La Diferenciabilidad en Procesos de

Control de Markov Descontados

Tesis presentada por:

Hugo Adán Cruz Suárez

Para obtener el grado de

Doctor en Ciencias (Matemáticas)

En la especialidad en
Matemáticas

Tesis dirigida por

Dr. Raúl Montes de Oca M.

Departamento de Matemáticas

Universidad Autónoma Metropolitana Iztapalapa
México D.F.

Diciembre 2006

Introducción

La presente tesis está relacionada con la teoría de **Procesos de Control de Markov** (PCMs). En ella se tratarán problemas a tiempo discreto, con horizonte infinito ([9], [25], [31] y [33]).

De manera general, un Proceso de Control de Markov (PCM) se encarga de modelar un sistema dinámico cuyos estados son observados de manera periódica por un controlador. El desarrollo de un PCM, a través del tiempo, está dado de acuerdo al siguiente procedimiento. En cada tiempo t, $t = 0, 1, ...$, el controlador decide el control que aplicará dependiendo del estado del sistema. Entonces, como consecuencia del estado y de haber aplicado el control, se paga un costo, y el sistema, mediante la ley de transición, se traslada a un nuevo estado en el instante de tiempo $t+1$. Al ocurrir un estado en $t + 1$, se repite el procedimiento anteriormente descrito.

A la sucesión de controles aplicados en cada tiempo se le llama **política**. Para evaluar la calidad de cada política se cuenta con un criterio de rendimiento o función objetivo.

En el trabajo consideraremos la función objetivo conocida como **costo total descontado**. A un PCM equipado con esta función objetivo le llamaremos un **PCM descontado**.

El **problema de control óptimo** consiste en encontrar una política que optimice el criterio de rendimiento. La política que optimiza el criterio de rendimiento se le llama **política óptima**. También, al criterio de rendimiento evaluado en la política óptima se le conoce como la **función de valores óptimos**.

En el desarrollo de la tesis se supondrá que se tienen condiciones para la existencia de una política óptima, lo cual es posible garantizar bajo condiciones bastantes generales ([33]). Además se supone que los espacios de estados y de controles son subconjuntos de espacios euclidianos (no necesariamente compactos).

Un procedimiento de solución para PCMs está basado en el principio de Bellman conocido como **Programación Dinámica (**[8]). La idea de este procedimiento es llevar el problema de control óptimo a un problema equivalente. Dicho problema equivalente consiste en resolver una ecuación funcional para la función de valores óptimos, conocida como **Ecuación de Programación Dinámica** (EPD) ([33]). En general, dicha ecuación, no es posible resolverla de forma directa, salvo en casos especiales. No obstante se encuentra establecido en la literatura de PCMs descontados, bajo condiciones bastantes generales, el **Algoritmo de Iteración de Valores** (ó método de Aproximaciones sucesivas) ([33]). Este Algoritmo permite construir una sucesión de funciones, conocidas como **Funciones de Iteración de Valores** que aproximan a la función de valores óptimos. Después, si con dicho algoritmo es posible determinar a la función de valores óptimos, entonces podemos encontrar la política óptima a partir de la EPD.

En esta tesis se analizarán cierta clase de problemas referentes a la teoría de PCMs descontados, a saber:

1. la diferenciabilidad para sistemas deterministas controlados. Ésta es establecida y usada para deducir una ecuación funcional, conocida como Ecuación de Euler que caracteriza (en particular) a las funciones de iteración de valores;

2. el análisis de problemas de control estocástico vía problemas deterministas controlados;

3. condiciones para determinar la diferenciabilidad de la política óptima y de la función de valores óptimos en problemas de horizonte infinito.

