \mathbb{E}^{\pi}_{nx}[r_n\big(x, f_n(x)\big)] + \mathbb{E}^{\pi}_{nx}\left[\sum_{k=n+1}^{N-1} r_k\big(X_k, f_k(X_k)\big) + g_N(X_N)\right] \\
&= r_n\big(x, f_n(x)\big) \\
&\quad + \mathbb{E}^{\pi}_{nx}\left[\mathbb{E}^{\pi}_{nx}\left[\sum_{k=n+1}^{N-1} r_k\big(X_k, f_k(X_k)\big) + g_N(X_N)|X_{n+1}\right]\right] \\
&= r_n\big(x, f_n(x)\big) \\
&\quad + \int \mathbb{E}^{\pi}_{n+1,x'}\left[\sum_{k=n+1}^{N-1} r_k\big(X_k, f_k(X_k)\big) + g_N(X_N)\right] Q_n\big(dx'|x, f_n(x)\big) \\
&= r_n\big(x, f_n(x)\big) + \int V_{n+1,\pi}(x') Q_n\big(dx'|x, f_n(x)\big).
\end{aligned}
$$

donde se han usados las propiedades de $\mathbb{P}^{\pi}_{nx}$ en la cuarta igualdad.

b) Se sigue de a) por inducción.

□

La siguiente definición es crucial para la solución de Problemas de Decisión de Markov.

Definición 1.3.2. Sea $v \in \mathbb{M}(E)$. Una regla de decisión $f \in F_n$ se llamará un *maximizador* de v en el tiempo n si $\mathscr{T}_{nf} v = \mathscr{T}_n v$, es decir, para todo $x \in E$, $f(x)$ es un punto máximo de la función $a \mapsto L_n v(x,a), a \in D_n(x)$.

El Teorema 1.3.3 da el método de solución para los Problemas de Decisión de Markov. Éstos se pueden resolver por aplicaciones sucesivas de los $\mathscr{T}_n$-operadores. Como se mencionó anteriormente, en general, no es cierto que $\mathscr{T}_n v \in \mathbb{M}(E)$, para $v \in \mathbb{M}(E)$. Sin embargo, se puede demostrar que V_n es medible analíticamente y que la sucesión (V_n) satisface la *Ecuación de Bellman*

$$
\begin{aligned}
V_N &= g_N \\
V_n &= \mathscr{T}_n V_{n+1}, \qquad n = 0, 1, \ldots, N-1,
\end{aligned}
$$

ver por ejemplo [5]. El siguiente teorema muestra que cuando una solución de la ecuación de Bellman existe junto con una sucesión de maximizadores, entonces ésta obtiene la solución al problema de optimización.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 19 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 20 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Teorema 1.3.2. *(Teorema de Verificación).* Sea $(v_n) \subset \mathbb{M}(E)$ una solución de la ecuación de Bellman. Entonces, se satisface:

a) $v_n \geq V_n$ para $n = 0, 1, \ldots, N$.

b) Si f_n^* es un maximizador de v_{n+1} para $n = 0, 1, \ldots, N-1$, entonces $v_n = V_n$ y la política $\pi^* = (f_0^*, f_1^*, \ldots, f_{N-1}^*)$ es óptima para el Problema de Decisión de Markov de N etapas.

Demostración. a) Para $n = N$ se tiene

$$v_N = g_N = V_N.$$

Suponer que $v_{n+1} \geq V_{n+1}$, entonces para toda $\pi = (f_0, f_1, \ldots, f_{N-1})$

$$v_n = \mathscr{T}_n v_{n+1} \geq \mathscr{T}_n V_{n+1} \geq \mathscr{T}_{n,f_n} V_{n+1,\pi} = V_{n\pi}.$$

Tomando el supremo sobre todas las políticas π resulta que $v_n \geq V_n$.

b) Se demuestra recursivamente que $v_n = V_n = V_{n\pi^*}$. Para $n = N$ es obvio. Suponer que la afirmación es verdadera para $n+1$, entonces

$$V_n \leq v_n = \mathscr{T}_{nf_n^*} v_{n+1} = \mathscr{T}_{nf_n^*} V_{n+1} = V_{n\pi^*} \leq V_n$$

y el teorema se cumple.

□

El Teorema de Verificación es suficiente para aplicaciones donde una solución de la ecuación de Bellman es fácil al igual que la existencia de maximizadores; por ejemplo, si los espacios de estados y acciones son finitos. En general, la existencia de una política óptima no está garantizada. Se tiene que realizar otras suposiciones acerca de la estructura del problema para asegurar esto. También notar que el valor de un problema de optimización siempre es único mientras que una política óptima puede no ser única. En lo que sigue se hará la siguiente suposición de estructura.

Suposiciones de Estructura (SE_N)**:** Existen conjuntos $\mathbb{M}_n \subset \mathbb{M}(E)$ y $\Delta_n \subset F_n$ tales que para todo $n = 0, 1, \ldots, N-1$:

(i) $g_N \in \mathbb{M}_N$.

(ii) Si $v \in \mathbb{M}_{n+1}$ entonces $\mathscr{T}_n v$ está bien definido y $\mathscr{T}_n v \in \mathbb{M}_n$.

(iii) Para todo $v \in \mathbb{M}_{n+1}$ existe un maximizador f_n de v con $f_n \in \Delta_n$.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 20 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 21 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

El siguiente teorema muestra cómo problemas de decisión de Markov pueden resolverse recursivamente, resolviendo N problemas de optimización de una etapa.

Teorema 1.3.3. *(Teorema de Estructura).* Suponer SE_N. Entonces:

a) $V_n \in \mathbb{M}_n$ y la sucesión (V_n) satisface la ecuación de Bellman, es decir, para $n = 0, 1, \ldots, N-1$

$$V_N(x) = g_N(x)$$

$$V_n(x) = \sup_{a \in D_n(x)} \left\{ r_n(x,a) + \int V_{n+1}(x') Q_n(dx'|x,a) \right\}, \quad x \in E.$$

b) $V_n = \mathscr{T}_n \mathscr{T}_{n+1} \cdots \mathscr{T}_{N-1} g_N$.

c) Para $n = 0, 1, \ldots, N-1$, existen maximizadores f_n de V_{n+1} con $f_n \in \Delta_n$, y cualquier otra sucesión de maximizadores f_n^* de V_{n+1} define una política óptima $(f_0^*, f_1^*, \ldots, f_{N-1}^*)$ para el Problema de Decisión de Markov de N etapas.

Demostración. Primero se demostrará *a)* y *c)*. Se demostrará por inducción sobre $n = N-1, \ldots, 0$ que $V_n \in \mathbb{M}_n$ y que

$$V_{n\pi^*} = \mathscr{T}_n V_{n+1} = V_n$$

donde $\pi^* = (f_0^*, \ldots, f_{N-1}^*)$ es la política generada por los maximizadores de $V_1, \ldots, V_N$ y $f_n^* \in \Delta_n$. Se tiene que $V_N = g_N \in \mathbb{M}_N$ por (SE_N)(i). Ahora, suponer que la afirmación es verdadera para $N-1, \ldots, n+1$. Puesto que $V_k \in \mathbb{M}_k$ para $k = N, \ldots, n+1$, los maximizadores $f_n^*, \ldots, f_{N-1}^*$ existen y con la iteración de recompensas y la hipótesis de inducción se obtiene

$$V_{n\pi^*} = \mathscr{T}_{nf_n^*} V_{n+1,\pi^*} = \mathscr{T}_{nf_n^*} V_{n+1} = \mathscr{T}_n V_{n+1}.$$

Por lo tanto $V_n \geq \mathscr{T}_n V_{n+1}$. De otra manera, para una política arbitraria π

$$V_{n\pi} = \mathscr{T}_{nf_n} V_{n+1,\pi} \leq \mathscr{T}_{nf_n} V_{n+1} \leq \mathscr{T}_n V_{n+1},$$

donde se ha usado la propiedad que preserva el orden de $\mathscr{T}_{nf_n}$. Tomando el supremo sobre todas las políticas se obtiene $V_n \leq \mathscr{T}_n V_{n+1}$. Con todo esto junto se sigue que

$$V_{n\pi^*} = \mathscr{T}_n V_{n+1} = V_n$$

y por (SE_N), $V_n \in \mathbb{M}_n$.

Para b), proceder nuevamente por inducción sobre $n = N-1, \ldots, 0$, y de a) y por (SE_N) se satisface el resultado.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 21 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 22 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

□

Del Teorema 1.3.3 se concluye directamente el siguiente corolario.

Corolario 1.3.1. Considerar que (SE_N) se satisface. Si $n \leq m \leq N$, entonces

$$V_n(x) = \sup_{\pi} \mathbb{E}^{\pi}_{n,x} \left[\sum_{k=n}^{m-1} r_k\big(X_k, f_k(X_k)\big) + V_m(X_m) \right], \quad x \in E.$$

El Teorema 1.3.3 implica el siguiente algoritmo recursivo para resolver Problemas de Decisión de Markov:

Algoritmo de inducción hacia atrás.

1. Sea $n := N$ y para $x \in E$:
$$V_N(x) := g_N(x).$$
2. Sustituir n por $n-1$ y calcular para toda $x \in E$
$$V_n(x) = \sup_{a \in D_n(x)} \left\{ r_n(x,a) + \int V_{n+1}(x') Q_n(dx'|x,a) \right\}.$$
Calcular un maximizador f_n^* de V_{n+1}.
3. Si $n = 0$, entonces la función de valor V_0 está calculada y la política óptima está dada por $\pi^* = (f_0^*, \ldots, f_{N-1}^*)$. En otro caso, ir al paso 2.

El Teorema 1.3.3 dice que los maximizadores producen una estrategia óptima. Sin embargo, lo inverso no es cierto.

El Principio de Programación Dinámica, informalmente, dice que cuando se tiene una política óptima π^* sobre un cierto horizonte N al considerar el proceso sólo en un subintervalo de $[0,N]$, entonces la política correspondiente que se obtiene al restringir π^* a este subintervalo es de nuevo óptima. Este principio se formaliza en el siguiente teorema:

Teorema 1.3.4. (Principio de Programación Dinámica). Suponer que (SE_N) se satisface. Entonces, para $n \leq m \leq N$

$$V_{n\pi^*}(x) = V_n(x) \quad \Rightarrow \quad V_{m\pi^*} = V_m \quad \mathbb{P}^{\pi^*}_{nx} - c.s.,$$

es decir, si $(f_n^*, \ldots, f_{N-1}^*)$ es óptima para el periodo de tiempo $[n,N]$ entonces $(f_m^*, \ldots, f_{N-1}^*)$ es óptima para $[m,N]$.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 22 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 23 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Demostración. Se sigue de la iteración de recompensas (Teorema 1.3.1) y la definición de V_m que

$$\begin{aligned} V_n(x) &= V_{n\pi^*}(x) \\ &= \mathscr{T}_{nf_n^*} \cdots \mathscr{T}_{m-1,f_{m-1}^*} V_{m\pi^*}(x) \\ &= \mathbb{E}_{nx}^{\pi^*} \left[\sum_{k=n}^{m-1} r_k(X_k, f_k^*(X_k)) + V_{m\pi^*}(X_m) \right] \\ &\leq \mathbb{E}_{nx}^{\pi^*} \left[\sum_{k=n}^{m-1} r_k\big(X_k, f_k^*(X_k)\big) + V_m(X_m) \right] \\ &\leq V_n(x), \end{aligned}$$

donde la última desigualdad es consecuencia del Corolario 1.3.1. Esto implica que

$$\mathbb{E}_{nx}^{\pi^*}[V_m(X_m) - V_{m\pi^*}(X_m)] = 0,$$

lo cual se entiende que $V_{m\pi^*} = V_m$, $\mathbb{P}_{nx}^{\pi^*} - c.s.$

□

Observación 1.3.2. En lugar de recompensas de una etapa r_n y una recompensa terminal g_N, algunas veces los problemas tienen costos de una etapa c_n y un costo terminal h_N. En este caso se quiere minimizar

$$\mathbb{E}_{nx}^{\pi} \left[\sum_{k=n}^{N-1} c_k(X_k, f_k(X_k)) + h_N(X_N) \right], \quad x \in E$$

para $\pi = (f_0, \ldots, f_{N-1})$. Pero este problema puede transformarse en un problema de maximización de recompensas $r_n(x,a) := -c_n(x,a)$, $g_N(x) := -h_N(x)$. Así, todas las afirmaciones hasta aquí permanecen válidas. Se usará la misma notación $V_{n\pi}$ y V_n para las funciones de costo bajo la política π y la función de costo mínimo, respectivamente. Además, el operador de costo mínimo tiene la forma

$$(\mathscr{T}_n v)(x) = \inf_{a \in D_n(x)} \left\{ c_n(x,a) + \int v(x') Q_n(dx'|x,a) \right\}.$$

En este caso V_n es también llamada función de costo.

◊

1.4. Modelos de Decisión de Markov Estructurados

En esta sección se dan condiciones suficientes sobre las cuales las suposiciones (SI_N) y (SE_N), de la sección anterior, se satisfacen y así implicar las validaciones de la ecuación

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 23 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 24 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

de Bellman y la existencia de la política óptima. Para (SE_N) se identificarán condiciones que impliquen conjuntos especiales $\mathbb{M}_n$ y Δ_n que puedan ser elegidos. Por supuesto que es interesante elegir los conjuntos $\mathbb{M}_n$ y Δ_n tan pequeños como sea posible (ver [5], [4] y [17]).
Considerar la Suposición de Integrabilidad (SI_N). Esto se cumple cuando el Modelo de Decisión de Markov tiene una función de "acotación superior".

Definición 1.4.1. Una función medible $b : E \to \mathbb{R}_+$ es llamada una *función de acotación superior* para el Modelo de Decisión de Markov si existen $c_r, c_g, \alpha_b \in \mathbb{R}_+$ tal que para todo $n = 0, 1, ..., N-1$:

(i) $r_n^+(x,a) \leq c_r b(x)$ para todo $(x,a) \in D_n$,

(ii) $g_N^+(x) \leq c_g b(x)$ para todo $x \in E$,

(iii) $\int b(x')Q_n(dx'|x,a) \leq \alpha_b b(x)$ para todo $(x,a) \in D_n$.

Si r_n y g_N están acotadas por arriba, entonces obviamente $b \equiv 1$ es una función de acotación superior.

Sea b una función cota superior para el Modelo de Decisión de Markov. Para $v \in \mathbb{M}(E)$, denotar la *norma suprema ponderada* por

$$\| v \|_b := \sup_{x \in E} \frac{| v(x) |}{b(x)}$$

(con la convención $\frac{0}{0} := 0$) y definimos el conjunto

$$\mathbb{B}_b := \{v \in \mathbb{M}(E) \mid \| v \|_b < \infty\}.$$

Equivalentemente $\mathbb{B}_b$ puede ser escrito como

$$\mathbb{B}_b := \{v \in \mathbb{M}(E) \mid |v(x)| \leq cb(x) \text{ para todo } x \in E \text{ y para algún } c \in \mathbb{R}_+\}.$$

El concepto de funciones de acotación superior es de particular interés para los Modelos de Decisión de Markov con horizonte infinito (Capítulo 2). El siguiente resultado es fundamental para muchas aplicaciones.

Proposición 1.4.1. Si el Modelo de Decisión de Markov tiene una función de acotación superior b, entonces $\delta_n^N \in \mathbb{B}_b$ y la Suposición de Integrabilidad (SI_N) se satisface.

Demostración. Puesto que $\delta_n^N \geq 0$, se debe mostrar que $\delta_n^N(x) \leq cb(x)$ para algún $c \in \mathbb{R}_+$. De las propiedades de una función de acotación superior se sigue que

$$\begin{aligned} \mathbb{E}_x^\pi \left[r_k^+ \left(X_k, f_k(X_k) \right) | X_{k-1} \right] &= \int r_k^+ \left(x', f_k(x') \right) Q_k \left(dx' | X_{k-1}, f_{k-1}(X_{k-1}) \right) \\ &\leq c_r \alpha_b b(X_{k-1}) \end{aligned}$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 24 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 25 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

y por inducción se obtiene

$$\mathbb{E}_x^{\pi}\left[r_k^+\left(X_k, f_k(X_k)\right)\right] \leq c_r \alpha_b^k b(x), \quad x \in E.$$

Análogamente se obtiene

$$\mathbb{E}_x^{\pi}\left[g_N^+(X_N))\right] \leq c_g \alpha_b^N b(x), \quad x \in E$$

y el resultado se sigue.

□

Para el resto de la Sección 1.4 asumir que E y A son espacios de Borel (ver la Observación 1.1.1 a)). También, asumir que D_n es un subconjunto de Borel de $E \times A$. Además, suponer que el Modelo de Decisión de Markov tiene una función de acotación superior b e introducir el siguiente conjunto

$$\mathbb{B}_b^+ := \{v \in \mathbb{M}(E) | \quad \| v^+ \|_b < \infty\}.$$

El conjunto $\mathbb{B}_b^{\pm}$ puede escribirse como

$$\mathbb{B}_b^+ := \{v \in \mathbb{M}(E) | \, v^+(x) \leq cb(x) \text{ para todo } x \in E \text{ y para algún } c \in \mathbb{R}_+\}.$$

1.4.1. Modelos de Decisión de Markov Semicontinuos

En los Modelos de Decisión de Markov *semicontinuos* la Suposición de Estructura (SE_N) se cumple cuando $\mathbb{M}_n$ es un subconjunto de funciones semicontinuas superiores.

Proposición 1.4.2. Sea $v \in \mathbb{B}_b^+$ semicontinua superior. Suponer que:

(i) $D_n(x)$ es compacto para todo $x \in E$ y $x \mapsto D_n(x)$ es semicontinua superior,

(ii) $(x,a) \mapsto L_n v(x,a)$ es semicontinua superior en D_n.

Entonces $\mathscr{T}_n v$ es semicontinua superior y existe un maximizador f_n de v.

Observación 1.4.1. La condición (i) en la Proposición 1.4.2 puede ser reemplazada por la siguiente condición: Para todo $x \in E$, los conjuntos de nivel $\{a \in D_n(x) | L_n v(x,a) \geq c\}$ son compactos para todo $c \in \mathbb{R}$ y el mapeo del conjunto de valores

$$x \mapsto \{a \in D_n(x) | L_n v(x,a) = \mathscr{T}_n v(x)\}$$

es semicontinuo superior.

◊

iThenticate Página 25 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 26 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Demostración. Definir

$$w(x,a) := L_n v(x,a), \quad w^*(x) := \mathscr{T}_n v(x)$$

y $D(x) := D_n(x)$. Para $x_0 \in E$ seleccionar una sucesión $(x_n) \subset E$ convergiendo a x_0 tal que el límite de $w^*(x_n)$ existe. Se debe demostrar que

$$\lim_{n\to\infty} w^*(x_n) \leq w^*(x_0).$$

Puesto que $a \mapsto w(x,a)$ es semicontinua superior en el conjunto compacto $D(x)$, alcanza su supremo en $D(x)$. Sea $a_n \in D(x_n)$ un punto máximo de $a \mapsto w(x_n,a)$ en $D(x_n)$. Por la semicontinuidad superior de $x \mapsto D(x)$, existe una subsucesión (a_{n_k}) de a_n convergiendo a algún $a_0 \in D(x_0)$. La semicontinuidad superior de w implica

$$\lim_{n\to\infty} w^*(x_n) = \lim_{k\to\infty} w^*(x_{n_k}) = \lim_{k\to\infty} w(x_{n_k}, a_{n_k}) \leq w(x_0,a_0) \leq w^*(x_0),$$

es decir, w^* es semicontinua superior. Puesto que w y w^* son medibles, se sigue fácilmente que

$$D^* := \{(x,a) \in D | w(x,a) = w^*(x)\}$$

es un subconjunto de Borel de $E \times A$ y cada $D^*(x)$ es compacto puesto que

$$D^*(x) := \{a \in D(x) | w(x,a) \geq w^*(x)\}.$$

Entonces, aplicando el teorema de selección se obtiene un selector f Borel medible para D^*. Este es el maximizador deseado de v.