La Ecuación de Euler (EE) ha sido usada especialmente para aplicaciones en Economía (véase [2], [12], [22], [23], [36], [37], [45], [46] y [49]). La finalidad de dar una versión de la EE para PCMs descontados (deterministas), es proveer un procedimiento complementario que facilite la solución de la EPD. La EE es una ecuación funcional la cual se puede expresar en términos de la función de valores óptimos o de la política óptima. En general, no existe un procedimiento para resolver la EE de forma directa. En la literatura se pueden encontrar procedimientos intuitivos para resolverla ([2], [23], [36] y [37]), cuya idea básica consiste en proponer la solución de la EE dentro de una familia de funciones, y a partir de la EE determinar una función en particular, la cual es la solución de ésta. En otros casos se han propuesto procedimientos para aproximar la solución de la EE ([45]). En resumen, no existe un procedimiento general y/o fundamentado teóricamente (hasta donde el autor conoce) para encontrar una solución de la EE. Además, se tienen referencias donde la EE no sólo se ha utilizado para dar solución al problema de control óptimo, sino también para hacer un análisis cualitativo de la solución óptima, como por ejemplo, para analizar la estabilidad de dicha solución ([3] y [12]).

La contribución de la presente tesis al tema de la EE está basada en los siguientes trabajos: [17], [18], [19] y [20]. En la tesis, primeramente se presentan condiciones, para modelos deterministas, que permiten garantizar la diferenciabilidad de la función de valores óptimos y de la política óptima. En modelos deterministas, el problema de la diferenciabilidad ha sido analizado principalmente por Santos en [43], [44] y [45]. La clase de modelos con los que Santos trabaja son problemas de crecimiento económico. El desarrollo de su teoría lo hace bajo suposiciones de concavidad, compacidad y acotabilidad en ciertas componentes del modelo de control. En contraste, las condiciones que presentamos en este trabajo no imponen restricciones de acotabilidad, ni compacidad en las componentes del modelo de control. Una vez establecida la propiedad de diferenciabilidad se procede a dar una versión de la EE para PCMs deterministas, en términos de las derivadas de las funciones de iteración de valores. Para resolver la EE se utiliza un procedimiento iterativo. Este procedimiento es usado para resolver algunos ejemplos de la teoría de control.

El segundo punto que se aborda en la tesis, consiste en analizar problemas de control estocástico que provienen de problemas deterministas. La idea consiste en considerar un problema de control determinista fijo, y a partir de éste, construir un modelo estocástico modificando la dinámica del sistema determinista mediante un ruido aleatorio, dejando fijas las otras componentes del modelo de control. Una pregunta natural al considerar el problema estocástico inducido es, ¿qué propiedades de su solución óptima son heredadas del problema original determinista?. La respuesta a esta cuestión se da en un teorema donde se afirma que, bajo ciertas condiciones, la función de valores óptimos del problema inducido es la suma de la función de valores óptimos del problema original, más una constante real adecuada, y las políticas de ambos problemas coinciden. También, en esta parte se presentan varios ejemplos. En especial se muestra un ejemplo clásico de Economía referente a problemas de crecimiento económico, con función de utilidad logarítmica ([36], [37] y [49]). Este ejemplo no ha sido analizado utilizando las técnicas desarrolladas en este trabajo.

Finalmente, el tercer problema que se analiza es referente a la diferenciabilidad de la función de valores óptimos y de la política óptima en problemas estocásticos, de horizonte infinito. La diferenciabilidad de la función de valores óptimos y de la política óptima son problemas que han sido trabajados para PCMs aplicados a modelos económicos por Araujo & Scheinkman ([4]), Benveniste & Scheinkman ([7]), Blume et al. ([11]), De la Fuente ([22]), Santos ([43], [44] y [45]), entre otros (véase también [12] y [49]). El estudio de la diferenciabilidad para PCMs no sólo es importante como una propiedad cualitativa, sino también hace posible el estudio: de la estabilidad de un PCM ([3] y [12]), del principio del máximo (o mínimo) ([5], [6], [26], [27], [29], [39] y [40]), de análisis comparativo ([45]), de métodos numéricos aplicados a PCMs ([34], [45], [46] y [47]), y de procedimientos de linealización ([48]).

La principal aportación de este trabajo referente a la diferenciabilidad, consiste en la presentación de condiciones que extienden los trabajos mencionados en el párrafo anterior ([18] y [21]). En especial, se extiende el trabajo de Blume et.al. ([11]), el cual es el antecedente más general acerca de la diferenciabilidad. En él, se trabajan modelos de control a tiempo discreto, en espacios euclidianos, con espacio de controles compactos, bajo condiciones de concavidad en el modelo y condiciones de acotabilidad en la función de recompensa. Primeramente la extensión que se hace, es con respecto a la compacidad del espacio de controles y a la acotabilidad de la función de costo. Es decir, con las condiciones dadas en este trabajo permitimos considerar PCMs con espacio de controles no compactos y función de costo posiblemente no acotada. En segundo lugar, se presenta una extensión para modelos no convexos.

Este trabajo está estructurado en cinco capítulos. El segundo, es un capítulo preliminar en el que se formula la teoría básica que se trabajará a lo largo de la tesis. Se definirán la clase de modelos con los que se trabajarán, dando una manera de interpretarlos. Se plantean las hipótesis específicas para nuestros modelos, el problema de control óptimo que nos interesa resolver y presentamos algunas técnicas conocidas en la teoría general de PCMs que se usan para resolverlo. También se incluye una sección referente a la clase de problemas controlados deterministas. En este caso, se presenta el problema de control óptimo correspondiente, y se proporcionan algunas referencias donde se ha trabajado la existencia de políticas óptimas, y también se proporciona una versión de la ecuación de programación dinámica.

Antes de iniciar la teoría desarrollada en la tesis, se presenta un capítulo (Capítulo 3) sobre los problemas propuestos. En este capítulo se dan los antecedentes que motivaron el estudio de la clase de problemas que se consideran en la tesis, y se establecen éstos.

En el Capítulo 4 se muestra la diferenciabilidad para PCMs deterministas y se da una versión de la EE, para dichos modelos. También se presentan ejemplos que muestran la efectividad del uso de la EE cuando se tienen condiciones de diferenciabilidad en el modelo de control. Después, en este mismo capítulo, se proporciona una forma de abordar problemas de control estocástico inducidos por problemas controlados deterministas. En este caso se dan las condiciones para considerar ambos problemas, y se proporcionan los resultados que sustentan la teoría desarrollada. También se presentan las versiones estocásticas de los problemas deterministas trabajados en un inicio, se verifican nuestras hipótesis y se da la solución correspondiente.

En el Capítulo 5 se provee la teoría referente a la diferenciabilidad de la función de valores óptimos y de la política óptima, para problemas de horizonte infinito. Primeramente se presenta un problema de optimización general, para el cual se establecen condiciones que determinan la propiedad de diferenciabilidad en el modelo y también se da una fórmula para calcular la derivada de la función de valores óptimos, conocida como fórmula de la envolvente (en inglés: envelope formula). Para este caso se presentan algunos ejemplos de la teoría de PCMs, donde se aplica directamente al algoritmo de iteración de valores la fórmula de la envolvente. Después, se presenta una sección donde se abordan condiciones referentes a la diferenciabilidad de PCMs descontados, tanto para modelos convexos, como no convexos. Toda esta teoría desarrollada se ejemplifica con algunos problemas de control.

En un último capítulo se dan las conclusiones del trabajo, y se presentan algunos problemas abiertos relacionados con los temas de la tesis.

Finalmente, presentamos un Apéndice A donde se da la notación utilizado en la tesis, y un Apéndice B en el cual se proporcionan algunos teoremas auxiliares usados en la misma.

Preliminares

0.1. Procesos de Control de Markov Descontados

En esta sección se dará la teoría, en forma sintetizada, de procesos de control de Markov. Para un desarrollo detallado de esta teoría y la demostración de los resultados se pueden consultar las referencias: [9], [25], [33], entre otras.

Un Modelo de Control de Markov (MCM), estacionario, a tiempo discreto, consiste de una quíntupla:

$$M = (X, A, \{A(x) \mid x \in X\}, Q, c),$$

donde,

a. X es un espacio de Borel no vacío, llamado el espacio de estados;

b. A es un espacio de Borel no vacío, llamado el conjunto de acciones o controles;

c. $\{A(x) \mid x \in X\}$ es una familia de subconjuntos medibles, no vacíos $A(x)$ de A, donde $A(x)$ denota el conjunto de controles admisibles cuando el sistema se encuentra en el estado $x \in X$. El conjunto $\mathbb{K}$ de parejas de estados acciones admisibles, está definido por

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

y se supone que es un conjunto medible del espacio producto $X \times A$;

d. Q es un kérnel estocástico definido en X dado $\mathbb{K}$, llamado la ley de transición, i.e. para cada $(x, a) \in \mathbb{K}$, $Q(\cdot\,|x, a)$ es una medida de probabilidad en X, y para cada $B \subset X$, medible, $Q(B\,|\cdot)$ es una función medible;

e. $c : \mathbb{K} \to \mathbb{R}$ es una función medible y se llama la función de costo de un paso.