□

Observación 1.4.2. Si $A \subset \mathbb{R}$ entonces existe un maximizador más pequeño y un maximizador más grande de $v \in \mathbb{B}_b^+$. Note que el conjunto

$$D_n^*(x) := \{a \in D_n(x) \mid L_n v(x,a) = \mathscr{T}_n v(x)\}$$

es compacto para $x \in E$. Entonces el máx $D_n^*(x)$ y el mín $D_n^*(x)$ son maximizadores de v por el Teorema de selección.

◊

Un conjunto de condiciones suficientes en los datos del Modelo de Decisión de Markov con el fin de asegurar que (SE_N) se satisfaga, se presenta más adelante, siendo $\mathbb{M}_n$ el conjunto de funciones semicontinuas superiores $v \in \mathbb{B}_b^+$ y con $\Delta_n := F_n$. La demostración se sigue directamente de la Proposición 1.4.2.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 26 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 27 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Teorema 1.4.1. Suponer que el Modelo de Decisión de Markov tiene una función de acotación superior b y que para todo $n = 0, 1, ..., N-1$ se cumple:

(i) $D_n(x)$ es compacto para todo $x \in E$ y $x \mapsto D_n(x)$ es semicontinua superior,

(ii) $(x,a) \mapsto \int v(x')Q_n(dx'|x,a)$ es semicontinua superior para toda $v \in \mathbb{B}_b^+$ semicontinua superior,

(iii) $(x,a) \mapsto r_n(x,a)$ es semicontinua superior,

(iv) $x \mapsto g_N(x)$ es semicontinua superior.

Entonces los conjuntos $\mathbb{M}_n := \{v \in \mathbb{B}_b^+ \mid v \text{ es semicontinua superior}\}$ y $\Delta_n := F_n$ satisfacen la Suposición de Estructura (SE_N). En particular, $V_n \in \mathbb{M}_n$ y existe un maximizador $f_n^* \in F_n$ de V_{n+1}. La política $(f_0^*, ..., f_{N-1}^*)$ es óptima.

En lugar de verificar directamente la condición (ii) del Teorema 1.4.1, se puede usar de manera alternativa el siguiente Lema:

Lema 1.4.1. Sea b una función continua de acotación superior. Entonces las siguientes afirmaciones son equivalentes:

(i) $(x,a) \mapsto \int v(x')Q(dx'|x,a)$ es semicontinua superior para toda $v \in \mathbb{B}_b^+$ semicontinua superior.

(ii) $(x,a) \mapsto \int b(x')Q(dx'|x,a)$ es continua y $(x,a) \mapsto \int v(x')Q(dx'|x,a)$ es continua y acotada, para toda función v continua y acotada sobre E.

Un kernel estocástico Q con las propiedades anteriores es llamado *débilmente continuo*.

Demostración. La prueba de (ii) implica (i) es como sigue: Sea $v \in \mathbb{B}_b^+$ semicontinua superior. Entonces se tiene $v - cb \leq 0$, para algún $c \in \mathbb{R}_+$, y $x \mapsto v(x) - cb(x)$ es semicontinua superior. Esto implica la existencia de una sucesión $(\tilde{v}_k)$ de funciones continuas y acotadas tales que $\tilde{v}_k \downarrow v - cb$. Debido a la suposición, la función $(x,a) \mapsto \int \tilde{v}(x')Q(dx'|x,a)$ ahora es continua y acotada. Además, la convergencia monótona implica que

$$\int \tilde{v}(x')Q(dx'|x,a) \downarrow \int (v - cb)(x')Q(dx'|x,a) \quad para \quad k \to \infty.$$

Así, se puede concluir que el límite es semicontinuo superiormente. Sin embargo, en vista de la suposición, esto implica que $(x,a) \mapsto \int \tilde{v}(x')Q(dx'|x,a)$ es semicontinua superiormente.
En seguida se prueba que (i) implica (ii): Puesto que b y $-b$ son $\mathbb{B}_b^+$, se obtiene que

$$(x,a) \mapsto \int b(x')Q(dx'|x,a)$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 27 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 28 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

es continua. Si v es acotada y continua, entonces las funciones $v- \parallel v \parallel$ y $-v- \parallel v \parallel$ pertenecen a $\mathbb{B}_b^+$ y son semicontinuas superiormente. Por lo tanto, la función $(x,a) \mapsto \int v(x')Q(dx'|x,a)$ es continua.

□

1.4.2. Modelos de Decisión de Markov Continuos

Enseguida se verá cuándo las Suposiciones de Estructura (SE_N) se satisfacen al ser $\mathbb{M}_n$ un subconjunto de funciones continuas.

Proposición 1.4.3. Sea $v \in \mathbb{B}_b^+$ continua. Suponer que las siguientes suposiciones se satisfacen:

(i) $D_n(x)$ es compacto para todo $x \in E$ y $x \mapsto D_n(x)$ es continuo,

(ii) $(x,a) \mapsto L_n v(x,a)$ es continua en D_n.

Entonces $\mathscr{T}_n v$ es continua y existe un maximizador $f_n \in F_n$ de v. Si v tiene un único maximizador $f_n \in F_n$ al tiempo n, entonces f_n es continua.

Demostración. Se usa la misma notación como en la prueba de la Proposición 1.4.2. En vista de la Proposición 1.4.2 es suficiente mostrar que w^* es semicontinua inferiormente, es decir, que $w^*(x_0) \leq \lim_{n\to\infty} w^*(x_n)$ para cada sucesión $(x_n) \subset E$ la cual converge a $x_0 \in E$ y para la cual el $\lim_{n\to\infty} w^*(x_n)$ existe. Se sabe, por la suposición, que $w^*(x_0) = w(x_0,a_0)$ para alguna $a_0 \in D_n(x_0)$. Puesto que $x \mapsto D(x)$ es continua, existe una subsucesión (x_{n_k}) de (x_n) y una sucesión de puntos $(a_{n_k}) \in D(x_{n_k})$ convergiendo a a_0. Por lo tanto, se tiene $(x_{n_k}, a_{n_k}) \to (x_0, a_0)$. Se sigue de la continuidad de w que

$$w^*(x_0) = w(x_0,a_0) = \lim_{k\to\infty} w(x_{n_k}, a_{n_k}) \leq \lim_{k\to\infty} w^*(x_{n_k}) = \lim_{n\to\infty} w^*(x_n).$$

Puesto que $x \mapsto D(x)$ es semicontinua superiormente, D es cerrado y se sigue que

$$D^* := \{(x,a) \in D \mid w(x,a) = w^*(x)\}$$

es un subconjunto cerrado de D. Entonces se obtiene que $x \mapsto D^*(x)$ también es semicontinua superiormente. Así, si v tiene un único maximizador f_n, es decir, $D_n^*(x) = \{f_n(x)\}$ para todo $x \in E$, entonces f_n debe ser continua.

□

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 28 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 29 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Observación 1.4.3. Si $A \subset \mathbb{R}$, entonces el maximizador más pequeño (o más grande) de v es semicontinuo inferiormente (o semicontinuo superiormente). Esto puede ser visto como sigue: $A \subset \mathbb{R}$, entonces $x \mapsto f(x) := \max D^*(x)$ es el maximizador más grande. Elegir $x_0, x_n \in E$ tal que $x_n \to x_0$ y el límite de $(f_n(x))$ existe. Puesto que $a_n := f(x_n) \in D^*(x_n)$ y $x \mapsto D^*(x)$ es semicontinua superiormente, (a_n) tiene un punto de acumulación en $D^*(x_0)$ el cual debe ser el $\lim_{n\to\infty} f(x_n)$. Se sigue que

$$\lim_{n\to\infty} f(x_n) \leq \max D^*(x_0) = f(x_0),$$

es decir, el maximizador más grande es semicontinuo superiormente. De la misma manera puede mostrarse que el maximizador más pequeño es semicontinuo inferiormente.

$\Diamond$

Teorema 1.4.2. Suponer que un Modelo de Decisión de Markov con una función de acotación superior b está dado y que, para toda $n = 0, 1..., N-1$, se cumple:

(i) $D_n(x)$ es compacto para todo $x \in E$ y $x \mapsto D_n(x)$ es continua,

(ii) $(x,a) \mapsto \int v(x')Q_n(dx'|x,a)$ es continua para toda $v \in \mathbb{B}_b^*$ continua,

(iii) $(x,a) \mapsto r_n(x,a)$ es continua,

(iv) $x \mapsto g_N(x)$ es continua.

Entonces los conjuntos $\mathbb{M}_n := \{v \in \mathbb{B}_b^* \mid v \text{ es continua}\}$ y $\Delta_n := F_n$ satisfacen la Suposición de Estructura (SE_N). Si el maximizaor de V_n es único, entonces Δ_n puede ser elegido como el conjunto de funciones continuas.

1.4.3. Modelos de Decisión de Markov Medibles

Algunas veces la suposición de estructura se puede cumplir con $\mathbb{M}_n = \mathbb{B}_b^+$, para este caso se usa el siguiente resultado.

Proposición 1.4.4. Sea $v \in \mathbb{B}_b^+$ y suponer que lo siguiente se satisface:

(i) $D_n(x)$ es compacto para todo $x \in E$,

(ii) $a \mapsto L_n v(x,a)$ es semicontinua superiormente en $D_n(x)$ para todo $x \in E$.

Entonces $\mathscr{T}_n v$ es medible y existe un maximizador $f_n \in F_n$ de v.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 29 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 30 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Demostración. Se usa la misma notación como en la prueba de la Proposición 1.4.2. Sea $\alpha \in \mathbb{R}$. Es suficiente probar que $\{x \in E | w^*(x) \geq \alpha\}$ es un conjunto de Borel. Pero

$$\begin{aligned}\{x \in E | w^*(x) \geq \alpha\} &= \{x \in E | w(x,a) \geq \alpha \text{ para alguna } a \in D(x)\} \\ &= \text{proy}_E\{(x,a) \in D | w(x,a) \geq \alpha\}.\end{aligned}$$

Este conjunto es de Borel por un resultado de Kunugui y Novikov (ver [32]), ya que D es Borel con valores compactos (es decir, secciones verticales compactas) y $\{(x,a) \in D | w(x,a) \geq \alpha\}$ es un subconjunto de Borel de D con valores cerrados (y por lo tanto compactos). En realidad, Kunugui y Novikov prueban que la proyección de un subconjunto de Borel de $E \times A$ con valores compactos es un subconjunto de Borel de E. La existencia de un maximizador puede ser mostrado de la misma manera como en la prueba de la Proposición 1.4.2.

□

Observación 1.4.4. *a)* Si A es numerable y $D_n(x)$ es finito para todo $x \in E$, entonces las suposiciones (i) y (ii) de la Proposición 1.4.4 se cumplen.

b) La condición (i) de la Proposición 1.4.4 puede ser reemplazada por: Para todo $x \in E$ el conjunto de nivel $\{a \in D_n(x) | L_n v(x,a) \geq c\}$ es compacto para todo $c \in \mathbb{R}$.

◊

De manera particular, el Teorema 1.3.3 se cumple bajo las condiciones del Teorema 1.4.3.

Teorema 1.4.3. Suponer que un Modelo de Decisión de Markov con una función b de acotación superiormente es dado y que para toda $n = 0, 1, ..., N-1$ se cumple:

(i) $D_n(x)$ es compacto para todo $x \in E$,

(ii) $a \mapsto \int v(x')Q_n(dx'|x,a)$ es semicontinua superiormente para toda $v \in \mathbb{B}_b^+$ y para toda $x \in E$,

(iii) $a \mapsto r_n(x,a)$ es semicontinua superiormente para todo $x \in E$.

Entonces los conjuntos $\mathbb{M}_n := \mathbb{B}_b^+$ y $\Delta_n := F_n$ satisfacen la Suposición de Estructura (SE_N).

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 30 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 31 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

1.5. Modelos de Decisión de Markov Estacionarios

El *modelo de decisión de Markov estacionario* es aquel donde los datos no dependen de n y está dado por (E, A, D, Q, r_n, g_N) con $r_n := \beta^n r$, $g_N := \beta^N g$ y $\beta \in (0,1]$.

Se denota por F al conjunto de todas las reglas de decisión $f : E \to A$ con $f(x) \in D(x)$ para $x \in E$. Entonces F^N es el conjunto de todas las políticas $\pi = (f_0, ..., f_{N-1})$ de N etapas.

La recompensa esperada descontada en n etapas bajo una política $\pi \in F^n$ está dada por

$$J_{n\pi}(x) := \mathbb{E}_x^{\pi}\left[\sum_{k=0}^{n-1} \beta^k r(X_k, f_k(X_k)) + \beta^n g(X_n)\right], \quad x \in E,$$

cuando el sistema inicia en el estado $x \in E$. La recompensa máxima esperada descontada en n etapas está definida mediante

$$\begin{aligned} J_0(x) &:= g(x) \\ J_n(x) &:= \sup_{\pi \in F^n} J_{n\pi}(x), \quad x \in E, 1 \le n \le N. \end{aligned}$$

Con el fin de obtener un problema de optimización estocástico bien definido se necesita la siguiente suposición de integrabilidad.

Suposición (S_N): Para $x \in E$

$$\delta_N(x) := \sup_{\pi} \mathbb{E}_x^{\pi}\left[\sum_{k=0}^{N-1} \beta^k r^+(X_k, f_k(X_k)) + \beta^N g^+(X_N)\right] < \infty.$$

Observación 1.5.1. Puesto que $\delta_0 := g^+ \le \delta_{n-1} \le \delta_n \le \delta_N$ y

$$\delta_n^N = \beta^n \delta_{N-n}, \quad n = 0, 1, ..., N,$$

la Suposición de Integrabilidad (SI_N) es equivalente a la Suposición de Integrabilidad de la Sección 1.2 cuando se tiene un Modelo de Decisión de Markov estacionario.

$\Diamond$

Como lo explicado en la Sección 1.4, es conveniente mostrar que (SI_N) se satisface probando la existencia de una función de acotación superior. La definición de una función de acotación superior para un Modelo de Decisión de Markov estacionario es como sigue.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 31 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 32 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Definición 1.5.1. Un mapeo medible $b : E \to \mathbb{R}_+$ es llamado *función de acotación superior* para el Modelo de Decisión de Markov estacionario si existe $c_r, c_g, \alpha_b \in \mathbb{R}_+$, tales que:

(i) $r^+(x,a) \leq c_r b(x)$ para todo $(x,a) \in D$,

(ii) $g^+(x,a) \leq c_g b(x)$ para todo $x \in E$,

(iii) $\int b(x')Q(dx'|x,a) \leq \alpha_b b(x)$ para todo $(x,a) \in D$.

Si el Modelo de Decisión de Markov estacionario tiene una función de acotación superior, entonces se tiene $\delta_N \in \mathbb{B}_b$ y la Suposición de integrabilidad (S_N) se satisface. Claramente cada modelo estacionario es un modelo no estacionario especial. Se obtiene la siguiente relación entre las funciones de valor J_n y V_n:

$$V_n(x) = \beta^n J_{N-n}(x), \quad x \in E, \; n = 0, 1, ..., N.$$

Pero, de otro modo, cada Modelo de Decisión de Markov no estacionario puede ser formulado como uno estacionario. La idea es ampliar el espacio de estados incluyendo el parámetro tiempo (ver [16]).

Como en la Definición 1.3.1, se introducen los siguientes operadores para $v \in \mathbb{M}(E)$:

$$\begin{aligned}(Lv)(x,a) &:= r(x,a) + \beta \int v(x')Q(dx'|x,a), \quad (x,a) \in D,\\ \mathscr{T}_f v(x) &:= (Lv)(x, f(x)), \quad x \in E,\\ \mathscr{T} v(x) &:= \sup_{a \in D(x)} Lv(x,a), \quad x \in E.\end{aligned}$$

$\mathscr{T}$ es llamado el *operador de recompensa máxima*.

Teorema 1.5.1. ***(Iteración de Recompensas).*** Para $\pi = (f_0, ..., f_{n-1})$ se cumple:

$$J_{n\pi} = \mathscr{T}_{f_0} \cdots \mathscr{T}_{f_{n-1}} \mathrm{g}.$$

La suposición de estructura (SE_N) tiene que ser modificada para los Modelos de Decisión de Markov Estacionarios.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 32 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Suposiciones de Estructura (SE_N): Existen conjuntos $\mathbb{M} \subset \mathbb{M}(E)$ y $\Delta \subset F$ tales que:

(i) $g \in \mathbb{M}$.

(ii) Si $v \in \mathbb{M}$ entonces $\mathscr{T}v(x)$ está bien definido y $\mathscr{T}v \in \mathbb{M}$.

(iii) Para toda $v \in \mathbb{M}$ existe un maximizador $f \in \Delta$ de v, es decir,

$$\mathscr{T}_f v(x) = \mathscr{T}v(x), \quad x \in E.$$

El Teorema 1.3.3, acerca del cálculo recursivo de las funciones de valor óptimas, ahora tiene la siguiente forma.

Teorema 1.5.2. ***(Teorema de Estructura).*** Supongamos que SE_N se satisface.

a) Entonces $J_n \in \mathbb{M}$ y la ecuación de Bellman $J_n = \mathscr{T}J_{n-1}$ se cumple, es decir,

$$J_0(x) = g(x)$$

$$J_n(x) = \sup_{a \in D(x)} \left\{ r(x,a) + \beta \int J_{n-1}(x') Q(dx'|x,a) \right\}, \quad x \in E.$$

Además, $J_n = \mathscr{T}^n g$.

b) Para $n = 1, \ldots, N$ existen maximizadores f_n^* de J_{n-1} con $f_n^* \in \Delta$ y cada sucesión de maximizadores f_n^* de J_{n-1} define una política óptima $(f_N^*, \ldots, f_1^*)$ para el problema de decisión de Markov estacionario de N-etapas.

Los casos simples surgen cuando E y A son finitos. Aquí la Suposición de Estructura (SE_N) se satisface con $\mathbb{M} := \{v : E \to \mathbb{R}\}$ porque cada función es medible y existen maximizadores. Además, el kernel de transición tiene una densidad discreta y se denota por

$$q(x'|x,a) := Q(\{x'\}|x,a)$$

para $x, x' \in E$ y $a \in D(x)$.

Análogamente para el caso no estacionario, el Teorema 1.5.2 da un algoritmo recursivo para resolver Problemas de Decisión de Markov. Debido a la estacionarieidad de los datos, no es necesario formular el algoritmo como un algoritmo hacia atrás.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 34 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Algoritmo de Inducción hacia adelante.

1. Sea $n := 0$ y para $x \in E$:
$$J_0(x) := g(x)$$
2. Sustituir n por $n+1$ y para todo $x \in E$
$$J_n(x) = \sup_{a \in D(x)} \left\{ r(x,a) + \int J_{n-1}(x') Q(dx'|x,a) \right\}.$$
Calcular un maximizador f_n^* de J_{n-1}.
3. Si $n = N$, entonces la función de valor J_N es calculada y la política óptima π^* está dada por $\pi^* = (f_N^*, ..., f_1^*)$. En otro caso, ir al paso 2.

El *algoritmo de inducción* calcula las funciones de valor de n etapas y las reglas de decisión recursivamente sobre las etapas, iniciando con la función de recompensa terminal.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 34 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco.
México.

iThenticate Página 36 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

Capítulo 2

Teoría de Procesos de Decisión de Markov de Horizonte Infinito

En este capítulo se consideran los Procesos de Decisión de Markov con horizonte de tiempo infinito. Existen situaciones donde los modelos con tiempo de horizonte infinito surgen de manera natural, por ejemplo, cuando se considera la variable tiempo de vida del inversionista. Sin embargo, es más importante el hecho de que los Modelos de Decisión de Markov finitos, pero grandes, pueden ser aproximados mediante modelos de horizonte infinito (ver [1], [8] y [17]).

2.1. Modelos de Decisión de Markov de Horizonte Infinito

Los Modelos de Decisión de Markov con horizonte infinito pueden verse como una aproximación de un modelo con horizonte finito, pero grande. Comúnmente el modelo de horizonte infinito es fácil resolver y su política óptima produce una política razonable para el modelo con horizonte finito. Esta afirmación se formaliza más adelante en esta sección. En lo que sigue siempre se asumirá que está dado un Modelo de Decisión de Markov (estacionario) con horizonte infinito.

Definición 2.1.1. Un *Modelo de Decisión de Markov estacionario* con horizonte infinito consiste de un conjunto de datos (E, A, D, Q, r, β), donde E, A, D, Q, r y β están dados en la Definición 1.1.1 (ver también la Observación 1.1.1 c) y la Sección 1.5). No hay recompensa terminal, es decir, $g \equiv 0$.

iThenticate Página 36 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 37 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Sea $\pi = (f_0, f_1, ...) \in F^\infty$ una política para el Modelo de Decisión de Markov de horizonte infinito. Entonces definir

$$J_{\infty\pi}(x) := \mathbb{E}_x^\pi \left[\sum_{k=0}^{\infty} \beta^k r\big(X_k, f_k(X_k)\big) \right], \quad x \in E,$$

lo cual da la recompensa esperada descontada bajo la política π sobre un horizonte a tiempo infinito cuando se inicia en el estado x. El proceso (X_k) es el correspondiente al *Proceso de Decisión de Markov* de horizonte infinito. El problema de control es entonces

$$J_\infty(x) := \sup_\pi J_{\infty\pi}(x), \quad x \in E. \tag{2.1.1}$$

La función $J_\infty(x)$ da la recompensa máxima esperada descontada sobre un horizonte a tiempo infinito, cuando se inicia en el estado x. Una política $\pi^* \in F^\infty$ es llamada *óptima* si $J_{\infty\pi^*}(x) = J_\infty(x)$ para todo $x \in E$.

Con el fin de obtener un problema bien definido asumir todo el tiempo la siguiente suposición

Suposición de Integrabilidad (*SI*):

$$\delta(x) := \sup_\pi \mathbb{E}_x^\pi \left[\sum_{k=0}^{\infty} \beta^k r^+\big(X_k, f_k(X_k)\big) \right] < \infty, \quad x \in E.$$

Además, es conveniente introducir el conjunto

$$\mathbb{B} := \{v \in \mathbb{M}(E) \mid v(x) \leq \delta(x) \text{ para todo } x \in E\}.$$

Claramente, $J_{\infty\pi} \in \mathbb{B}$ para toda política π. Con el fin de garantizar que el problema de horizonte infinito es una aproximación del modelo de horizonte finito, se hace la siguiente suposición de convergencia en todo este capítulo.

Suposición de Convergencia (*SC*):

$$\lim_{n\to\infty} \sup_\pi \mathbb{E}_x^\pi \left[\sum_{k=n}^{\infty} \beta^k r^+\big(X_k, f_k(X_k)\big) \right] = 0, \quad x \in E.$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 37 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 38 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Más adelante se discutirá la suposición de convergencia y se darán algunas condiciones simples, las cuales implican (SC). Para este propósito las *funciones de acotación superior* serán importantes. Una función de acotación superior para un Modelo de Decisión de Markov con horizonte infinito es definido como sigue.

Definición 2.1.2. Una función medible $b : E \to \mathbb{R}_+$ es llamada *función de acotación superior* para el Modelo de Decisión de Markov con horizonte infinito si existen constantes $c_r, \alpha_b \in \mathbb{R}_+$ tales que:

(i) $r^+(x,a) \leq c_r b(x)$ para todo $(x,a) \in D$,

(ii) $\int b(x')Q(dx'|x,a) \leq \alpha_b b(x)$ para todo $(x,a) \in D$.

Introducimos el operador $\mathscr{T}_\circ$ mediante

$$\mathscr{T}_\circ v(x) := \sup_{a \in D(x)} \beta \int v(x')Q(dx'|x,a), \quad x \in E,$$

cuando la integral existe. El operador $\mathscr{T}_\circ$ causa un tiempo de falla mediante un paso del tiempo, es decir, $\mathscr{T}_\circ v(x)$ es la esperanza máxima descontada de $v(X_1)$ cuando el estado actual es x. En particular se cumple:

$$\sup_\pi \mathbb{E}_x^\pi \left[\sum_{k=n}^{\infty} \beta^k r^+\big(X_k, f_k(X_k)\big) \right] = \beta^n \sup_\pi \mathbb{E}_x^\pi [\delta(X_n)] = \mathscr{T}_\circ^n \delta(x).$$

La última ecuación es cierta puesto que δ es una función de valor de un Modelo de Decisión de Markov positivo (ver Sección 2.4) y se aplica la Observación 2.3.14. En particular si $v \in \mathbb{B}$ entonces $\mathscr{T}_\circ v \leq \delta$. La Suposición de Convergencia (SC) también puede ser expresada mediante la condición

$$\lim_{n\to\infty} \mathscr{T}_\circ^n \delta(x) = 0, \quad x \in E.$$

En particular, ahora suponer que el Modelo de Decisión de Markov tiene una función de acotación superior b y recordar la definición del conjunto

$$\mathbb{B}_b^+ := \{v \in \mathbb{M}(E) | v^+(x) \leq cb(x) \quad \text{para algún} \quad c \in \mathbb{R}_+\}.$$

Si $\| \delta \|_b < \infty$ y $\mathscr{T}_\circ^n b(x) \to 0$, entonces la Suposición de Integrabilidad (SI) y la Suposición de Convergencia (SC) se cumplen. En seguida definimos α_b mediante

$$\alpha_b := \sup_{(x,a)\in D} \frac{\int b(x')Q(dx'|x,a)}{b(x)}.$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 38 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 39 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Entonces obtenemos

$$\mathscr{T}_\circ b(x) = \sup_{a\in D(x)} \beta \frac{\int b(x')Q(dx'|x,a)}{b(x)} b(x) \leq \beta\alpha_b b(x)$$

y por inducción

$$\mathscr{T}_\circ^n b(x) \leq (\beta\alpha_b)^n b(x).$$

Además se tiene

$$\delta(x) \leq c_r \sum_{k=0}^{\infty} (\beta\alpha_b)^k b(x).$$

Estas desigualdades implican que (SI) y (SC) se satisfacen cuando $\beta\alpha_b < 1$. Tal situación aparece en el *caso descontado*, es decir, si la función de recompensa está acotada por arriba y $\beta < 1$. En este caso se tiene $\alpha_b \leq 1$.
Las Suposiciones (SI) y (SC) también se satisfacen para los Modelos de Decisión de Markov *negativos*, es decir, si la función de recompensa satisface $r \leq 0$.

Observación 2.1.1. Se puede demostrar que las políticas dependientes de la historia no mejoran la recompensa máxima esperada descontada. Más precisamente, sea Π_∞ el conjunto de las políticas dependientes de la historia para el Modelo de Decisión de Markov de horizonte infinito. Entonces se cumple que:

$$J_\infty(x) = \sup_{\pi\in F^\infty} J_{\infty\pi}(x) = \sup_{\pi\in\Pi_\infty} J_{\infty\pi}(x), \quad x \in E.$$

Para una prueba ver Hinderer (1970), Teorema 18.4. Este resultado también es válido para Modelos de Decisión de Markov positivos los cuales son examinados en la Sección 2.4.

$\Diamond$

La suposición de convergencia implica las siguientes propiedades de monotonicidad de las funciones de valor

$$J_{n\pi} = \mathscr{T}_{f_0} \ldots \mathscr{T}_{f_{n-1}} 0,$$
$$J_n = \sup_\pi J_{n\pi}.$$

Lema 2.1.1. Para $n, m \in \mathbb{N}_0$ con $n \geq m$ se cumple:

a) $J_{n\pi} \leq J_{m\pi} + \mathscr{T}_\circ^m \delta$.

b) $J_n \leq J_m + \mathscr{T}_\circ^m \delta$.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 39 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 40 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Demostración. Para la parte a) sea $\pi = (f_0, f_1, ...)$ un política de infinitas etapas. Mediante la iteración de recompensa, para $n \geq m$ se obtiene

$$\begin{aligned} J_{n\pi} &= \mathscr{T}_{f_0} ... \mathscr{T}_{f_{n-1}} 0, \\ &\leq \mathscr{T}_{f_0} ... \mathscr{T}_{f_{m-1}} 0 + \sup_{\pi'} \mathbb{E}_x^{\pi'} \left[\sum_{k=m}^{n-1} \beta^k r^+ \left(X_k, f'_k(X_k) \right) \right] \\ &\leq \mathscr{T}_{f_0} ... \mathscr{T}_{f_{m-1}} 0 + \mathscr{T}_\circ^m \delta = J_{m\pi} + \mathscr{T}_\circ^m \delta. \end{aligned}$$

La parte b) se sigue al tomar el supremo sobre todas las políticas en la parte a)

□

En vista del Lema 2.1.1, la Suposición de Convergencia (SC) implica que las sucesiones $(J_{n\pi})$ y (J_n) son decrecientes débilmente. Se deduce que los límites $\lim_{n\to\infty} J_{n\pi}$ y $\lim_{n\to\infty} J_n$ existen. Además, para $\pi \in F^\infty$ se obtiene

$$J_{\infty\pi} = \lim_{n\to\infty} J_{n\pi}.$$

Definir el *límite de la función de valor* mediante

$$J(x) := \lim_{n\to\infty} J_n(x) \leq \delta(x), \quad x \in E.$$

Por definición se cumple que $J_{n\pi} \leq J_n$ para todo $n \in \mathbb{N}$, entonces $J_{\infty\pi} \leq J$ para toda política π. Tomando el supremo sobre todas las políticas π implica

$$J_\infty(x) \leq J(x), \quad x \in E. \tag{2.1.2}$$

Notar que en general $J \neq J_\infty$ y las funciones J_n, J y J_∞ no están en $\mathbb{B}$. Sin embargo, J_∞ y J son analíticamente medibles (ver [5]). El siguiente Lema muestra que el operador L y el límite pueden ser intercambiados.

Lema 2.1.2. Asumir (SC) y sea $f \in F$. Entonces se cumple:

a) $\lim_{n\to\infty} LJ_n(x,a) = LJ(x,a)$ para todo $(x,a) \in D$.

b) $\lim_{n\to\infty} \mathscr{T}_f J_n(x) = \mathscr{T}_f J(x)$ para todo $x \in$.

Demostración. Primero, recordar la definición del operador L

$$Lv(x,a) := r(x,a) + \beta \int v(x') Q(dx'|x,a), \quad (x,a) \in D, v \in \mathbb{B}.$$

Debido a la definición de $\mathscr{T}_f$ es suficiente probar la parte a). Si la sucesión (J_n) es monótona decreciente, la afirmación se sigue directamente del teorema de convergencia monótona. En el caso general considerar

$$w_m(x) := \sup_{n\geq m} J_n(x)$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 40 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 41 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

para $m \in \mathbb{N}$. Claramente las sucesiones (w_m) y (Lw_m) son monótonas decrecientes y por lo tanto $\lim_{m\to\infty} w_m$ y $\lim_{m\to\infty} Lw_m$ existen. Puesto que, por el Lema 2.1.1,

$$w_m(x) \geq J_m(x) \geq w_m(x) - \mathscr{T}_\circ^m \delta(x)$$
$$Lw_m(x,a) \geq LJ_m(x,a) \geq Lw_m(x,a) - \mathscr{T}_\circ^{m+1} \delta(x)$$

para toda $m \in \mathbb{N}$, se obtiene debido a la Condición (C)

$$\lim_{m\to\infty} w_m(x) = \lim_{m\to\infty} J_m(x) \quad y \quad \lim_{m\to\infty} Lw_m(x,a) = \lim_{m\to\infty} LJ_m(x,a).$$

Finalmente, puesto que $w_m \leq \delta$, se obtiene por el Teorema de convergencia monótona

$$\lim_{m\to\infty} LJ_m(x,a) = \lim_{m\to\infty} Lw_m(x,a) = (L \lim_{m\to\infty} w_m)(x,a) = (L \lim_{m\to\infty} J_m)(x,a).$$

□

Si $f^\infty := (f,f,\ldots) \in F^\infty$ es una *política estacionaria*, es decir, se usa la misma regla de decisión en cada etapa, entonces escribir

$$J_f := J_{\infty f^\infty} = \lim_{n\to\infty} \mathscr{T}_f^n 0.$$

Del Lema 2.1.2 se obtiene la *iteración de recompensas* de horizonte infinito.

Teorema 2.1.1. ***(Iteración de Recompensas).*** Asumir (SC) y considerar $\pi = (f,\sigma) \in F \times F^\infty$. Entonces se cumple:

a) $J_{\infty\pi} = \mathscr{T}_f J_{\infty\sigma}$.

b) $J_f \in \mathbb{B}$ y $J_f = \mathscr{T}_f J_f$.

Como en el Capítulo 1, primero se formula un teorema de verificación con el fin de evitar los problemas generales de medibilidad. El teorema afirma que los candidatos para la solución óptima del problema (2.1.1) están dados por puntos fijos del operador $\mathscr{T}$ de recompensa máxima.

Teorema 2.1.2. ***(Teorema de Verificación).*** Asumir (SC) y sea $v \in \mathbb{B}$ un punto fijo de $\mathscr{T}$ tal que $v \geq J_\infty$. Si f^* es un maximizador de v, entonces

$$v = J_\infty$$

y la política estacionaria $(f^*,f^*,\ldots)$ es óptima para el Modelo de Decisión de Markov de etapa infinita.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 41 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 42 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Demostración. Por suposición se obtiene

$$v = \mathscr{T}_{f^*} v = \mathscr{T}_{f^*}^n v \leq \mathscr{T}_{f^*}^n 0 + \mathscr{T}_{\circ}^n \delta.$$

Haciendo $n \to \infty$ se concluye $v \leq J_{f^*} \leq J_\infty$. Así, la afirmación se sigue ya que $J_\infty \leq v$.

□

En lo que sigue se quiere resolver el problema (2.1.1) y al mismo tiempo se quiere interpretar como un "caso límite", es decir, se quiere tener $J = J_\infty$. Con el fin de garantizar lo anterior se requiere de la siguiente suposición de estructura. Adicionalmente se asumen algunas propiedades de la función de valor límite $J = \lim_{n\to\infty} J_n$.

Suposición de Estructura (SE): Existen conjuntos $\mathbb{M} \subset \mathbb{M}(E)$ y $\Delta \subset F$ tales que:

(i) $0 \in \mathbb{M}$.

(ii) Si $v \in \mathbb{M}$ entonces

$$\mathscr{T} v(x) := \sup_{a \in D(x)} \left\{ r(x,a) + \beta \int v(x') Q(dx'|x,a) \right\}, \quad x \in E$$

está bien definida y $\mathscr{T} v \in \mathbb{M}$.

(iii) Para toda $v \in \mathbb{M}$ existe un maximizador $f \in \Delta$ de v.

(iv) $J \in \mathbb{M}$ y $J = \mathscr{T} J$.

En la Sección 2.2 se darán condiciones importantes las cuales implican (SE).

Teorema 2.1.3. ***(Teorema de Estructura).*** Suponer que (SC) y (SE) se satisfacen. Entonces se cumple:

a) $J_\infty \in \mathbb{M}$, $J_\infty = \mathscr{T} J_\infty$ y $J_\infty = J = \lim_{n\to\infty} J_n$ **(Iteración de Valor)**.

b) J_∞ es la función v en $\mathbb{M} \cap \mathbb{B}$ r-subarmónica más grande, es decir, J_∞ es la función v en $\mathbb{M}$ más grande con $v \leq \mathscr{T} v$ y $v \leq \delta$.

c) Existe un maximizador $f \in \Delta$ de J_∞, y cada maximizador f^* de J_∞ define una política óptima estacionaria $(f^*, f^*, \ldots)$ para el Modelo de Decisión de Markov de etapa infinita.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 42 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 43 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Demostración. Observar que se tiene $J_\infty \leq J$. Las partes (iii) y (iv) de la Suposición de Estructura (SE) implican que existe una $f^* \in \Delta$ tal que $\mathscr{T}_{f^*}J = \mathscr{T}J = J$. Así, iterando el operador $\mathscr{T}_{f^*}$ como en el Lema 2.1.1, se obtiene:

$$J = \mathscr{T}_{f^*}^n J \leq \mathscr{T}_{f^*}^n 0 + \mathscr{T}_\circ^n \delta$$

para toda $n \in \mathbb{N}$. Tomando el límite $n \to \infty$ resulta

$$J \leq J_{f^*} \leq J_\infty,$$

ya que siempre $J_{f^*} \leq J_\infty$. Por lo tanto se obtiene

$$J_\infty = J = J_{f^*}$$

y la política estacionaria $(f^*, f^*, \ldots)$ es óptima. Esto implica las partes $a)$ y $c)$. Para la parte $b)$ notar que $J_\infty \in \mathbb{M}, J_\infty = \mathscr{T}J_\infty$ y $J_\infty \leq \delta$. Ahora, sea $v \in \mathbb{M} \cap \mathbb{B}$ otra función r-subarmónica. Entonces

$$v \leq \mathscr{T}^n v \leq \mathscr{T}^n 0 + \mathscr{T}_\circ^n \delta = J_n + \mathscr{T}_\circ^n \delta$$

para toda $n \in \mathbb{N}$. Para $n \to \infty$ se tiene: $v \leq J = J_\infty$.

□

La ecuación $J_\infty = \mathscr{T}J_\infty$ es llamada la *ecuación de Bellman* para el Modelo de Decisión de Markov de horizonte infinito. A veces esta ecuación de punto fijo también es llamada la *ecuación de optimalidad*. La parte $a)$ del teorema anterior muestra que J_∞ es aproximado por J_n para n grande, es decir, el valor del Problema de Decisión de Markov de horizonte infinito se puede obtener iterando el $\mathscr{T}$-operador. Esto justifica la noción de *iteración de valor*.

Observación 2.1.2. *a)* La parte $c)$ del Teorema 2.1.3 muestra que una política óptima se puede encontrar entre las mismas estacionarias. Formalmente se tiene

$$J_\infty = \sup_{f \in F} J_f.$$

b) Bajo las suposiciones del Teorema 2.1.3 la afirmación inversa de la parte $c)$ también es cierta: si f^∞ es óptima, entonces f es un maximizador de J_∞. Esto se puede ver como sigue: Suponer que f^∞ es óptima, es decir, $J_\infty = J_f$. Puesto que $J_f = \mathscr{T}_f J_f$ por el Teorema 2.1.1 se sigue que $\mathscr{T}_f J_\infty = J_\infty = \mathscr{T}J_\infty$, es decir, es un maximizador de J_∞.

◊

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 43 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 44 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

2.2. Modelos de Decisión de Markov Semicontinuos

En esta sección se muestran algunas condiciones de continuidad y compactación que implican la Suposición de Estructura (SE) y las cuales son usadas en las aplicaciones. En lo que sigue considerar E y A espacios de Borel, D como un subconjunto de Borel de $E \times A$ y definir

$$D_n^*(x) := \{a \in D(x) | a \text{ es un punto máximo de } a \mapsto LJ_{n-1}(x,a)\}$$

para $n \in \mathbb{N} \cup \{\infty\}$ y $x \in E$, y definir por

$$LsD_n^*(x) := \{a \in A | a \text{ es un punto de acumulación de una sucesión } (a_n) \text{ con}$$
$$a_n \in D_n^*(x) \text{ para } n \in \mathbb{N}\}$$

el *límite superior de la sucesión conjunto* $(D_n^*(x))$. El conjunto $D_n^*(x)$ consiste de las acciones óptimas en el estado x para un Problema de Decisión de Markov de n etapas y $D_\infty^*(x)$ son las acciones óptimas para el problema de horizonte infinito. Se puede plantear el siguiente teorema (ver también el Teorema 1.4.1.)