Podemos pensar un MCM estacionario a tiempo discreto como un sistema estocástico controlado que se observa de manera periódica en los tiempos $t = 0, 1, 2....$ Sean x_t y a_t el estado del sistema y el control en el tiempo t. La dinámica que describe a este sistema estocástico funciona de la forma siguiente: si el sistema se encuentra en el estado $x_t = x \in X$, en el tiempo t y la acción $a_t = a \in A(x)$ es aplicada; entonces ocurren dos cosas:

1. se paga un costo $c(x, a)$; y
2. el sistema se traslada a un nuevo estado x_{t+1}, mediante la distribución de probabilidad $Q(\cdot\,|x, a)$ sobre X, es decir,

$$Q(B\,|x, a) = \Pr(x_{t+1} \in B\,|x_t = x, a_t = a)\,,$$

$B \in \mathcal{B}(X)$, donde $\mathcal{B}(X)$ denota la σ-álgebra de Borel de X.

Una vez hecha esta transición a un nuevo estado, se elige una nueva acción y la dinámica anteriormente descrita se repite.

En algunos casos la ley de transición Q está inducida por una ecuación en diferencias de la forma siguiente

$$x_{t+1} = L(x_t, a_t, \xi_t), \tag{1}$$

$t = 0, 1, ...,$ con el estado inicial x_0 conocido. Aquí $\{\xi_t\}$ es una sucesión de variables aleatorias, independientes e idénticamente distribuidas (i.i.d.), con valores en algún espacio de Borel S, con densidad común Δ, e independientes del estado inicial x_0. Se tiene entonces que

$$Q(B\,|x, a) = \int I_B(L(x, a, s))\Delta(s)ds, \tag{2}$$

donde $B \in \mathcal{B}(X)$, $(x, a) \in \mathbb{K}$ y I_B es la función indicadora del conjunto B.

Políticas de control. Para introducir el concepto de estrategia o política, considérese un MCM M y defina $\mathbb{H}_t$, *el espacio de las historias observadas del proceso de control hasta el tiempo* t, como

$$\mathbb{H}_0 = X, \text{ y } \mathbb{H}_t = \mathbb{K} \times \mathbb{H}_{t-1},$$

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

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

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

Sea $\mathbb{F}$ el conjunto de todas las funciones medibles $f : X \to A$ tal que $f(x) \in A(x)$ para todo $x \in X$.

Definition 1 *Una política de control es una sucesión $\pi = \{\pi_t\}$ de kérneles estocásticos definidos en A dado $\mathbb{H}_t$, satisfaciendo la condición*

$$\pi_t(A(x_t)\,|h_t) = 1,$$

para todo $h_t \in \mathbb{H}_t$ y $t = 0, 1, ...$

El conjunto de todas las políticas será denotado por Π. Una clase particular de políticas son las siguientes.

Una política de control $\pi = \{\pi_t\}$ se dice que es:

1. **política determinista** si existe una sucesión $\{g_t\}$ de funciones medibles $g_t : \mathbb{H}_t \to A$ tal que, para toda $h_t \in \mathbb{H}_t$ y $t = 0, 1, ...$, $g_t(h_t) \in A(x_t)$ y $\pi_t(\cdot\,|h_t)$ está concentrado en $g_t(h_t)$;

2. **política Markoviana determinista** si existe una sucesión $\{f_t\}$ de funciones $f_t \in \mathbb{F}$ tal que $\pi_t(\cdot\,|h_t)$ está concentrado en $f_t(x_t) \in A(x_t)$ para toda $h_t \in \mathbb{H}_t$ y $t = 0, 1, ...$ El conjunto de políticas Markovianas será denotado por $\mathbb{M}$;

3. **política estacionaria** si existe una función $f \in \mathbb{F}$ tal que $\pi_t(\cdot\,|h_t)$ está concentrado en $f(x_t) \in A(x_t)$ para toda $h_t \in \mathbb{H}_t$ y $t = 0, 1, ...$ El conjunto de políticas estacionarias será denotado por $\mathbb{F}$.