Teorema 2.2.1. Suponer que existe una función de acotación superior b con $\mathscr{T}_\circ^n b \to 0$, cuando $n \to \infty, \| \delta \|_b < \infty$ y que las siguientes condiciones se satisfacen:

(i) $D(x)$ es compacto para todo $x \in E$ y $x \mapsto D(x)$ es semicontinua superiormente.

(ii) $(x,a) \mapsto \int v(x')Q(dx'|x,a)$ es semicontinua superiormente para toda $v \in \mathbb{B}_b^+$ semicontinua superiormente.

(iii) $(x,a) \mapsto r(x,a)$ es semicontinua superiormente.

Entonce se cumple:

a) $J_\infty \in \mathbb{B}_b^+$, $J_\infty = \mathscr{T} J_\infty$ y $J_\infty = J$ **(Iteración de Valor)**.

b) Si b es semicontinua superiormente entonces J_∞ es semicontinua superiormente.

c) $\emptyset \neq LsD_n^*(x) \subset D_\infty^*(x)$ para todo $x \in E$ **(Iteración de Política)**.

d) Existe una $f^* \in F$ con $f^*(x) \in LsD_n^*(x)$ para todo $x \in E$, y la política estacionaria $(f^*, f^*, \ldots)$ es óptima.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 44 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 45 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Demostración. La suposición (SC) se satisface puesto que $\delta \in \mathbb{B}_b^+$ y $\mathscr{T}_\circ^n b \to 0$ cuando $n \to \infty$. Entonces considerar

$$D^*(x) := \{a \in D(x) |\ a \text{ es un punto máximo de } a \longmapsto LJ(x,a)\}$$

para $x \in E$. Primero se prueba que $\emptyset \neq LsD_n^*(x) \subset D(x)^*$, $x \in E$. Considerar $x \in E$ fijo y definir para $n \in \mathbb{N}$ las funciones $v_n : D(x) \to \mathbb{R} \cup \{\infty\}$ por

$$v_n(a) := LJ_{n-1}(x,a).$$

Con el Lema 2.1.2 se sigue que el $\lim_{n\to\infty} v_n(a) = LJ(x,a)$, para toda $a \in D(x)$. El Teorema 1.4.1 y las suposiciones i)-iii) implican que v_n es semicontinua superior, puesto que $J_{n-1} \in \mathbb{B}_b^+$ y J_{n-1} es semicontinua superior. Del Lema 2.1.1 b) obtenemos para $n \geq m$

$$\begin{aligned} v_n(a) &\leq L(J_{m-1} + \mathscr{T}_\circ^{m-1}\delta)(x,a) \leq LJ_{m-1}(x,a)\mathscr{T}_\circ^m\delta(x) \\ &= v_m(a) + \mathscr{T}_\circ^m\delta(x). \end{aligned}$$

Así, se obtiene $\emptyset \neq LsD_n^*(x) \subset D^*(x)$ y

$$\begin{aligned} \mathscr{T}J(x) &= \sup_{a\in D(x)} \lim_{n\to\infty} v_n(a) = \lim_{n\to\infty} \sup_{a\in D(x)} v_n(a) \\ &= \lim_{n\to\infty} \mathscr{T}J_{n-1}(x) = \lim_{n\to\infty} J_n(x) = J(x). \end{aligned}$$

La función límite J es medible ya que J_n es medible, así $J \in \mathbb{B}$. Por el Teorema 1.4.1 existen reglas de decisión $f_n^* \in F$ con $f_n^* \in D_n^*(x)$ para todo $x \in E$ y $n \in \mathbb{N}$. En vista al teorema de selección medible existe una regla de decisión $f_n^* \in F$ tal que

$$f_n^*(x) \in Ls\{f_n^*(x)\} \subset LsD_n^*(x) \subset D^*(x).$$

Por lo tanto f_n^* es un maximizador de J y se obtiene

$$J(x) = \mathscr{T}_{f^*}^n J(x) \leq \mathscr{T}_{f^*}^n 0(x) + \mathscr{T}_\circ^n\delta(x)$$

para todo $n \in \mathbb{N}$. Tomando el límite $n \to \infty$ se tiene $J(x) \leq J_{f^*}(x) \leq J_\infty(x)$. Así, se obtiene $J(x) = J_\infty(x)$ y $D^*(x) = D_\infty(x)$, para todo $x \in E$, y la política estacionaria $(f^*, f^*, \ldots)$ es óptima. Además, se obtiene para $n \geq m$

$$J_n \leq J_m + \mathscr{T}_\circ^m\delta \leq J_m + c\mathscr{T}_\circ^n b,$$

para algún $c > 0$. Puesto que J_m y $\mathscr{T}_\circ^m b$ son semicontinuos superiormente por el Teorema 1.4.1, el Lema A.1.2 implica que J_∞ es semicontinua superiormente.

□

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 45 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 46 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Corolario 2.2.1. Suponer que el Modelo de Decisión de Markov tiene una función b de acotación superior semicontinua superiormente con $\beta\alpha_b < 1$. Si las condiciones (i), (ii) y (iii) del Teorema 2.2.1 se satisfacen, entonces todas las afirmaciones a)-d) del Teorema 2.2.1 son validas.

La demostración se sigue del Teorema 2.2.1, ya que $\beta\alpha_b < 1$ implica $\| \delta \|_b < \infty$ y $\mathscr{T}_\circ^n b \to 0$ cuando $n \to \infty$. La condición de convergencia en el Corolario 2.2.1 es cierta para el *caso descontado* (es decir, r está acotada por arriba y $\beta < 1$) y también para el *caso negativo* (es decir, $r \leq 0$).

Teorema 2.2.2. Suponer que existe una función de acotación superior b con $\mathscr{T}_\circ^n b \to 0$ cuando $n \to \infty$, $\| \delta \|_b < \infty$ y que se satisfacen las siguientes condiciones:

(i) $D(x)$ es compacto para todo $x \in E$,

(ii) $a \mapsto \int v(x')Q(dx'|x,a)$ es semicontinua superiormente para toda $v \in \mathbb{B}_b^+$ y para todo $x \in E$,

(iii) $a \mapsto r(x,a)$ es semicontinua superiormente para todo $x \in E$.

Entonces se cumple:

a) $J_\infty \in \mathbb{B}_b^+$, $J_\infty = \mathscr{T} J_\infty$ y $J_\infty = J$ **(Iteración de Valor)**.

b) $\emptyset \neq LsD_n^*(x) \subset D_\infty^*(x)$ para todo $x \in E$ **(Iteración de Política)**.

c) Existe una $f^* \in F$ con $f^*(x) \in LsD_n^*(x)$ para todo $x \in E$, y la política estacionaria $(f^*, f^*, \ldots)$ es óptima.

Demostración. Notar que las v_n también son semicontinuas superiormente pero aquí no se necesita mostrar que J es semicontinua superiormente. □

El teorema anterior da condiciones que garantizan que se satisface (SE) para el conjunto $\mathbb{B}_b^+$. Es posible derivar otro conjunto de condiciones involucrando funciones continuas (Usando el Teorema 1.4.2.), las cuales implican la iteración de valores y de políticas pero aquí se omiten.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 46 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 47 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

2.3. Modelos de Decisión de Markov Restringidos

Una situación ventajosa e importante surge cuando el operador $\mathscr{T}$ es restringido. Aquí se asume que el Modelo de Decisión de Markov tiene una función de acotación en lugar de una función de acotación superior la cual, hasta ahora, se ha considerado.

Definición 2.3.1. Una función medible $b : E \to \mathbb{R}_+$ es llamada una *función de acotación* para el Modelo de Decisión de Markov si existen $c_r, \alpha_b \in \mathbb{R}_+$ tales que

(i) $|r(x,a)| \leq c_r b(x)$ para todo $(x,a) \in D$.

(ii) $\int b(x')Q(dx'|x,a) \leq \alpha_b b(x)$ para todo $(x,a) \in D$.

Si b es una función de acotación entonces $(\mathbb{B}_b, \|\cdot\|_b)$ es un espacio de Banach. La norma del supremo ponderada $\|\cdot\|_b$ se ha definido en la Sección 1.4. Sólo para recordar esta definición, aquí para $v \in \mathbb{M}(E)$

$$\| v \|_b := \sup_{x \in E} \frac{| v(x) |}{b(x)}.$$

Los Modelos de Decisión de Markov con una función de acotación b y $\beta\alpha_b < 1$ son llamados *restringidos*. Se verá en el Lema 2.3.1 que $\beta\alpha_b < 1$ es el módulo de $\mathscr{T}$.

Observación 2.3.1. *a)* Si r está acotada, entonces $b \equiv 1$ es una función de acotación. Si además $\beta < 1$, entonces el Modelo de Decisión de Markov es restringido (el *caso descontado* clásico).

b) Para cualquier Modelo de Decisión de Markov restringido las suposiciones (SC) y (SI) se satisfacen, puesto que $\delta \in \mathbb{B}_b$ y existe una constante $c > 0$ con

$$\lim_{n\to\infty} \sup_{\pi} \mathbb{E}_x^{\pi} \left[\sum_{k=n}^{\infty} \beta^k r^+\big(X_k, f_k(X_k)\big) \right] \leq c \lim_{n\to\infty} (\beta\alpha_b)^n b(x) = 0.$$

◊

Lema 2.3.1. Suponer que el Modelo de Decisión de Markov tiene una función de acotación b y sea $f \in F$.

a) Para $v, w \in \mathbb{B}_b$ se cumple:

$$\| \mathscr{T}_f v - \mathscr{T}_f w \|_b \leq \beta\alpha_b \| v - w \|_b,$$

$$\| \mathscr{T} v - \mathscr{T} w \| \leq \beta\alpha_b \| v - w \|_b .$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 47 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 48 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

b) Sea $\beta\alpha_b < 1$. Entonces $J_f = \lim_{n\to\infty} \mathscr{T}_f^n g$ para toda $g \in \mathbb{B}_b$, y J_f es el único punto fijo de $\mathscr{T}_f$ en $\mathbb{B}_b$.

Demostración.

a) Para $f \in F$ se obtiene

$$\begin{aligned}\mathscr{T}_f v(x) - \mathscr{T}_f w(x) &\leq \beta \sup_{a\in D(x)} \int \Big(v(x') - w(x')\Big) Q(dx|x,a) \\ &\leq \beta ||v-w||_b \sup_{a\in D(x)} \int b(x') Q(dx'|x,a).\end{aligned}$$

Al intercambiar v y w se concluye

$$\mathscr{T}_f w(x) - \mathscr{T}_f v(x) \leq \beta ||v-w||_b \sup_{a\in D(x)} \int b(x') Q(dx'|x,a).$$

Tomando la norma del supremo ponderada resulta

$$||\mathscr{T}_f v - \mathscr{T}_f w||_b \leq \beta\alpha_b ||v-w||_b.$$

Para el segundo enunciado notar que para funciones $g, h : M \to \mathbb{R}$ se cumple

$$\sup_{x\in M} g(x) - \sup_{x\in M} h(x) \leq \sup_{x\in M} \Big(g(x) - h(x)\Big),$$

luego

$$\mathscr{T} w(x) - \mathscr{T} v(x) \leq \beta \sup_{a\in D(x)} \int \Big(v(x') - w(x')\Big) Q(dx'|x,a).$$

Procediendo de la misma manera como para $\mathscr{T}_f$ se deriva la segunda desigualdad.

b) En vista de la parte a), el operador $\mathscr{T}_f$ está restringido en $\mathbb{B}_b$, y los enunciados se siguen del teorema del punto fijo de Banach.

□

El siguiente teorema es una versión ligeramente fuerte del Teorema de Verificación 2.1.2 para los Modelos de Decisión de Markov restringidos.

Teorema 2.3.1. *(Teorema de Verificación).* Sea b una función de acotación, $\beta\alpha_b < 1$ y $v \in \mathbb{B}_b$ un punto fijo de $\mathscr{T} : \mathbb{B}_b \to \mathbb{B}_b$. Si f^* es un maximizador de v, entonces $v = J_\infty$ y $(f^*, f^*, \ldots)$ es una política óptima estacionaria.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 48 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 49 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Demostración. Primero se obtiene $v = \mathscr{T}_{f^*}^n v \leq \mathscr{T}_{f^*}^n 0 + \mathscr{T}_{\circ}^n \delta$ de lo cual se concluye que $v = J_{f^*}$, cuando $n \to \infty$. De otra manera, puesto que por el Teorema 1.3.2 y el teorema de punto fijo de Banach $J \leq \lim_{n\to\infty} \mathscr{T}^n 0 = v$, se obtiene completamente que

$$J_\infty \leq J \leq v \leq J_{f^*} \leq J_\infty.$$

□

A continuación se da el resultado importante para los Modelos de Decisión de Markov restringidos.

Teorema 2.3.2. *(Teorema de Estructura).* Sea b una función de acotación superior y $\beta\alpha_b < 1$. Si existe un subconjunto cerrado $\mathbb{M} \subset \mathbb{B}_b$ y un conjunto $\Delta \subset F$ tales que:

(i) $0 \in \mathbb{M}$,

(ii) $\mathscr{T} : \mathbb{M} \to \mathbb{M}$,

(iii) para todo $v \in \mathbb{M}$ existe un maximizador $f \in \Delta$ de v,

entonces se cumple:

a) $J_\infty \in \mathbb{M}$, $J_\infty = \mathscr{T} J_\infty$ y $J_\infty = J$ **(Iteración de Valor)**.

b) J_∞ es el único punto fijo de $\mathscr{T}$ en $\mathbb{M}$.

c) J_∞ es la función r-superarmonica más pequeña $v \in \mathbb{M}$, es decir, J_∞ es la función más pequeña $v \in \mathbb{M}$ con $v \geq \mathscr{T} v$.

d) Sea $h \in \mathbb{M}$. Entonces

$$\| J_\infty - \mathscr{T}^n h \|_b \geq \frac{(\beta\alpha_b)^n}{1-\beta\alpha_b} \| \mathscr{T} h - h \|_b .$$

e) Existe un maximizador $f \in \Delta$ de J_∞, y para cada maximizador f^* de J_∞ define una política óptima estacionaria $(f^*, f^*, ...)$.

Demostración. Por el teorema de punto fijo de Banach (Teorema A.3.1), existe una función $v \in \mathbb{M}$ con

$$v = \mathscr{T} v \quad \text{y} \quad v = \lim_{n\to\infty} \mathscr{T}^n 0.$$

Así, todas las suposiciones del Teorema 2.1.3 se satisfacen, es decir, se ha demostrado a) y e). La parte b) y d) se siguen directamente del teorema de punto fijo de Banach. Para c),

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 49 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 50 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

notar que si $f \in \mathbb{M}$ es una solución de $v \geq \mathscr{T} v$ entonces por iteración se obtiene

$$v \geq \mathscr{T}^n v \to J_\infty,$$

cuando $n \to \infty$, de donde se sigue que $v \geq J_\infty$.

□

Asumir que E y A son espacios de Borel y D es un subconjunto de Borel de $E \times A$.

Teorema 2.3.3. Sea b una función de acotación superior y $\beta \alpha_b < 1$. Si se satisfacen las siguientes condiciones:

(i) $D(x)$ es compacto para todo $x \in E$ y $x \mapsto D(x)$ es continua,

(ii) $(x,a) \mapsto \int v(x')Q(dx'|x,a)$ es continua para toda $v \in \mathbb{B}_b$ continua,

(iii) $(x,a) \mapsto r(x,a)$ es continua,

entonces se cumple:

a) J_∞ es continua, $J_\infty \in \mathbb{B}_b^+$ y $J_\infty = J$ **(Iteración de Valor)**.

b) J_∞ es el único punto fijo continuo de $\mathscr{T}$ en $\mathbb{B}_b$.

c) $\emptyset \neq LsD_n^*(x) \subset D_\infty^*(x)$ para todo $x \in E$ **(Iteración de Política)**.

d) Existe una $f^* \in F$ con $f^*(x) \in LsD_n^*(x)$ para todo $x \in E$, y la política estacionaria $(f^*, f^*, \ldots)$ es óptima.

Demostración. El Teorema 1.4.2 implica que (i)-(iii) del Teorema 2.3.2 se satisfacen con el conjunto $\mathbb{M} := \{v \in \mathbb{B}_b | v \text{ es continua }\}$. Falta demostrar que $\mathbb{M}$ es cerrado. Suponer que $(v_n) \subset \mathbb{M}$ con $\lim_{n\to\infty} \| v_n - v \|_b = 0$. Por definición de $\| \cdot \|_b$ se sigue que $\frac{v_n}{b}$ converge uniformemente a $\frac{v}{b}$. Así, puesto que las $\frac{v_n}{b}$ son continuas por suposición, el límite $\frac{v}{b}$ es continuo lo cual implica que $\mathbb{M}$ es cerrado. La iteración de política se sigue como en el Teorema 2.2.1.

□

Universidad Juárez Autónoma de Tabasco. México

iThenticate Página 50 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco.
México.

iThenticate Página 52 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México

Capítulo 3

Los Mercados Financieros

En esta sección se introducen los mercados financieros, los cuales aparecen en nuestras aplicaciones. Se presenta un mercado financiero a tiempo discreto. Un ejemplo importante es el modelo binomial (ver [2], [7] y [24]).

3.1. Dinámica de Activos y Estrategias de Portafolios

Se asume que los precios de los activos están monitoreados a tiempo discreto. Suponer que el tiempo es dividido en periodos de tiempo de longitud Δt y $t_n = n\Delta t$. La forma más común de un precio de un activo es un modelo multiplicativo, es decir, si S_n es el precio al tiempo $t_n > 0$ entonces

$$S_{n+1} = S_n \tilde{R}_{n+1}.$$

La variable aleatoria positiva $\tilde{R}_{n+1}$ define el cambio de precio relativo S_{n+1}/S_n entre el tiempo t_n y t_{n+1}. Para un bono sin riesgo, el cambio de precio relativo S^0_{n+1}/S^0_n es elegido para ser $1 + i_{n+1}$ con la tasa de interés $i_{n+1} \in \mathbb{R}_+$.

Existen dos casos especiales importantes del modelo multiplicativo. El primero es el *modelo binomial* o *modelo de Cox-Ross-Rubinstein*. Aquí se supone que existe un bono con $i_{n+1} \equiv i$ y una acción con cambios de precios relativos, los cuales son independientes e idénticamente distribuidos y pueden tomar dos valores: ya sea $\mathbf{u} > 0$ para subir o $\mathbf{d} > 0$ para bajar, donde se asume que $\mathbf{d} < \mathbf{u}$. Las probabilidades para subir y para bajar son p y $1 - p$ respectivamente, es decir,

iThenticate Página 52 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 53 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

$$\mathbb{P}(\tilde{R}_n = \mathbf{u}) = p, \quad \mathbb{P}(\tilde{R}_n = \mathbf{d}) = 1 - p.$$

La evolución del precio de la acción se muestra en la Figura 1. Si se quiere modelar d activos simultáneamente un *modelo multinomial* es adecuado, es decir, $\tilde{R}_n$ puede tomar valores en $\{\mathbf{u}, \mathbf{d}\}^d = \{z_1, \ldots, z_{2^d}\}$ y

$$\mathbb{P}(\tilde{R}_n = z_k) = p_k, \quad k = 1, \ldots, 2^d$$

con $\sum_{k=1}^{2^d} p_k = 1$, $p_k \geq 0$.