Nótese que $\mathbb{F} \subset \mathbb{M} \subset \Pi$.

Para cada política π y estado inicial $x_0 = x \in X$, una medida de probabilidad P_x^π es definida sobre el espacio $\Omega = (X \times A)^\infty$ en forma canónica ([33]). E_x^π denotará el correspondiente operador esperanza. El proceso estocástico obtenido será llamado Proceso de Control de Markov (PCM).

Criterio de Rendimiento. Después de definir un MCM y de presentar la definición de política, es necesario definir un indicador que mida las consecuencias de utilizar una política determinada. Para ello es necesario definir un criterio de rendimiento o índice de funcionamiento que mida el rendimiento del sistema cuando una política dada $\pi \in \Pi$ se usa y el estado inicial es $x \in X$.

Definition 2 *El costo descontado total esperado se define como*

$$v(\pi, x) := E_x^\pi \left[\sum_{t=0}^{\infty} \alpha^t c(x_t, a_t) \right], \tag{3}$$

cuando la política $\pi \in \Pi$ *es usada, y* $x \in X$ *es el estado inicial. Aquí,* $\alpha \in (0,1)$ *es un factor de descuento dado.*

Definition 3 *Una política* π^* *se dice que es **óptima** si*

$$v(\pi^*, x) = V(x), \tag{4}$$

$x \in X$*, donde*

$$V(x) := \inf_{\pi \in \Pi} v(\pi, x). \tag{5}$$

$x \in X$*. La función* V *definida en (5) es llamada **función de valores óptimos**.*

Problema de control óptimo. El problema de control óptimo consiste en determinar una política π^* que satisfaga (4).

Ahora se enunciarán algunas suposiciones y resultados que se usarán en los siguientes capítulos.

Hipótesis 1

1. El costo de una etapa c es no negativo, semicontinuo por abajo (s.c.a.) e inf-compacto en $\mathbb{K}$. (c es inf-compacto en $\mathbb{K}$ si el conjunto

$$\{a \in A(x) \mid c(x,a) \leq \overline{s}\}$$

es compacto para cualquier $x \in X$ y $\overline{s} \in \mathbb{R}$).

2. La ley de transición Q es fuertemente continua, i.e.

$$\mu'(x,a) := \int \mu(y)\, Q(dy \mid x, a)$$

es continua y acotada en $\mathbb{K}$, para cualquier función medible y acotada $\mu : X \to \mathbb{R}$.

3. Existe una política π tal que $v(\pi, x) < \infty$, para cada $x \in X$.

Definition 4 *Las funciones de* ***iteración de valores*** *se definen como*

$$V_n(x) = \min_{a \in A(x)} \left[c(x,a) + \alpha \int V_{n-1}(y)\, Q(dy \mid x,a) \right], \tag{6}$$

para todo $x \in X$ *y* $n = 1, 2, ...,$ *con* $V_0(\cdot) = 0$.

Remark 1 *Usando la Hipótesis 1, es posible demostrar ([33]) que, para cada* $n = 1, 2, \cdots$ *, existe una política estacionaria* $f_n \in \mathbb{F}$ *tal que el mínimo en (6) se alcanza, i.e.*

$$V_n(x) = c(x, f_n(x)) + \alpha \int V_{n-1}(y)\, Q(dy \mid x, f_n(x)), \tag{7}$$

$x \in X$.