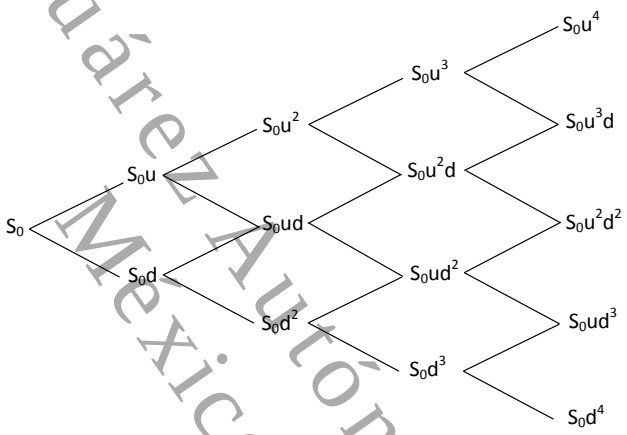


Figura 3.1: Evolución del precio de un activo en el modelo binomial

El segundo caso especial importante es una discretización del modelo de Black-Scholes-Merton. En el modelo de mercado de Black-Scholes-Merton estándar se asume que el precio de la acción evoluciona de acuerdo a

$$dS_t = S_t(\mu dt + \sigma dW_t)$$

donde (W_t) es un proceso de Wiener, $\mu \in \mathbb{R}$ y $\sigma > 0$ son parámetros dados. Así, si se toma un paso del tiempo Δt se obtiene

$$S_{n+1} = S_n exp\left((\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma W_{\Delta t}\right).$$

En este caso $\tilde{R}_n$ tiene una *distribución lognormal* o $\log\tilde{R}_n$ está distribuida normalmente. Por lo tanto la versión multivariada es $\log\tilde{R}_n \sim \mathrm{N}(a, \sum)$, es decir, $\log\tilde{R}_n$ tiene una distribución normal multivariada (el *log* es tomado componente a componente). Notar que en ambos casos especiales se tiene $\tilde{R}_n > 0$, y de este modo se garantiza que el precio de las acciones se mantendrá positivo con probabilidad uno.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 53 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 54 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

En lo que sigue se considera un *mercado financiero de N periodos* con d activos con riesgo y un bono sin riesgo. Asumir que todas las variables aleatorias están definidas en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ con una filtración $(\mathscr{F}_n)$ y $\mathscr{F}_0 := \{\emptyset, \Omega\}$. El mercado financiero está dado por:

- Un bono sin riesgo con $S_0^0 \equiv 1$ y

$$S_{n+1}^0 := S_n^0(1+i_{n+1}), \quad n = 0,1,...,N-1$$

donde i_{n+1} es la tasa de interés determinista para el periodo de tiempo $[n, n+1)$. Si la tasa de interés es constante, es decir, $i_n \equiv i$, entonces $S_n^0 = (1+i)^n$.

- Existen d activos con riesgo y el proceso del precio del activo k está dado por $S_0^k = s_0^k$ y

$$S_{n+1}^k = S_n^k \tilde{R}_{n+1}^k, \quad n = 0,1,...,N-1.$$

Los procesos (S_n^k) se asumen adaptados con respecto a la filtración $(\mathscr{F}_n)$ para todo k. Además, se supone que $\tilde{R}_{n+1}^k > 0$ $\mathbb{P}$-c.s. para todo k y n, y que s_0^k es determinístico. $\tilde{R}_{n+1}^k$ es el cambio de precio relativo en el intervalo de tiempo $[n, n+1)$ para el activo k.

En lo que sigue, se denota por $S_n := (S_n^1, ..., S_n^d)$, $\tilde{R}_n := (\tilde{R}_n^1, ..., \tilde{R}_n^d)$ y $\mathscr{F}_n^S := \sigma(S_0, ..., S_n)$. Puesto que (S_n) es $(\mathscr{F}_n)$-adaptado se cumple: $\mathscr{F}_n^S \subset \mathscr{F}_n$ para $n = 0,1,...,N$.

Definición 3.1.1. Un *portafolio* o una *estrategia de inversión* es un proceso estocástico $\phi = (\phi_n^0, \phi_n)$ $(\mathscr{F}_n)$-adaptado, donde $\phi_n^0 \in \mathbb{R}$ y $\phi_n = (\phi_n^1, ..., \phi_n^d) \in \mathbb{R}^d$ para $n = 0,1,...,N-1$. La cantidad ϕ_n^k denota el monto de dinero que es invertido en el activo k durante el intervalo de tiempo $[n, n+1)$.

Observación 3.1.1. Notar que en general no se realiza restricción en el signo de ϕ_n^k. En particular, a ϕ_n^k se le permite ser negativo, en tal caso $k = 0$ implica que un préstamo se ha tomado y que la tasa de interés es la misma para el endeudamiento y prestar. En el caso $\phi_n^k < 0$, para $k \neq 0$, esto corresponde a la venta corta del activo k.

El vector (ϕ_0^0, ϕ_0) es llamado el *portafolio inicial* del inversionista. El valor del portafolio inicial está dado por:

$$X_0 := \sum_{k=0}^{d} \phi_0^k = \phi_0^0 + \phi_0 \cdot e,$$

donde $x \cdot y = \sum_{k=1}^{d} x_k y_k$ denota el producto interno de los vectores $x, y \in \mathbb{R}^d$ y $e := (1, ..., 1) \in \mathbb{R}^d$.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 54 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 55 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Sea ϕ una estrategia de portafolio y denotar por X_{n-} el valor del portafolio al tiempo n antes de invertir. Entonces

$$X_n := X_{n-} := \sum_{k=0}^{d} \phi_{n-1}^k \tilde{R}_n^k = \phi_{n-1}^0(1+i_n) + \phi_{n-1} \cdot \tilde{R}_n.$$

El valor del portafolio al tiempo n después de invertir está dado por

$$X_{n+} := \sum_{k=0}^{d} \phi_n^k = \phi_n^0 + \phi_n \cdot e.$$

Se escribe X_n^ϕ cuando se quiera tomar explicitamente la dependencia en la estrategia de portafolio ϕ.

Definición 3.1.2. Una estrategia de portafolio ϕ es llamada *autofinanciada* si

$$X_{n-}^\phi := X_{n+}^\phi, \quad \mathbb{P}-c.s.$$

para todo $n = 1, ..., N-1$, es decir, la riqueza presente es simplemente reasignada a los activos.

Considerar estrategias de portafolios autofinanciadas. La condición autofinanciada implica que las siguientes equivalencias se cumplen para todo $n = 1, ..., N-1$:

$$X_n := X_{n+}, \qquad \Longleftrightarrow \qquad \phi_n^0 + \phi_n \cdot e = \phi_{n-1}^0(1+i_n) + \phi_{n-1} \cdot \tilde{R}_n.$$

Esta ecuación se puede usar para derivar una fórmula recursiva importante para la evolución de la riqueza. Esta ecuación recursiva se deriva de la siguiente manera (la condición de autofinanciada es aplicada en la última igualdad):

$$\begin{aligned} X_{n+1} &= X_0 + \sum_{t=1}^{n+1} (X_t - X_{t-1}) \\ &= X_0 + \sum_{t=1}^{n+1} \left(\phi_t^0 - \phi_{t-1}^0 + \phi_t \cdot e - \phi_{t-1} \cdot e \right) \\ &= X_0 + \sum_{t=1}^{n+1} \left(\phi_{t-1}^0 i_t + \phi_{t-1} \cdot (\tilde{R}_t - e) \right). \end{aligned}$$

Puesto que $\phi_n^0 = X_n - \phi_n \cdot e$, se puede eliminar ϕ_n^0 para obtener

$$\begin{aligned} X_{n+1} &= X_n + (\phi_n^0 i_{n+1} + \phi_n \cdot (\tilde{R}_{n+1} - e)) \\ &= X_n(1+i_{n+1}) + \sum_{k=1}^{d} \phi_n^k \left(\tilde{R}_{n+1}^k - 1 - i_{n+1} \right). \end{aligned}$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 55 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 56 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Cuando se introduce el *proceso de riesgo relativo* (R_n), definido por $R_n := (R_n^1, ..., R_n^d)$, y

$$R_n^k := \frac{\tilde{R}_n^k}{1+i_n} - 1, \quad k = 1, ..., d,$$

se obtiene la fórmula recursiva

$$X_{n+1} = (1+i_{n+1})\Big(X_n + \phi_n \cdot R_{n+1}\Big). \tag{3.1.1}$$

La ventaja de la ecuación (3.1.1) es que sólo la inversión en las acciones entra en la ecuación y (ϕ_n) puede completarse a una estrategia de portafolio autofinanciada. Como es usual, se tienen que excluir oportunidades de arbitraje en el mercado financiero. Una oportunidad de arbitraje es definida formalmente de la siguiente manera (ver [1] y [6]).

Definición 3.1.3. Una *oportunidad de arbitraje* es una estrategia de portafolio autofinanciada $\phi = (\phi_n^0, \phi_n)$ con la siguiente propiedad: $X_0^\phi = 0$ y

$$\mathbb{P}(X_N^\phi \geq 0) = 1, \qquad y \qquad \mathbb{P}(X_N^\phi > 0) > 0.$$

Ligeramente hablando, una oportunidad de arbitraje es una estrategia de inversión sin riesgo con la posibilidad de una ganancia. En mercados financieros reales tales oportunidades algunas veces existen pero una vez detectadas los inversionistas pronto las disiparán. El siguiente teorema caracteriza la ausencia de oportunidades de arbitraje. Muestra que el mercado financiero es libre de oportunidades de arbitraje si y sólo si no existe localmente oportunidad de arbitraje. Esta propiedad es importante cuando se reducen problemas de optimización multiperiodo a problemas de un periodo.

Teorema 3.1.1. Considerar un mercado financiero de N periodos. Los siguientes dos enunciados son equivalentes:

a) No existen oportunidades de arbitraje

b) Para $n = 0, 1, ..., N-1$ y para todo $\phi_n \in \mathbb{R}^d$, $\mathscr{F}_n$ – medible, se cumple:

$$\phi_n \cdot R_{n+1} \geq 0 \quad \mathbb{P}-\text{c.s.} \qquad \Rightarrow \qquad \phi_n \cdot R_{n+1} = 0 \quad \mathbb{P}-\text{c.s.}$$

Demostración. Primero suponer que existe un $n \in \{1, ..., N\}$ y un $\varphi \in \mathbb{R}^d$, $\mathscr{F}_{n-1}$–medible, tal que

$$\varphi \cdot R_n \geq 0 \, \mathbb{P}-\text{c.s} \quad \text{y} \quad \mathbb{P}(\varphi \cdot R_n > 0) > 0.$$

Entonces se puede construir una estrategia arbitraria ϕ como sigue:

$$\phi_k := \begin{cases} \varphi & \text{si } k = n-1, \\ 0 & \text{de otro modo.} \end{cases}$$

Universidad Juárez Autónoma de Tabasco. México

iThenticate Página 56 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 57 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Entonces $X_{n-1}^{\phi} = \phi_{n-1}^{0} + \phi_{n-1} \cdot e = 0$ y $X_n^{\phi} = (1+i_n)(\varphi \cdot R_n)$. Luego, $\mathbb{P}(X_n^{\phi} \geq 0) = 1$ y $\mathbb{P}(X_n^{\phi} > 0) > 0$. Puesto que $X_N^{\phi} = (1+i_{n+1}) \cdots (1+i_N) X_n^{\phi}$, la estrategia ϕ es una estrategia de arbitraje.
Con el fin de mostrar la implicación inversa, suponer que existe una estrategia de arbitraje $\phi = (\phi_n^0, \phi_n)$. Sea

$$m := \text{mín}\{\, n \mid \mathbb{P}(X_n^{\phi} \geq 0) = 1 \text{ y } \mathbb{P}(X_n^{\phi} > 0) > 0\}.$$

Entonces $m \geq N$ por suposición. Definir $\varphi := 1_{[X_{m-1}^{\phi} \leq 0]} \phi_{m-1}$. Entonces φ es $\mathscr{F}_{m-1}-$medible y se puede considerar la estrategia de portafolio ψ la cual invierte las mismas cantidades como ϕ desde el tiempo 0 a $m-2$ y al tiempo $m-1$ la cantidad ϕ en los activos con riesgo. Se tiene

$$X_m^{\psi} = (1+i_m)(X_{m-1}^{\phi} + \varphi \cdot R_m).$$

Puesto que $\mathbb{P}(X_m^{\psi} \geq 0) = 1$ se sigue que $\varphi \cdot R_m \geq 0$, $\mathbb{P}-$c.s. Por definición de m se tiene $\mathbb{P}(X_{m-1}^{\phi} = 0) = 1$ ó $\mathbb{P}(X_{m-1}^{\phi} < 0) > 0$. En el primer caso se obtiene

$$\mathbb{P}(\varphi \cdot R_m > 0) = \mathbb{P}(X_m^{\phi} > 0) > 0$$

y en el segundo caso se tiene $\mathbb{P}(\varphi \cdot R_m > 0) \geq \mathbb{P}(X_{m-1}^{\phi} < 0) > 0$. Con todo esto se ha mostrado $\mathbb{P}(\varphi \cdot R_m \geq 0) = 1$ y $\mathbb{P}(\varphi \cdot R_m > 0) > 0$, lo cual implica el enunciado.

□

Observación 3.1.2. Cuando las variables aleatorias $R_1, \ldots, R_N$ son independientes y $\mathscr{F}_n := \mathscr{F}_n^S := \sigma(S_0, \ldots S_n) = \sigma(R_1, \ldots, R_n)$ se obtiene la siguiente caracterización de la ausencia de arbitraje: Para toda $n = 0, 1, \ldots, N-1$ y para todo $\phi \in \mathbb{R}^d$ se cumple:

$$\phi \cdot R_{n+1} \geq 0 \quad \mathbb{P}-c.s. \qquad \Rightarrow \qquad \phi \cdot R_{n+1} = 0 \quad \mathbb{P}-c.s. \tag{3.1.2}$$

Notar que aquí ϕ es determinístico. No es difícil ver que este enunciado es equivalente al Teorema 3.1.1 parte b). La demostración es como sigue: El Teorema 3.1.1 parte b) implica (3.1.2) ya que cada $\phi \in \mathbb{R}^d$ es $\mathscr{F}_n$-medible. Ahora suponer que (3.1.2) se cumple y que $\phi \in \mathbb{R}^d$ es $\mathscr{F}_n$-medible tal que

$$\phi_n \cdot R_{n+1} \geq 0 \quad c.s. \quad \text{y} \quad \mathbb{P}(\phi_n \cdot R_{n+1} > 0) > 0.$$

Definir $\tilde{\phi}_n := \mathbb{E}[\phi_n] \in \mathbb{R}^d$. Entonces se obtiene $\mathbb{E}[\phi_n \cdot R_{n+1} \mid R_{n+1}] = \tilde{\phi}_n \cdot R_{n+1} \geq 0$ $\mathbb{P}$-c.s. y $\mathbb{P}(\tilde{\phi}_n \cdot R_{n+1} > 0) > 0$, lo cual es una contradicción.

◊

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 57 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 58 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Ejemplo 3.1.1. En el mercado financiero binomial, una condición necesaria y suficiente para la falta de oportunidades de arbitraje es que los parámetros del modelo satisfagan

$$\mathbf{d} < 1+i < \mathbf{u}.$$

Esto puede ser visto como sigue. Primero notar que ϕR_1 puede tomar los valores

$$\phi(\frac{\mathbf{d}}{1+i}-1) \qquad \text{y} \qquad \phi(\frac{\mathbf{u}}{1+i}-1).$$

Así, si $\mathbf{d} < 1+i < \mathbf{u}$, entonces

$$\phi R_1 \gtrless 0 \quad \Longrightarrow \quad \phi \equiv 0$$

y por lo tanto se cumple que $\phi R_1 = 0$. Si $\mathbf{u} > \mathbf{d} \geqslant 1+i$, se obtiene para todo $\phi > 0$ que

$$\phi R_1 > 0 \quad \text{y} \quad \mathbb{P}(\phi R_1 > 0) > 0.$$

Similarmente, si $1+i \geqslant \mathbf{u} > \mathbf{d}$ se tiene para todo $\phi < 0$ que $\phi R_1 > 0$ y $\mathbb{P}(\phi R_1 > 0) > 0$. Todo esto junto demuestra el enunciado.

♦

3.2. Funciones de Utilidad y Utilidad Esperada

En general una función de utilidad es un mapeo $U : domU \to \mathbb{R}$ el cual es aplicado a un resultado variable de una inversión. En particular, si se tienen dos variables aleatorias X y Y, podemos compararlas mediante la comparación de $\mathbb{E}U(X)$ con $\mathbb{E}U(Y)$ donde el valor más grande es preferido. Una U razonable debe ser creciente, lo cual significa que más dinero es mejor que menos dinero. Muchas veces también es asumido que U es cóncava, lo que significa que para un individuo la utilidad marginal de riqueza es decreciente.

Definición 3.2.1. Una función $U : domU \to \mathbb{R}$ es llamada *función de utilidad* si U es estrictamente creciente, estrictamente cóncava y continua en $domU$.

Si $domU$ es un intervalo abierto, entonces la concavidad de U implica que U también es continua. Si $domU = [0,\infty)$, entonces U es continua en $(0,\infty)$ y suponer que U también es continua en 0 por la derecha.
Si un inversionista elige $U(x) = x$ (lo cual no es una función de utilidad por definición) entonces se dice ser *neutral al riesgo* ya que no se hace ninguna consideración por el riesgo. Las siguientes funciones de utilidad son estándar:

a) *Utilidad logarítmica*. Aquí se tiene $U(x) = \log(x)$ y $domU = (0,\infty)$.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 58 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 59 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

b) *Utilidad potencia*. Aquí se tiene $U(x) = \frac{1}{\gamma}x^{\gamma}$ y $domU = [0,\infty)$ cuando $0 < \gamma < 1$. Si $\gamma < 0$ tenemos $domU = (0,\infty)$.

c) *Utilidad exponencial*. Aquí se tiene $U(x) = -\frac{1}{\gamma}e^{-\gamma x}$ con $\gamma > 0$ y $domU = \mathbb{R}$.

d) *Utilidad cuadrática*. Aquí se tiene $U(x) = x - \gamma x^2$ para $\gamma > 0$ y $domU = (-\infty, (2\gamma)^{-1})$. Esta función sólo es creciente para $x < (2\gamma)^{-1}$.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 59 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 60 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

Capítulo 4

Aplicaciones a Finanzas

4.1. Problemas de Optimización Financiera

La teoría de Procesos de Decisión de Markov, que ha sido presentada en el Capítulo 1, ahora será aplicada a algunos problemas seleccionados de optimización dinámica en finanzas. El modelo básico subyacente es el mercado financiero del Capítulo 3. Asumir que los inversionistas son pequeños y no pueden influir en el proceso de precios de los activos.

4.1.1. El Problema de Optimización de un Periodo

Se verá el problema de maximización de utilidad de un periodo (ver [6]). Para facilitar la notación omitir el índice de tiempo en i_1, $\tilde{R}_1$ y R_1 y escribir i, $\tilde{R}$ y R. Suponer que se tiene un inversionista con función de utilidad $U : domU \to \mathbb{R}$ y riqueza inicial $x > 0$. Denotar por

$$\Delta(x) := \left\{ (\phi^0, \phi) \in \mathbb{R}^{1+d} | \phi^0(1+i) + \phi \cdot \tilde{R} \in domU \;\; c.s, \;\; \phi^0 + \phi \cdot e \leq x \right\}$$

el conjunto de estrategias de portafolios admisibles de un periodo. Una formulación sencilla del problema de maximización de utilidad de un periodo es

$$\sup_{\phi \in \Delta(x)} \mathbb{E}U(\phi^0(1+i) + \phi \cdot \tilde{R}). \tag{4.1.1}$$

Por lo tanto, el inversionista maximiza la utilidad esperada de su inversión bajo la restricción de que el retorno está en el dominio de la función de utilidad con probabilidad uno

iThenticate Página 60 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 61 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

y que se satisface el presupuesto restringido $\phi^0 + \phi \cdot e \leq x$. Con el fin de tener un modelo bien definido hacer la siguiente suposición de integrabilidad, donde $\| x \| := |x_1| + \cdots + |x_d|$ para $x \in \mathbb{R}^d$.

Suposición: $\mathbb{E} \| R \| < \infty$.

Obviamente $\mathbb{E} \| R \| < \infty$ es equivalente a $\mathbb{E} \| \tilde{R} \| < \infty$. Puesto que U es cóncava, la función de utilidad puede ser acotada por arriba mediante una función lineal afín $c_u(1+x)$ con $c_u \in \mathbb{R}_+$. Así la suposición implica que

$$\mathbb{E}U(\phi^0(1+i) + \phi \cdot \tilde{R}) \leq c_u(1 + \phi^0(1+i) + \phi \cdot \mathbb{E} \| \tilde{R} \|) < \infty$$

para todo $(\phi^0, \phi) \in \mathbb{R}^{1+d}$. La siguiente observación es simple, pero crucial. Suponer que (ϕ^0, ϕ) es un portafolio con $\phi^0 + \phi \cdot e < x$. Entonces se puede construir un nuevo portafolio invirtiendo la riqueza restante $x - \phi^0 - \phi \cdot e$ en el bono. Este nuevo portafolio se denota por $(\varphi^0, \varphi) := (x - \phi \cdot e, \phi)$. Puesto que las funciones de utilidad son estrictamente crecientes se obtiene que

$$\mathbb{E}U(\varphi^0(1+i) + \varphi \cdot \tilde{R}) > \mathbb{E}U(\phi^0(1+i) + \phi \cdot \tilde{R}).$$

En vista del problema (4.1.1) se puede concluir, sin pérdida de generalidad, que la restricción presupuestaria puede ser formulada como $\phi^0 + \phi \cdot e = x$ lo cual permite eliminar la componente del bono y obtener un problema no restringido. De las consideraciones anteriores se puede escribir

$$\phi^0(1+i) + \phi \cdot \tilde{R} = X_1^{\phi} = (1+i)(x + \phi \cdot R).$$

Por lo tanto, las inversiones admisibles de un periodo en los activos con riesgo están dadas por

$$D(x) := \left\{ a \in \mathbb{R}^d | (1+i)(x + a \cdot R) \in domU \quad \mathbb{P} - c.s. \right\}.$$

Para $a \in D(x)$, la cantidad invertida en el bono está dada por $x - a \cdot e$. Por lo tanto cuando se define

$$u(x,a) := \mathbb{E}U\big((1+i)(x + a \cdot R)\big)$$

una formulación equivalente de (4.1.1) es

$$v(x) := \sup_{a \in D(x)} u(x,a). \tag{4.1.2}$$

El siguiente resultado muestra que la ausencia de oportunidad de arbitraje es equivalente a la existencia de una solución óptima para el problema (4.1.2).

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 61 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 62 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Teorema 4.1.1. Sea U una función de utilidad con $domU = [0,\infty)$ o $domU = (0,\infty)$. Entonces se cumple:

a) No existen oportunidades de arbitraje si y sólo si existe una función medible $f^* : domU \to \mathbb{R}^d$ tal que

$$u\big(x, f^*(x)\big) = v(x), \quad x \in domU.$$

b) La función $v(x)$ es estrictamente creciente, estrictamente cóncava y continua en $domU$.

Observación 4.1.1. La solución óptima del Teorema 4.1.1 del problema de maximización de utilidad de un periodo es única si el mercado financiero es *no redundante*, es decir, no existe un activo que pueda ser perfectamente replicado mediante una combinación lineal de los otros activos.

$\Diamond$

Considerar que la función de utilidad en el problema (4.1.2) está dada por

$$U(x) = \log x, \quad \text{con} \quad domU = (0,\infty)$$

y una riqueza inicial $x > 0$. Puesto que será más conveniente trabajar con fracciones de dinero invertido, esto es, considerar la transformación $a = \alpha x$, se define el conjunto de fracciones admisibles por

$$A := \Big\{ \alpha \in \mathbb{R}^d | (1+i)(1+\alpha \cdot R) > 0 \quad \mathbb{P}-c.s. \Big\},$$

en lugar de considerar

$$D(x) := \Big\{ a \in \mathbb{R}^d | (1+i)(x+a \cdot R) \in domU \quad \mathbb{P}-c.s. \Big\}.$$

Lo anterior se debe a que

$$(1+i)(x+\alpha x \cdot R) \in (0,\infty) \qquad \Leftrightarrow \qquad (1+\alpha \cdot R) > 0.$$

Ahora, el problema de optimización de un periodo (4.1.2) es:

$$v := \sup_{\alpha \in A} \mathbb{E} \log(1+\alpha \cdot R). \tag{4.1.3}$$

En el caso de invertir en una acción y que el procesos de precios siga el *modelo binomial*, el problema de optimización (4.1.2) se reduce a

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 62 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 63 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

$$\sup_{\alpha_0<\alpha<\alpha_1} p\log\Big(1+i\alpha(\mathbf{u}-1-i)\Big)+(1-p)\log\Big(1+i\alpha(\mathbf{d}-1-i)\Big)$$

donde

$$\alpha_0 := \frac{1+i}{1+i-\mathbf{u}} < 0 \quad \text{y} \quad \alpha_1 := \frac{1+i}{1+i-\mathbf{d}} > 0.$$

Así, la fracción óptima la cual es invertida en la acción es de la forma

$$\alpha^* := (1+i)\Big(\frac{p}{1+i-\mathbf{d}} - \frac{1-p}{\mathbf{u}-1-i}\Big). \tag{4.1.4}$$

Considerar una riqueza inicial de $x = 1000$, con una tasa de interés de $i = 0.04$. Suponer el proceso de precios de la acción sigue el *modelo binomial*, con $\mathbb{P}(\tilde{R} = \mathbf{1.05}) = .7$ y $\mathbb{P}(\tilde{R} = \mathbf{1.035}) = .3$. Entonces, $\alpha^* := 114.4$ y

$$\phi = (209.7528, 790.2472).$$

Ver Apéndice B (Código 1).

4.1.2. El Problema de Consumo-Inversión

Se considera el problema de consumo e inversión con horizonte de tiempo infinito (Ver Cap. 2).

Un inversionista desea asignar su riqueza actual x entre consumir la cantidad a e invertir la cantidad $x-a$. Asumir que los préstamos no son permitidos.

Al inicio de cada periodo $k = 0, 1, 2, \ldots$, puede decidir cuánto del capital consumir y cuánto invertir. La cantidad que se invierte está evaluada por una función de utilidad

$$U(x) := \frac{3}{0.75}x^{0.75}, \quad \text{para } x \geq 0.$$

Aquí, el estado x del sistema es el capital disponible. La acción $a = f(x)$ es la cantidad de dinero la cual es consumida. La función de recompensa está dada por

$$r(x,a) := U(x-a),$$

y el proceso de riqueza evoluciona de la siguiente manera

$$X_{n+1} = (X_n - f(X_n)) \cdot Z_n,$$

esto es, la riqueza al tiempo $n+1$ es proporcional a la cantidad invertida en el tiempo n, donde se supone que Z_n son variables aleatorias independientes e idénticamente distribuidas independientes de la riqueza inicial X_0 y además

$$Z_0 \sim N(1.05, 0.6) \quad \text{si } z \geq 0,$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 63 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 64 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

en cualquier otro caso, $Z_0 = 0$. Luego, $(a, x-a) = (a_n, x_n - a_n)$ es la estrategia de consumo-inversión, esto es, $x_n - a_n$ y a_n son $\mathscr{F}$-medibles, donde $0 \leq a \leq x$ para toda $n \in \mathbb{N}$. El objetivo es maximizar la recompensa total esperada descontada

$$J_{\infty\pi}(x) := \mathbb{E}_x^{\pi}\left[\sum_{k=0}^{\infty} \beta^k r\big(X_k, f_k(X_k)\big)\right] \tag{4.1.5}$$

sobre todas las reglas de inversión permitidas $\pi = (f_0, f_1, \ldots) \in F^{\infty}$, considerando el factor de descuento $\beta = 0.99$.

Se describe el problema como un Problema de Decisión de Markov de la siguiente manera.

- Tomar $E = A = [0, \infty)$ y $D(x) = [0, x]$, para toda $x \in X$;
- La ley de transición Q está definida como

$$Q(B|x,a) = \begin{cases} \int_B \frac{1}{a}\xi(y/a)dy & \text{si } a \neq 0, \\ I_B(0) & \text{si } a = 0, \end{cases}$$

  donde $(x,a) \in D$, $B \in \mathscr{B}(E)$, ξ es la densidad de Z_t y I_B es la función indicadora de $B \in \mathscr{B}(E)$;
- La función de recompensa $r : D \longrightarrow \mathbb{R}$ está dada por

$$r(x,a) = \begin{cases} \frac{3}{0.75}(x-a)^{0.75} & \text{si } x \neq a, \\ 0 & \text{si } x = a. \end{cases}$$

Notar que $D(x)$ es compacto y $r(x,a)$ es continua en $a \in D(X)$, para todo $x \in E$. También, observar que r no está acotada y que

$$\begin{aligned} \sup_{a\in[0,x]} r(x,a) &= \sup_{a\in[0,x)} r(x,a) \\ &= \sup_{a\in[0,x)} \frac{3}{0.75}(x-a)^{0.75} \\ &= \frac{3}{0.75}x^{0.75} \end{aligned}$$

si $x > 0$, y cero en otro caso. Por lo que

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 64 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 65 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

$$\sup_{a\in D(x)} r(x,a) = \begin{cases} \frac{3}{0.75}x^{0.75} & \text{si } x \neq a, \\ 0 & \text{si } x = a. \end{cases}$$

Luego,

$$\begin{aligned}
\mathscr{T}_\circ r(x,a) &= \sup_{a\in[0,x)} \beta \int r(x,a)Q(dx'|x,a) \\
&= \sup_{a\in[0,x)} 0.99 \int_{(0,\infty)} \frac{3}{0.75}(x)^{0.75} Q(x'|x,a) \\
&= \frac{(3)(0.99)}{0.75} \sup_{a\in[0,x)} \int_{(0,\infty)} (a\cdot s)^{0.75}\xi(s)ds \\
&= \frac{(3)(0.99)}{.75}\left(\int_{(0,\infty)} \frac{s^{0.75}}{0.6\sqrt{2\pi}} e^{-0.5((x-1.05)/0.6)^2} ds\right)\left(\sup_{a\in[0,x)} a^{0.75}\right) \\
&= (0.99)(1.00667)\left(\frac{3}{0.75}\right)x^{0.75} \\
&= (0.9966033)\left(\frac{3}{0.75}\right)x^{0.75}.
\end{aligned}$$

De manera similar, por inducción, se tiene

$$\mathscr{T}_\circ^n r(x,a) = (0.9966033)^n \left(\frac{3}{0.75}\right)x^{0.75}, \quad \text{para todo } n = 0,1,2,\ldots, \text{y } x > 0.$$

Entonces, para $x > 0$,

$$\sum_{n=0}^{\infty} \mathscr{T}_\circ^n r(x,a) = \left(\frac{3}{0.75}\right)x^{0.75} \sum_{n=0}^{\infty} (0.9966033)^n,$$

lo cual converge, esto es,

$$\lim_{n\to\infty} \mathscr{T}_\circ^n r(x,a) = 0.$$

Por lo tanto se tiene SI y SC.

El siguiente lema permite obtener las funciones de valor J_n y su demostración se puede encontrar en [17, Sección 3.6].

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 65 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 66 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Lema 4.1.1. Sea $\theta := [\beta m_\gamma]^{\frac{1}{\gamma-1}}$. Entonces

$$J_0(x) = 0 \qquad \text{para todo } x \in E,$$

$$J_n(x) = \begin{cases} \left[\theta^{n-1}\dfrac{1-\theta}{1-\theta^n}\right]^{\gamma-1} \dfrac{b}{\gamma}x^\gamma & \text{si } x > 0, \\ 0 & \text{si } x = 0, \end{cases} \tag{4.1.6}$$

para $n = 1, 2, \ldots$. Además, las reglas de decisión de f_n están dadas por

$$f_n(x) = \frac{1-\theta^{n-1}}{1-\theta^n}x \text{ para todo } x \in E,\ n = 1, 2, \ldots. \tag{4.1.7}$$

□

Por lo tanto, de (4.1.6) se tiene que la función de valor óptima J_∞ del problema de control en (4.1.5) es, para $x > 0$,

$$\begin{aligned} J_\infty(x) &= \exp\left[(\gamma-1)\ln\left(\frac{1-\theta}{-\theta}\right)\right]\frac{b}{\gamma}x^\gamma \\ &= \exp\left[(0.75-1)\ln\left(\frac{1-1.013703}{-1.013703}\right)\right]\frac{3}{0.75}x^{0.75} \\ &= 11.73097x^{0.75} \end{aligned}$$

y $J_\infty(0) = 0$, y la política óptima está dada por $\pi = (f^*, f^*, \ldots)$, con

$$f^*(x) = \lim_{n\to\infty} f_n(x) = \frac{x}{\theta} = \frac{x}{1.013703},$$

ver [15, Teorema 4.5].

En el caso particular de que se tenga una riqueza inicial $x = 10,000$, entonces

$$\pi = (9864.823, 9864.823, \ldots),$$

es decir, se debe consumir solo el 1.35 % de la riqueza.

Universidad Juárez Autónoma de Tabasco. México

iThenticate Página 66 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 67 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

4.2. Problemas de Paro con Horizonte Finito

Se considera un problema de paro estacionario de horizonte finito N. Notar que siempre se tiene que parar, es decir, si no se para antes del tiempo N se debe parar al tiempo N. Asumir que el proceso de estados está dado por un proceso de Markov (X_n) (estacionario) sobre un espacio de estados general E con una σ-álgebra $\mathscr{E}$ y con kernel de transición (Q^X). Todo el tiempo que el proceso no es parado, se obtiene una recompensa medible $\beta^n r(X_n)$ la cual depende de la etapa n, para un factor de descuento $\beta \in (0,1]$. Si el proceso es parado al tiempo n, la recompensa medible $\beta^n g$ es recibida. Como en cualquier problema de decisión se tiene que definir la información con respecto a la cuál se elige el tiempo de paro. También, se asume que la decisión sólo puede depender de la observación del proceso (X_n). Así, denotar por $(\mathscr{F}_n)$ la filtración generada por el proceso (X_n), es decir, $\mathscr{F}_n := \sigma(X_0,...,X_n)$.

Definición 4.2.1. Un tiempo aleatorio $\tau : \Omega \to \mathbb{N}_0 \cup \{\infty\}$ es un $(\mathscr{F}_n)$- tiempo de paro si para toda $n \in \mathbb{N}_0$

$$\{\tau \leq n\} \in \mathscr{F}_n.$$

Esta condición significa que se requiere observar el proceso (X_n) hasta el tiempo n, para poder decidir si τ ya ocurrió o no. En tiempo discreto la condición es equivalente a $\{\tau = n\} \in \mathscr{F}_n$, para toda $n \in \mathbb{N}_0$. En las aplicaciones, considerar la filtración generada por (X_n). Así, si se elige un tiempo de paro τ con $\mathbb{P}_x(\tau \leq N) = 1$, para todo $x \in E$, se obtiene la recompensa

$$R_\tau := \sum_{k=0}^{\tau-1} \beta^k r(X_k) + \beta^\tau g(X_\tau).$$

Con el fin de obtener un problema bien definido, se necesita la siguiente suposición general:

Suposición (B_N): Para $x \in E$

$$\sup_{n \leq \tau \leq N} \mathbb{E}_{nx}\left[\sum_{k=n}^{\tau-1} \beta^k r^+(X_k) + \beta^\tau g^+(X_{n\leq\tau})\right] < \infty, \quad 0 \leq n \leq N.$$

Cuando se escribe $\mathbb{E}_{nx}$ (o $\mathbb{E}_x$ para la forma corta de $\mathbb{E}_{0x}$), la esperanza es tomada con respecto a la medida de probabilidad inducida por el proceso de Markov (X_n) dado $X_n = x$. El problema es encontrar el valor

$$J_N^*(x) := \sup_{\tau \leq N} \mathbb{E}_x[R_\tau] \tag{4.2.1}$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 67 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 68 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

donde el supremo es tomado sobre todos los tiempos de paro τ con $\mathbb{P}_x(\tau \leq N) = 1$, para todo $x \in E$. Un tiempo de paro $\tau^* \leq N$ es *óptimo* si para todo $x \in E$

$$J_N^*(x) := \mathbb{E}_x[R_{\tau^*}].$$

Con el fin de establecer la conexión al problema de paro en (4.3.1), la siguiente relación entre tiempos de paro y políticas en el Modelo de Decisión de Markov es crucial:
Suponer que $\pi = (f_0, ..., f_{N-1})$ es una política arbitraria para el Modelo de Decisión de Markov y definir

$$\tau_\pi := \text{ínf}\{n \in \mathbb{N}_0 | f_n(X_n) = 1\} \wedge N.$$

El tiempo aleatorio τ_π es un tiempo $(\mathscr{F}_n)$ de paro porque

$$\{\tau_\pi = n\} = \{f_0(X_0) = 0, ..., f_{n-1}(X_{n-1}) = 0, f_n(X_n) = 1\} \in \mathscr{F}_n.$$

Por construcción está acotado por N. τ_π es llamado un *tiempo de paro de Markov* puesto que $\pi = (f_n)$ es Markoviana.
De otra manera, suponer que τ es un $(\mathscr{F}_n)$ tiempo de paro. Entonces, por definición existe una función medible $f_n : E^{n+1} \to \{0, 1\}$, para toda $n \in \mathbb{N}_0$, tal que

$$1_{[\tau=n]} = f_n(X_0, X_1, ..., X_n).$$

Así, τ se puede representar por una política dependiente de la historia $\pi = (f_0, ..., f_{N-1})$. En particular se cumple que

$$J_{0\pi}(x) = \mathbb{E}_x \left[\sum_{k=0}^{\tau_\pi - 1} \beta^k r(X_k) + \beta^{\tau_\pi} g(X_{\tau_\pi}) \right] = \mathbb{E}_x[R_{\tau_\pi}] = \mathbb{E}_x[R_\tau].$$

Del Teorema 1.2.1 ya se conoce que en Modelos de Decisión de Markov la política óptima puede ser encontrada entre las políticas de Markov. Por lo tanto se puede concluir que el tiempo óptimo de paro, si existe, puede encontrarse entre los tiempos de paro de Markov. Así, se sigue que $J_0 := \sup_\pi J_{0\pi}$ (el valor del modelo de Decisión de Markov) es el mismo que J_N^* (el valor del problema de paro) y que una política π^*, si existe, define un tiempo de paro óptimo. Además, esta discusión revela que la Suposición (B_N) es equivalente a la Suposición (A_N) para el correspondiente Modelo de Decisión de Markov. Con todo esto junto se tiene el siguiente teorema.

Teorema 4.2.1. Se cumple que

a) $J_0(x) = J_N^*(x)$ para todo $x \in E$.

b) Suponer que $\pi^* = (f_0^*, ..., f_{N-1}^*)$ es una política óptima para el Problema de Decisión de Markov y definir el tiempo de paro

$$\tau^* := \text{ínf}\{n \in \mathbb{N}_0 | f_n^*(X_n) = 1\} \wedge N,$$

entonces τ^* es óptimo para el problema de paro (4.2.1).

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 68 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 69 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

De manera análoga se puede mostrar que Jn es la recompensa máxima esperada para el problema de paro y el problema a resolver es

$$J_N(x) := \sup_{\tau \leq N} \mathbb{E}_x \left[\sum_{k=0}^{\tau-1} \beta^k r(X_k) + \beta^\tau g(X_\tau) \right], \quad x \in E,$$

donde el supremo es tomado sobre todos los tiempos de paro τ, los cuales satisfacen $\mathbb{P}_x(\tau \leq N) = 1$ para todo $x \in E$. Este problema también se puede formular como un Problema de Decisión de Markov (caso no estacionario).