Lemma 1 *Supongamos que la Hipótesis 1 se cumple. Entonces:*

1. *La función de valores óptimos* V *definida en (5) satisface la* ***Ecuación de Optimalidad****, i.e., para todo* $x \in X$:

$$V(x) = \min_{a \in A(x)} \left[c(x,a) + \alpha \int V(y)\, Q(dy \mid x,a) \right]. \tag{8}$$

2. *Existe* $f \in \mathbb{F}$ *tal que*

$$V(x) = c(x, f(x)) + \alpha \int V(y)\, Q(dy \mid x, f(x)), \tag{9}$$

$x \in X$, *y* f *es óptima.*

3. *Para cualquier* $x \in X$, $V_n(x) \uparrow V(x)$ *cuando* $n \to +\infty$ *, con* V_n *definida en (6).*

0.2. Sistemas Controlados Deterministas

Consideremos un MCM M fijo.

Para sistemas de control determinista se supondrá que la ley de transición Q está dada por una ecuación del tipo

$$x_{t+1} = F(x_t, a_t), \tag{10}$$

$t = 0, 1, ...,$ donde $F : \mathbb{K} \to X$ es una función medible conocida. Obsérvese que en este caso Q está dada por

$$Q(B\,|x,a) = I_B(F(x,a)), \tag{11}$$

para todo $B \in \mathcal{B}(X)$ y $(x,a) \in \mathbb{K}$, donde $I_B(\cdot)$ denota la función indicadora de B. Se denotará a Q por Q_F para enfatizar que Q está inducida por F.

El conjunto de políticas admisibles para sistemas deterministas que se considerarán, serán las políticas Markovianas $\mathbb{M}$, aunque en general es posible tomar otra clase de políticas (ver [30] p. 9).

A continuación se presenta nuevamente los conceptos trabajados en la sección anterior (criterio de rendimiento, políticas, problema de control óptimo, etc.) para el caso determinista. La razón de esto, es debido a que en las secciones siguientes se trabajarán tanto problemas deterministas como estocásticos, por lo tanto consideramos importante dar la notación correspondiente para ambos problemas.

Para cada política $\pi \in \mathbb{M}$ y estado inicial $x \in X$, se define

$$w(\pi, x) := \sum_{t=0}^{\infty} \alpha^t c(x_t, a_t).$$

$w(\pi, x)$ es llamado el **costo total descontado**, donde $\alpha \in (0,1)$ es el factor de descuento.

El problema de control óptimo **determinista** consiste en obtener una política $\pi^* \in \mathbb{M}$, tal que

$$w(\pi^*, x) = \inf_{\pi \in \mathbb{M}} w(\pi, x),$$

$x \in X$, y π^* será llamada una **política óptima** (**determinista**). La función W definida por

$$W(x) := \inf_{\pi \in \mathbb{M}} w(\pi, x),$$

$x \in X$, será llamada la **función de valores óptimos** (**determinista**).

Remark 2 *En aplicaciones de PCMs a Economía, es común presentar el siguiente problema de optimización ([12], [22], [24], [35], [41], [43], [45] y [49]). Sea $\widehat{X} \subset \mathbb{R}^p$. Considérese una multifunción de $\widehat{X}$ a $\widehat{X}$, i.e., para cada $x \in \widehat{X}$, existe un conjunto no vacío $\widehat{X}$, denotado por $\Gamma(x)$ (Γ es conocido como el conjunto de tecnologías). $U : Graph(\Gamma) \to \mathbb{R}$ es la función de utilidad* *(se supone que* $U \leq 0$ *o* $U \leq \tau$*, donde* τ *es un número real fijo, y* $Graph(\Gamma) := \{(x,y) \in \widehat{X} \times \widehat{X} : x \in \widehat{X},\ y \in \Gamma(x)\}$*), y* $\alpha \in (0,1)$ *es el factor de descuento. Entonces el problema de maximización se presenta de la forma siguiente:*

$$\max_{\{x_{t+1}\}} \sum_{t=0}^{\infty} \alpha^t U(x_t, x_{t+1})$$

sujeto a $x_{t+1} \in \Gamma(x_t)$*,* $t = 0, 1, ...,$ *con* $x_0 \in \widehat{X}$ *fijo.*

De hecho, es posible establecer este problema de maximización en términos de un PCM descontado, tomando $X = A = \widehat{X}$*,* $A(x) = \Gamma(x)$*,* $x \in X$*,* $x_{t+1} = F(x_t, a_t) = a_t$ *y* $c(x_t, a_t) = -U(x_t, x_{t+1})$*,* $t = 0, 1, ...$