Teorema 4.2.2. Suponer que un problema de paro estacionario con horizonte finito N está dado. Entonces se cumple:

a) $J_0 = g$ y $J_n = \mathscr{T} J_{n-1}$, para $n = 1, ..., N$ donde

$$\mathscr{T} v(x) = \text{máx} \left\{ g(x), r(x) + \beta \int v(x') Q^X(dx'|x) \right\}, \quad x \in E.$$

b) $g \leq J_n \leq J_{n+1}$, para toda $n \in \mathbb{N}_0$.

c) Definir $d_n(x) := g(x) - r(x) - \beta \int J_{n-1}(x') Q^X(dx'|x)$ para $n = 1, ..., N$. Entonces tenemos la siguiente recursión

$$d_{n+1}(x) = d_1(x) - \beta \int d_n^-(x') Q^X(dx'|x).$$

d) Considerar $S_n^* := \{x \in E | J_n(x) = g(x)\}$ y definir $f_n^*(x) := 1_{S_n^*}$. Entonces $S_0^* = E$, $S_n^* = \{x \in E | d_n(x) \geq 0\}$ y $S_{n+1}^* \subset S_n^*$. La política $(f_N^*, ..., f_1^*)$ es óptima y el tiempo de paro

$$\tau^* := \text{mín}\{n \in \{0, 1, ..., N\} | X_n \in S_{N-n}^*\}$$

es óptimo para el problema de paro (4.2.1).

La inclusión $S_{n+1}^* \subset S_n^*$ significa que la tendencia a parar es no decreciente conforme pasa el tiempo.

Demostración.

a) Se sigue del Teorema 1.5.2 (caso estacionario) como en el caso no estacionario.

b) Ya que por la parte a) J_n es el máximo de g y alguna otra función, el enunciado $J_n \geq g$ es claro. Puesto que $g = J_0 \leq J_1$ y debido al hecho de que el operador de recompensa máxima está preservando el orden (Lema 1.3.1c)) se sigue que $J_n \leq J_{n+1}$.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 69 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 70 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

c) Primero notar que

$$d_1(x) = g(x) - c(x) - \beta \int g(x')Q^X(dx'|x).$$

Por la parte a) y la definición de d_n se tiene

$$J_n(x) = g(x) + d_n^-(x). \tag{4.2.2}$$

Así, se obtiene

$$\begin{aligned} d_{n+1}(x) &= g(x) - c(x) - \beta \int J_n(x')Q^X(dx'|x) \\ &= g(x) - c(x) - \beta \int g(x')Q^X(dx'|x) - \beta \int d_n^-(x')Q^X(dx'|x) \end{aligned}$$

y el enunciado se sigue.

d) Esta parte se sigue de nuevo del Teorema 1.5.2 y el resultado de monotonicidad en la parte b).

4.2.1. Precio de Opciones Americanas

Una subclase muy importante de los Problemas de Decisión de Markov que se consideran son los problemas de paro óptimo. Ahí, un Proceso de Markov $(X_n)_{n\in\mathbb{N}}$ está dado, el cual no puede ser influenciado por el decisor. Sin embargo, este proceso tiene que ser parado en algún punto del tiempo n y entonces una recompensa $g_n(X_n)$ es obtenida. Así, la sola decisión en cada punto del tiempo es si el proceso debe continuar o parar. Una vez parado, no es necesario una decisión adicional. Algunas veces se paga un costo o se obtiene una recompensa adicional, siempre que el proceso no sea parado. Por supuesto, el fin es encontrar un tiempo de paro tal que la recompensa de paro esperada sea maximizada.
Se considera el caso de un problema de paro con un *horizonte finito*. Se asume que el decisor tiene que parar a más tardar al tiempo N. El problema se introduce usando tiempos de paro y entonces se muestra cómo puede ser formulado como un Problema de Decisión de Markov. Debido al simple espacio de acciones se desprendería que la suposición de estructura siempre se satisface con el conjunto de todos los mapeos medibles e inmediatamente se obtiene un algoritmo.
Una aplicación clásica de problemas de paro óptimo en finanzas son las opciones Americanas. Con el propósito de encontrar el precio justo de una opción Americana y su tiempo óptimo de ejercicio, tenemos que resolver el problema de paro óptimo de horizonte finito N. En contrates a una opción Europea, el comprador de una opción Americana puede elegir cualquier tiempo, incluyendo el tiempo de expiración N, para ejercer la opción. Se considera el *modelo binomial* como un mercado financiero subyacente con la suposición $\mathbf{d} < 1 + i < \mathbf{u}$ lo cual implica la no oportunidad de arbitraje y la existencia de una única

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 70 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 71 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

medida $\mathbb{Q}$. Esta medida es usada para precios y también es llamada medida de probabilidad neutral al riesgo (ver [23] y [29]). Bajo $\mathbb{Q}$, la probabilidad de un movimiento de la acción está dada por

$$q = \frac{1+i-\mathbf{d}}{\mathbf{u}-\mathbf{d}}. \tag{4.2.3}$$

Se consideran opciones Americanas generales con fecha de expiración finita y se concentra en el caso de opciones con trayectorias independientes.

Opciones Americanas

El análisis se concentra en trayectorias independientes de opciones Americanas, es decir, el proceso (X_n) que tiene que ser parado está dado por el proceso del precio de la acción $X_n = S_n$. Una trayectoria independiente de una opción Americana da el pago $h(S_n)$ si es ejercida en el tiempo n, es decir, la función de pago depende sólo del precio actual de la acción y no de su trayectoria. La fecha de expiración se asume que es N. Sin embargo, la opción nunca puede ser ejercida, en dicho caso el pago es cero. Así, es fácil ver que no es óptimo ejercer cuando $h(S_n) < 0$ y, de manera equivalente, podemos elegir $h^+(S_n)$ como pago. Denotar $\beta := (1+i)^{-1}$. El tiempo de esta opción es entonces calculado como

$$\sup_{\tau \leqq N} \mathbb{E}_x^{\mathbb{Q}}[\beta^\tau h^+(S_\tau)],$$

donde el supremo es tomado sobre todos los tiempos de paro τ con $\mathbb{P}(\tau \leq N) = 1$. $S_0 = x$ es el precio de la acción al tiempo cero y la esperanza es tomada respecto a la medida neutral al riesgo $\mathbb{Q}$. Este problema de paro puede ser formulado como un Problema de Decisión de Markov. Los datos del problema de paro son:

- $E := \mathbb{R}_+$, donde x representa el precio actual de la acción,
- $A := \{0,1\}$, donde $a = 0$ significa continuar y $a = 1$ significa ejercer,
- $\mathbb{Q}^X(B|x) = q\delta_{x\mathbf{u}}(B) + (1-q)\delta_{x\mathbf{d}}(B)$, $x \in E$ para conjuntos de Borel B donde q es dado por (4.2.3),
- $r_N(x) := h^+(x)$ y $r(x) \equiv 0$,
- $\beta := (1+i)^{-1} \in (0,1]$ es el factor de descuento.

Notar que cuando S_0 es el precio inicial de la acción, entonces al tiempo n en el modelo binomial los únicos precios de la acción están dados por

$$\{S_0 \mathbf{u}^k \mathbf{d}^{n-k} | k = 0, \ldots, n\}.$$

iThenticate Página 71 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 72 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

La suposición (B_N) se satisface ya que

$$\sup_{n\leq\tau\leq N} \mathbb{E}^{\mathbb{Q}}_{nx}[\sum_{k=n}^{\tau-1}\beta^k c^+(X_k)+\beta^\tau h^+(X_\tau)] \leq \mathbb{E}^{\mathbb{Q}}_{nx}[\sum_{k=n}^{N} h^+(X_k)] < \infty$$

porque X_k sólo puede tomar un número finito de valores posibles con probabilidad positiva para todo $k = 1,\ldots,N$. Además, la siguiente iteración de valor se cumple para este problema.

Algoritmo de Precios de Opciones Americanas.
Suponer que la función de pago está dada por h.

1. Considerar $n := 0$ y definir para $x \in \{S_0\mathbf{d}^k\mathbf{u}^{N-k} | 0 \leq k \leq N\}$:

$$J_0(x) := h^+(x).$$

   Considerar $f_0^*(x) := 1$ si $h(x) \geq 0$ y $f_0^*(x) := 0$ si $h(x) < 0$.

2. Sustituir n por $n+1$ y calcular para todo $x \in \{S_0\mathbf{d}^k\mathbf{u}^{N-n-k} | 0 \leq k \leq N-n\}$:

$$J_n(x) := \text{máx}\left\{h^+(x), \beta\left(qJ_{n-1}(x\mathbf{u})+(1-q)J_{n-1}(x\mathbf{d})\right)\right\}.$$

   Considerar $f_n^*(x) := 1$ si $J_n(x) = h^+(x)$ y cero de otro modo.

3. Si $n = N$, entonces la función de valor $J_N(S_0)$ está calculado y una política óptima π^* está dada por $\pi^* = (f_N^*,\ldots,f_1^*,f_0^*)$. De otra manera, ir al paso 2.

El precio de la opción Americana, al tiempo n, está dado por $\pi_n(S_n) := J_{N-n}(S_n)$ y un tiempo óptimo de ejercicio para el periodo de tiempo $[n,N]$ es

$$\tau_n^* := \text{ínf}\{k \in \{n,\ldots,N\} | f_{N-k}^*(S_k) = 1\}.$$

■

Notar que si se considera ínf$\emptyset := N+1$, entonces $\tau_n^* = N+1$ significa que la opción nunca se ejerce. La siguiente iteración es equivalente y da el mismos valor.

$$J_0(x) := h^+(x)$$
$$J_n(x) = \text{máx}\left\{h(x), \beta\left(qJ_{n-1}(x\mathbf{u})+(1-q)J_{n-1}(x\mathbf{d})\right)\right\}$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 72 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 73 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Desde un punto de vista numérico es importante que, en cada punto del tiempo n, sólo un subconjunto de los precios de la acción en E pueda ser alcanzado y que es razonable calcular el valor de la función sólo en esos precios. De hecho, ésto está en el algoritmo. El pago inmediato $h(x)$ que es obtenido cuando se ejerce la opción es llamado el *valor intrínseco* de la opción.

Opción de compra Americana

Se considera una opción Americana de compra con precio de ejercicio de \$118. Se sabe que el precio del activo es de \$100 y se considera el modelo binomial como el mercado financiero. ¿Cuál es el momento óptimo para hacer uso de la opción Americana de compra?

Puesto que se considera el modelo binomial como mercado financiero, se deben estimar los valores de **u** y **d**. Además, se debe conocer la tasa de interés en cuestión.

Después de realizar un análisis estadístico de los precios del activo se ha determinado que el cambio de precio relativo es de 1.1 y 0.9, para subir y para bajar, respectivamente. Se sabe que el precio del activo actual es de \$100 y la tasa de interés en el mercado es del 3 % por periodo. Luego, se tienen los siguientes datos:

- $S_0 = 100$,
- **u**$= 1.1$ y **d**$= 0.9$,
- $i = 0.03$,
- $N = 4$,
- $K = 118$.

Por lo tanto, se considera $\beta = 0.97$, $q = 0.65$ y la función de utilidad $h(x) = x - 118$, donde x es el precio del activo.

Se utiliza el Algoritmo de Precios para Opciones Americanas y el Programa R para calcular los posibles precios del activo en cada etapa del tiempo n. También, se calcula π^*, $\pi_n(S_n)$ y τ_n^*. Los 15 precios posibles del activo son

$$\{146.41, 133.10, 121.00, 119.79, 110.00, 108.90, 100.00,$$
$$99.00, 98.01, 90.00, 89.10, 81.00, 80.19, 72.90, 65.61\}$$

y se muestran en la Figura 4.1 (ver Apéndice B, Código 2-1).
Considerar $n = 0$ y definir para $x \in \{146.41, 119.79, 98.01, 80.19, 65.61\}$:

$$J_0(x) := h^+(x),$$

de donde se tiene

$$J_0(146.41) = 28.41 \quad \text{y} \quad J_0(119.79) = 1.79$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 73 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 74 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

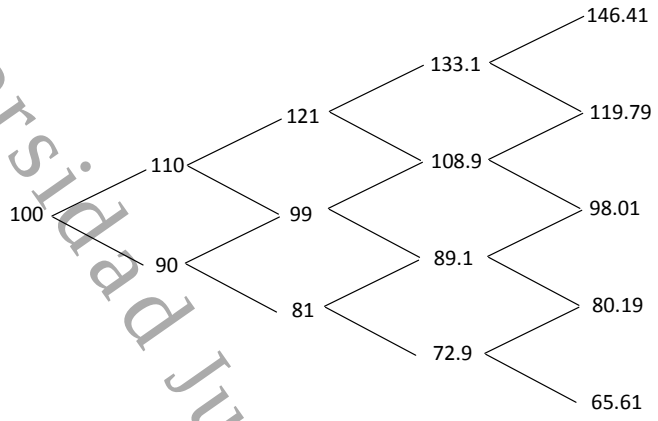


Figura 4.1: Evolución del precio de un activo en el modelo binomial en 4 etapas con precio inicial $S_0 = 100$, $\mathbf{u}$= 1.1 y $\mathbf{d}$= 0.9.

y para $x = 98.01, 80.19, 65.61$ se tiene $J_0(x) = 0$. Por lo tanto, $f_0^+(x) := 1$, si $x = 146.41$ ó $x = 119.79$, y en cualquier otro caso se tiene $f_0^+(x) := 0$ (ver Apéndice B, Código 2-2). Ahora, calcular

$$J_n(x) := \text{máx}\left\{h^+(x), \beta\big(qJ_{n-1}(x\mathbf{u}) + (1-q)J_{n-1}(x\mathbf{d})\big)\right\},$$

para todo $x \in \{100.00, 90.00, 110.00, 81.00, 99.00, 121.00, 72.90, 89.10, 108.90, 133.10\}$, con $\beta \approx 0.97$, $q = 0.65$ y $h(x) = x - 118$, para $n = 1, 2, 3, 4$. El Cuadro 4.1 muestra los valores de $J_n(x)$ para cada posible precio de x, para las etapas $n = 0, 1, 2, 3, 4$. (ver Apéndice B, Código 2-3)

J_0	J_1	J_2	J_3	J_4
28.41	18.54	12.08	7.87	5.12
1.79	1.13	0.71	0.45	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00

Cuadro 4.1: Valores de J_n

El Cuadro 4.2 muestra que la política óptima es $\pi^* = (f_4^*, f_3^*, f_2^*, f_1^*, f_0^*) = (0, 0, 0, 0, 1)$ si el precio es $x = 146.41$. También, del Cuadro 4.2 se observa que si en la etapa 4 el precio del activo es de 146.41, el precio de la opción al tiempo $n = 0$ es

$$\pi_0(S_0) = J_4(S_0) = J_4(100) = 5.12$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 74 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 75 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

y el tiempo óptimo de ejercicio es $\tau^* = 4$, es decir, en la etapa $N = 4$, con una ganancia

$$h(146.41) = 28.41.$$

N	n	x	$h(x)$	$\beta\left(qJ_{n-1}(x\mathbf{u}) + (1-q)J_{n-1}(x\mathbf{d})\right)$	$J_n(x)$	$f_n^*(x)$
4	0	146.41	28.41		28.41	1
3	1	133.1	15.1	18.54	18.54	0
2	2	121	3	12.08	12.08	0
1	3	110	-8	7.87	7.87	0
0	4	100	-18	5.12	5.12	0

Cuadro 4.2: Valores de f^* si el precio del activo es $x = 146.41$

El Cuadro 4.3, en el caso de que el precio del activo en la etapa 4 sea $x = 119.79$ y en la etapa 3 haya sido de 133.1, muestra que la política óptima también es $\pi^* = (f_4^*, f_3^*, f_2^*, f_1^*, f_0^*) = (0,0,0,0,1)$. Del Cuadro 4.3 se puede observar que el tiempo óptimo de ejercicio es $\tau^* = 4$, es decir, en la etapa $N = 4$, con una ganancia

$$h(119.79) = 1.79.$$

Sin embargo, el precio de la opción al tiempo $n = 3$ es

$$\pi_3(S_3) = J_1(S_3) = J_1(133.1) = 18.54.$$

N	n	x	$h(x)$	$\beta\left(qJ_{n-1}(x\mathbf{u}) + (1-q)J_{n-1}(x\mathbf{d})\right)$	$J_n(x)$	$f_n^*(x)$
4	0	119.79	1.79		1.79	1
3	1	133.1	15.1	18.54	18.54	0
2	2	121	3	12.08	12.08	0
1	3	110	-8	7.87	7.87	0
0	4	100	-18	5.12	5.12	0

Cuadro 4.3: Valores de f^* si el precio del activo es $x = 119.79$

Observar que si el precio del activo en la etapa 4 es de 119.79, el precio en la etapa 3 pudo haber sido 108.9, es decir, $S_4 = 119.79$ y $S_3 = 108.9$. En este caso, también se tiene la política óptima $\pi^* = (f_4^*, f_3^*, f_2^*, f_1^*, f_0^*) = (0,0,0,0,1)$ y el precio de la opción en la etapa 3 es

$$\pi_3(S_3) = J_1(S_3) = J_1(108.9) = 1.13.$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 75 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 76 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Para una opción de venta Americana la función de utilidad será $h(x) = 118 - x$. Por lo que se hará uso de la opción si el precio del activo es $98.01, 80.19$ ó 65.61. En el caso de que el activo cueste $x = 146.41$ en la etapa 4, entonces la política óptima es $\pi^* = (f_4^*, f_3^*, f_2^*, f_1^*, f_0^*) = (0,0,0,1,0)$, es decir, hacer uso de la opción en la etapa $n = 3$ con una ganancia de

$$h(133.1) = 15.1$$

y el precio de la opción al tiempo $n = 0$ es

$$\pi_0(S_0) = J_4(S_0) = J_4(100) = 14.82.$$

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 76 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco.
México.

Universidad Juárez Autónoma de Tabasco. México.

Conclusiones

Los Procesos de Decisión de Markov con recompensa total esperada descontada, como criterio de rendimiento, permiten modelar adecuadamente problemas financieros tales como el precio de opciones Americanas y problemas de consumo inversión considerando espacios de estados y acciones de Borel y recompensas no acotadas (ver [1], [8] y [15]).

A continuación se enlistan los resultados que se observan en el trabajo de la tesis.

1. **Problema de optimización de un periodo**

   Al considerar la función de utilidad $U = \log x$, la fracción óptima la cual se debe invertir es

   $$\alpha^* := (1+i)\Big(\frac{p}{1+i-\mathbf{d}} - \frac{1-p}{\mathbf{u}-1-i}\Big).$$

   Notar que $\alpha^* \in (\alpha_0, \alpha_1)$, $\alpha^* = \alpha^*(p)$ es creciente en p y que $\alpha^*(p) = 0$ si $p = \dfrac{1+i-\mathbf{d}}{\mathbf{u}-\mathbf{d}}$. Puesto que

   $$\mathbb{E}R = p(\frac{\mathbf{u}}{1+i} - 1) + (1-p)(\frac{\mathbf{d}}{1+i} - 1),$$

   si $p = \dfrac{1+i-\mathbf{d}}{\mathbf{u}-\mathbf{d}}$, entonces $\mathbb{E}R = 0$. En este caso, lo óptimo es únicamente invertir en el bono. Como consecuencia, si $p < \dfrac{1+i-\mathbf{d}}{\mathbf{u}-\mathbf{d}}$ es óptimo vender la acción, y si $p > \dfrac{1+i-\mathbf{d}}{\mathbf{u}-\mathbf{d}}$ es óptimo comprar la acción (ver [1]).

2. **Problema de Consumo-Inversión**

   Se presentó un caso particular del problema de consumo-inversión, al considerar el proceso de riqueza

   $$X_{n+1} = (X_n - f(X_n)) \cdot Z_n,$$

iThenticate Página 79 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

donde se asume que las variables aleatorias Z_n son independientes e idénticamente distribuidas independientes de la riqueza inicial X_0 y además

$$Z_0 \sim N(1.05, 0.6) \quad \text{si } z \geq 0,$$

en cualquier otro caso, $Z_0 = 0$. En [15] se hace el análisis cuando las variables aleatorias Z_n son no negativas con densidad común ξ.

El problema de consumo-inversión con horizonte aleatorio es tratado en [1, Capítulo 9], donde se considera un mercado financiero con d activos con riesgo y un bono sin riesgo y el procesos de riqueza es

$$X_{n+1} = X_n - f(X_n) + \phi_n \cdot R_{n+1},$$

donde se considera que no hay arbitraje y $\mathbb{E}||R_1|| < \infty$.

3. **Opción de compra Americana**

Si se considera el caso especial de opción de compra Americana con precio de ejercicio K, el pago cuando se ejerce la opción está dado por $h(x) = x - K$. El precio es el mismo como para una opción de Compra Europea donde elegir, si ejercer o no, es sólo dado al tiempo N. Se prueba esta declaración en el marco de Modelos de Decisión de Markov: Se afirma que para $n = 0, 1, \ldots, N-1$ y $x \in E$:

$$h^+(x) \leq \beta \left(qJ_n(x\mathbf{u}) + (1-q)J_n(x\mathbf{d}) \right)$$

lo que implica que $f_1^*(x) = \cdots = f_N^*(x) = 0$ es una estrategia óptima de ejercicio, es decir, no ejercemos hasta el tiempo N. La desigualdad anterior es cierta puesto que

$$\begin{aligned} \beta \left(qJ_n(x\mathbf{u}) + (1-q)J_n(x\mathbf{d}) \right) &\geq \beta \left(qJ_0(x\mathbf{u}) + (1-q)J_0(x\mathbf{d}) \right) \\ &\geq qJ_0\left(\beta x\mathbf{u}\right) + (1-q)J_0\left(\beta x\mathbf{d}\right) \\ &\geq J_0\left(q\beta x\mathbf{u} + (1-q)\beta x\mathbf{d}\right) \\ &= J_0(x) \geq h^+(x) \end{aligned}$$

donde se usa la convexidad de J_0 y la propiedad de martingala del precio descontado de una acción bajo $\mathbb{Q}$, es decir, considerar

$$q = \frac{1 + i - \mathbf{d}}{\mathbf{u} - \mathbf{d}}.$$

Ver [1].

iThenticate Página 79 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 80 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

Apéndice A

Herramientas de Matemática Financiera

A.1. Teoría de Precios bajo No Arbitraje

En esta sección resumimos algunos hechos de la teoría fundamental de precios bajo no arbitraje y algo de cobertura de martingalas en el precios de opciones. En lo que sigue suponer que un espacio de probabilidad filtrado $(\Omega, \mathscr{F}, (\mathscr{F}_n), \mathbb{P})$ está dado, donde $\mathscr{F}_0 = \{\emptyset, \Omega\}$. En este espacio existen $d+1$ activos y el precio al tiempo $n = 0, 1, ..., N$ de k activos está modelado por una variable aleatoria S_n^k (ver Sección 3.1). Un activo S^0 es un bono sin riesgo el cual se usa como numerario.

Definición A.1.1. *a)* Una medida de probabilidad $\mathbb{Q}$ en $(\Omega, \mathscr{F})$ es llamada *equivalente* a $\mathbb{P}$ si $\mathbb{Q}$ y $\mathbb{P}$ tienen los mismos conjuntos nulos.

b) Una medida de probabilidad $\mathbb{Q}$ en $(\Omega, \mathscr{F}_N)$ es llamada una *medida martingala* si el proceso de precio descontado de acciones $(\frac{S_n^k}{S_n^0})$ es una $(\mathscr{F}_n)$-martingala bajo $\mathbb{Q}$ para toda $k = 1, ..., d$.

La siguiente caracterización de no arbitraje es crucial.

Teorema A.1.1. El mercado financiero es libre de arbitraje si y sólo si existe una medida martingala equivalente.

iThenticate Página 80 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 81 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Definición A.1.2. *a)* Un *crédito eventual* es una variable aleatoria H no negativa $\mathscr{F}$-medible.

b) Un crédito eventual se dice ser *alcanzable* si existe una estrategia de portafolio ϕ la cual replica el pago H, es decir,

$$H = X_N^{\phi}, \quad \mathbb{P}-c.s.$$

La estrategia de portafolio ϕ es llamada una *estrategia replicante* o *estrategia de cobertura*.

c) El mercado financiero es llamado *completo* si cada crédito eventual es alcanzable.

Teorema A.1.2. Suponer que el mercado financiero no admite oportunidades de arbitraje. Entonces el mercado es completo si y sólo si existe una única medida martingala equivalente.

Ejemplo A.1.1. (Modelo Binomial o Cox-Ross-Rubinstein). Si los parámetros en el modelo binomial satisface

$$\mathbf{d} < 1+i < \mathbf{u}$$

entonces el mercado no admite oportunidades de arbitraje (ver Ejemplo 3.1.1) y también es completo. Si se denota el espacio muestral en este modelo por

$$\Omega := \{(\omega_1, ..., \omega_N) | \omega_i \in \{\mathbf{d}, \mathbf{u}\}\}$$

y $\tilde{R}_n(\omega) = \omega_n$, entonces la única medida martingala $\mathbb{Q}$ está determinada por

$$\mathbb{Q}(\tilde{R}_n = \mathbf{u}) = q = \frac{1+i-\mathbf{d}}{\mathbf{u}-\mathbf{d}}$$

y el hecho de que los cambios de precios relativos aleatorios $\tilde{R}_1, .., \tilde{R}_N$ son independientes bajo $\mathbb{Q}$.

♦

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 81 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

Apéndice B

Códigos de Programación

Se utiliza el Programa R para los cálculos.

Universidad Juárez Autónoma de Tabasco. México

Código 1

Las siguientes líneas de código calculan la cantidad a inversión en el bono y en el activo con riesgo en un periodo de inversión, dada una riqueza $x > 0$.

```
utilidad.log<- function(x,i,u,d,p){
  c<- 1+i
  A<- 1+i-u
  B<- 1+i-d
  p0<- B/(u-d)
  R1<- -A/c
  R2<- -B/c
  S1<- 1+ alphax*R1
  S2<- 1+ alphax*R2
  alphax <- (1+i)*(p/B + (1-p)/A)
  alpha0 <- c/A
  alpha1 <- c/B

  inversion<- c(x-alphax*log(x),alphax*log(x))

  S <- c(S1,S2)
  alpha<- c(alpha0,alpha1)

  z<- data.frame(Neutral=p0, alfa0=alpha0, alfa1=alpha1, alfa=alphax)
  z1<- data.frame(Inversión=x, bono = inversion[1], acción=inversion[2])
  print(z)
return(z1)
}
```

iThenticate Página 84 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Código 2-1

Las siguientes líneas de código calculan los $(N+1)\big((N+1)+1\big)/2$ posibles precios del activo en un mercado binomial, con S_0 como valor actual del activo, **u** y **d** como cambios de precios relativos y N como el número de etapas a considerar:

```
precios <- function(S0, u, d, N){
        S<-c()
        S[1] <- S0
        count <- 2
        for (i in 1:N) {
              for (k in 0:i) {
                    S[count] <- S0*u^k*d^(i-k)
                    count <- count + 1
                     }
              }
        return(S)
        }
```

Por ejemplo,

```
> sort(precios(100,1.1,.9,4))
```

proporciona los posibles precios de un activo con precio inicial $S_0 = 100$, **u**$= 1.1$, **d**$= 0.9$, como cambios de precios relativos en $N = 4$ etapas en orden ascendente.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 84 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 85 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Código 2-2

Las siguientes líneas de código calculan la regla de decisión $f_0^*(x)$, el precio x del activo, el valor de $J_0(x)$ y la ganancia $h(x)$ en el tiempo $n = 0$.

```
decision0<- function(S0,u,d,K,N){
     f=rep(0,N+1)
    X=rep(0,N+1)
    J0<-c()
    x<-c()
    for(k in 0:N)
    {
      X[k+1]= S0*d^k*u^(N-k)
      h=X-K
      J0[k+1]=max(0,h[k+1])
      if(h[k+1]>=0)
        f[k+1]=1
    }
    y<-c(f,X,J0,h)
    Y<-matrix(y,ncol=4)
  return(Y)
}
```

Por ejemplo,

```
> decision0(100,1.1,.9,118,4)
      [,1]       [,2]        [,3]
[1,]    1        146.41       28.41
[2,]    1        119.79        1.79
[3,]    0         98.01      -19.99
[4,]    0         80.19      -37.81
[5,]    0         65.61      -52.39
```

proporciona la decisión $f_0^* = 1$, cuando $x = 146.41$ ó $x = 119.79$.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 85 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 86 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

Código 2-3

Las siguientes líneas de código calculan la política π^*, el precio π_n y el momento óptimo de ejercicio τ_n^* al tiempo n para una opción de compra Americana (o venta si se considera $h(x) = K - x$ como la ganancia al momento de ejercer), donde se consideran N etapas, un mercado financiero binomial con **d** y **u** como los cambios de precios relativos, el precio S_0 de un activo al tiempo $n = 0$ (al momento de pactar la opción), una tasa de interés i por etapa y K como precio del ejercicio.

```
politica<- function(S0,u,d,i,K,N,n){
  b<-c()
  b<-(1+i)^(-1)
  q<-c()
  q<-(1+i-d)/(u-d)
  f=rep(0,N+1)
  S<-rep(0,(N*(N+1)/2))
  S[1]<-S0
  X=rep(0,N+1)
  J0<-c()
  x<-c()
  count<- 2
  for(t in 1:N){
    for(k in 0:i){
      S[count]= S0*d^(t-k)*u^k
      count<- count +1
    }
  }
  for(k in 0:N){
    X[k+1]= S0*d^k*u^(N-k)
    h=X-K
    J0[k+1]=max(0,h[k+1])
    if(h[k+1]>=0)
      f[k+1]=1
    lanbda<-which(f>0)
    for(j in 1:length(lanbda)){
      x[j]=X[j]
    }
  }
  y<-c(f,X,J0,h)
  Y<-matrix(y,ncol=4)
```

iThenticate Página 86 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

```
A<- matrix(0,N+1,N+1)
for(j in 1:N){
 Z<-c()
 for(k in 0:(N-j)){
  Z[k+1]=S0*d^k*u^(N-j-k)
 }
 A[1:length(Z),j+1]=Z
 A[,1]=X
}
J<-matrix(0,N+1,N+1)
J[,1]<-J0
for(t in 2:(N+1)){
 D<-c()
 for(j in 1:(N+2-t)){
 D[j]= max(A[,t][j]-K, b*(q*(J[,t-1][j]) + (1-q)*(J[,t-1][j+1])))
 }
 J[1:length(D),t]=D
}
W<-matrix(0,N+1,N+1)
for(t in 2:(N+1)){
 D<-c()
 for(j in 1:(N+2-t)){
  D[j]= b*(q*(J[,t-1][j]) + (1-q)*(J[,t-1][j+1]))
 }
 W[1:length(D),t]=D
}
H<-matrix(0,N+1, N+1)
H[,1]<-J0
for(t in 2:(N+1)){
 D<-c()
 for(j in 1:(N+2-t)){
  D[j]= A[,t][j]-K
 }
 H[1:length(D),t]=D
}
P<-matrix(0,N+1,N+1)
for(j in 1:(N+1)){
 for(r in 1:(N+1)){
  if(H[j,r]> W[j,r])
   P[j,r]=1
 }
}
```

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 88 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964


```
  Politica<-matrix(0,length(x), N+1)
  for(j in 1:length(x)){
    Politica[j,]<-P[j,]
  }
  precio.opcion<-c()
  precio.opcion<-J[,n+1]
  tau<-rep(0,(N+1)*(N+1))
  for(j in 1:(N+1)){
    for(i in 1:(N+1)){
      if(P[i,j]>0)
        tau[j*i]<-j
    }
  }
   tau.estrella1<-max(tau)
  tau.estrella<- N+1-tau.estrella1
  print(tau.estrella)
  print(precio.opcion)
  return(Politica)
}
```

Por ejemplo,

```
politica(100,1.1,.9,.03,118,4,1)
   [1] 4
   [1] 18.536893  1.129612  0.000000  0.000000  0.000000
         [,1] [,2] [,3] [,4] [,5]
   [1,]    1    0    0    0    0
   [2,]    1    0    0    0    0
```

muestra que en la etapa $n = 1$ el tiempo óptimo de ejercicio es hasta la etapa 4 y que el precio de la opción es 18.536, 1.129 ó 0, dependiendo del precio del activo. La política es la misma, $f^* = (1,0,0,0,0)$, en los casos en que se ejerce la opción.

Para ver los valores de x para los cuales se hace uso de la opción de compra Americana y para obtener los valores del Cuadro 4.1, agregarle al Código 2-3 *print(x)* y *print(J)* con los datos $S_0 = 100$, $\mathbf{u} = 1.1$, $\mathbf{d} = .9$, $i = .03$, $K = 118$, $N = 4$ y $n = 0$. A saber,

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 88 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México.

78

```
politica(100,1.1,.9,.03,118,4,0)
[1] 146.41 119.79
      [,1]      [,2]       [,3]      [,4]     [,5]
[1,] 28.41 18.536893 12.0818880 7.8667270 5.117306
[2,]  1.79  1.129612  0.7128617 0.4498642 0.000000
[3,]  0.00  0.000000  0.0000000 0.0000000 0.000000
[4,]  0.00  0.000000  0.0000000 0.0000000 0.000000
[5,]  0.00  0.000000  0.0000000 0.0000000 0.000000

[1] 4
[1] 28.41  1.79  0.00  0.00  0.00
     [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    0    0    0
[2,]    1    0    0    0    0
```

iThenticate Página 90 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Universidad Juárez Autónoma de Tabasco. México

Bibliografía

[1] N. Bäuerle and U. Rieder. *Markov Decision Processes with Applications to Finance.* Springer-Verlag Berlin Heidelberg, 2011.

[2] P. Willmott, S. Howison and J. Dewynne *The Mathematics of Financial Derivatives.* A Student Introduction, 1995.

[3] D. Bertsekas *Dynamic Programming and Optimal Control. Vol I.* Athena Scientific, Belmont, MA, third edition, 2005.

[4] E. Dynkin and A. Yushkevich. *Controlled Markov Processes.* Springer-Verlag, Berlin, 1979.

[5] D. Bertsekas and S. Shreve *Stochastic Optimal Control.* Academic Press, New York, 1978.

[6] H. Föllmer and A. Schied. *Stochastic Finance.* Walter de Gruyter, Berlin, 2004.

[7] L. van der Hoek and R. Elliott. *Binomial Models in Finance.* Springer-Verlag, New York, 2006.

[8] M. Puterman. *Markov DecisionProcesses: Discrete Stochastic Dynamic Programming.* John Wiley & Sons, New York, 1994.

[9] M. Dammon, C. Spatt and H. Zhang. *"Optimal Consumption and Investment with Capital Gains Taxes".* The Review of Financial Studies Fall, 2001, 14, N. 3,pp. 583-616, 2001.

[10] B. Ferguson and G. Lim. *"Discete Time Dynamic Models".* Taylor Francis, London and New York, 2003.

[11] L. Ljungqvist and T. Sargent. *"Recursive Macroeconomic Theory".* The MIT Press, U.S., 2004.

[12] D. Luenberger. *"Investment Science".* Oxford University Press, New York, 1998.

iThenticate Página 90 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 91 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

[13] A. Shiryaev. *"Essenttials of Stochastic Finance"*. World scientific, Singapore, 1999.

[14] S. Sundaresan. *Continuous-Time Methods in Finance.* A Review and an Assessment, The Journal of Finance, V. LV, N. 4, 2009

[15] H.D. Cruz-Suárez. *An stochastic consumption-investment problem with unbounded utility function.* Morfismos, Vol. 4, No. 1, pp. 19-30, 2000.

[16] O. Hernández-Lerma and J.B. Laserre. *Adaptive Markov Control Processes.* Springer-Verlag New York, 1989.

[17] O. Hernández-Lerma and J.B. Laserre. *Discrete-time Markov Control Processes.* Springer-Verlag New York, 1996.

[18] O. Hernández-Lerma and J.B. Laserre. *Further Topics on Discrete-Time Markov Control Processes.* Springer-Verlag New York, 1999.

[19] W. Powell . *"Approximate Dynamic Programming"*. John Wiley Sons, Inc., New York, 2007.

[20] R. Bellman. *Dynamic programming.* Princeton University Press, Priceton, NJ, 1957.

[21] S. Ross. *Introduction to stochastic dynamic programming.* Academic Press [Harcourt Brace Jonanovich Publishers], New York, 1983.

[22] S. Ross. *Applied Probability Models With Optimization Aplications.* Dover, New York, 1992.

[23] S. Ross. *An Introduction to Mathematical Finance: Options and Other Topics.* Cambridge University Press, 1999.

[24] S. Shreve. *Stochastic calculus for finance. I.* Springer-Verlag New York, 1996.

[25] I. Karatzas and S. Shevre. *"Brownian Motion and Stochastic Calculus"*. Springer-Verlag New York, 1988.

[26] X. Lin. *"Introductory Stochastic Analysis for Finance and Insurance"*. John Wiley & Sons, Inc., New York, 2006.

[27] M. Meyer. *"Continuous Stochastic Calculus with Applications to Finance"*. Chapman&Hall/CRC, US, 2001.

[28] A. Mundt. *"Dynamic Risk Management with Markov Decision Processes"*. Universitasverlag Karlsruhe, 2008.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 91 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

iThenticate Página 92 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

[29] S. Pliska. *Introduction to Mathematical Finance. Discrete Time Models.* Blackwell Publishers, Oxford, 2000.

[30] R. Elliott, L. Aggoun and J. Moore. *Hidden Markov Models.* Springer-Verlag New York, 1995.

[31] R. Elliott and P. Kopp. *Mathematics of Financial Markets.* Springer-Verlag New York, 2005.

[32] C.J. Himmelberg, Parthasarathy, T. and VanVleck, F. S. *Optimal plans for dynamic programming problems. Math. Oper. Res. 1, 390–394*, 1976.

[33] K. Hinderer. *Foundations of non-stationary dynamic programming with discrete time parameter.* Springer-Verlag, Berlin, 1970.

[34] G. Peskir and A. Shiryaev. *Optimal Stopping and Free-Boundary Problems.* Birkhäuser Verlag, Basel, 2006.

[35] J. Hull *Options, Futures, and Other Derivates.* Prentice Hall, Inc, 2009.

Universidad Juárez Autónoma de Tabasco. México.

iThenticate Página 92 de 92 - Engrega de integridad Identificador de la entrega trn:oid:::21:202594964

Rubén Zavala Camargo

PROCESOS DE DECISIÓN DE MARKOV CONAPLICACIONES AFINANZAS

Universidad Juárez Autónoma de Tabasco

Detalles del documento

Identificador de la entrega
trn:oid:::21:202594964

Fecha de entrega
24 mar 2025, 6:40 p.m. GMT-6

Fecha de descarga
9 sep 2026, 12:22 p.m. GMT-6

Nombre del archivo
Rubén Zavala Camargo.pdf

Tamaño del archivo
2.0 MB

89 páginas

20.194 palabras

100.818 caracteres

19% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- Bibliografía
- Texto citado
- Coincidencias menores (menos de 25 palabras)
- Abstract

Fuentes principales

19% Fuentes de Internet
0% Publicaciones
0% Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

Fuentes principales

19% Fuentes de Internet
0% Publicaciones
0% Trabajos entregados (trabajos del estudiante)

Fuentes principales

Las fuentes con el mayor número de coincidencias dentro de la entrega. Las fuentes superpuestas no se mostrarán.

1	Internet	
hdl.handle.net		10%
2	Internet	
revistas.ujat.mx		6%
3	Internet	
www.fcfm.buap.mx		<1%
4	Internet	
ri.ujat.mx		<1%
5	Internet	
andresarango.weebly.com		<1%
6	Internet	
es.wikipedia.org		<1%
7	Internet	
docplayer.es		<1%
8	Internet	
finanzas-mpc.blogspot.com		<1%
9	Internet	
www.amestad.mx		<1%
10	Internet	
1library.co		<1%
11	Internet	
Nicole Bäuerle, Ulrich Rieder. "Markov Decision Processes", 'Springer Fachmedien...		<1%