Definition 5 *Las* ***funciones de iteración de valores*** *son definidas como*

$$W_n(x) = \min_{a \in A(x)} \left[c(x,a) + \alpha W_{n-1}(F(x,a)) \right], \tag{12}$$

para todo $x \in X$ *y* $n = 1, 2, ...,$ *con* $W_0(\cdot) = 0$*.*

Remark 3 *Bajo ciertas condiciones (véase Observación 2.2.6 b)), es posible demostrar que para cada* $n = 1, 2, ...,$ *existe una política estacionaria* $g_n \in \mathbb{F}$ *tal que el mínimo en (12) se alcanza, i.e.*

$$W_n(x) = c(x, g_n(x)) + \alpha W_{n-1}(F(x, g_n(x))), \tag{13}$$

$x \in X$*.*

Lemma 2 *Bajo ciertas condiciones (véase Observación 2.2.6 a) para a) y c) de este Lema, y Observación 2.2.6 b) para b) de este Lema), resulta que:*

a) La función de valores óptimos W *es una solución de la siguiente ecuación (conocida como la* ***Ecuación de Programación Dinámica****):*

$$W(x) = \min_{a \in A(x)} \{ c(x,a) + \alpha W(F(x,a)) \}, \tag{14}$$

$x \in X$*.*

b) Existe $g \in \mathbb{F}$ *tal que* $g(x) \in A(x)$ *alcanza el mínimo en el lado derecho de (14), i.e.,*

$$W(x) = c(x, g(x)) + \alpha W(F(x, g(x))), \tag{15}$$

$x \in X$ *y* g *es óptima.*

c) *Para cada* $x \in X$, $W_n(x) \to W(x)$ *cuando* $n \to +\infty$.

Remark 4 *Para sistemas controlados deterministas, donde la ley de transición es inducida por una función continua* F, *se tiene que* Q_F *es débilmente continua, i.e.*

$$\int v(y) Q_F(dy \,|x, a) = v(F(x, a))$$

es una función continua en $(x, a) \in \mathbb{K}$ *para cada* $v \in \{\zeta : X \to \mathbb{R} : \zeta$ *es una función continua y acotada*$\}$. *La propiedad de la continuidad débil en la ley de transición* Q_F *ha sido usada como una suposición en algunos de las condiciones dadas en a), y b) de la Observación 2.2.6.*

Remark 5

a) *Para costos acotados, véase condiciones (a), (b) y (c) en Teorema 2.8 en [31] p. 23; para costos no acotados, véase las condiciones de los Teoremas 3, 4, 5, y 6 en [41].*

b) *Véase Condiciones 3.3.3 (a), 3.3.3 (b), y 3.3.3 (c1) ó Condiciones 3.3.4 (a) y 3.3.4 (b1), y Teorema 3.3.5 en [33] pp. 27-30.*

Problemas Propuestos

En este capítulo se presentan los problemas que serán abordados y resueltos en los capítulos siguientes. Se supone que contamos con PCMs descontados a tiempo discreto, con horizonte infinito, para los cuales se garantiza la existencia de políticas óptimas estacionarias. A continuación se recuerda la notación presentada en el capítulo anterior. La función de valores óptimos para el caso estocástico es denotada por V (y para el caso determinista se denota por W) y la política óptima por f (en el caso determinista se denota por g).

Los problemas que se trabajan en la tesis están motivados en la búsqueda de procedimientos complementarios que faciliten la solución de la EPD, pensando en PCMs con espacios de estados y de acciones, ambos subconjuntos de espacios euclidianos. Principalmente se consideran problemas donde las componentes del MCM presentan condiciones de diferenciabilidad. Bajo esas consideraciones, los temas que se tratan son los siguientes:

(a) La diferenciabilidad de la función de valores óptimos y de la política óptima en problemas controlados, para los casos determinista y estocástico.

(b) Métodos de solución para el problema de control óptimo como son: la Ecuación de Euler y la Fórmula de la Envolvente.

(c) El tipo de relaciones existentes entre problemas controlados deterministas y estocásticos.

A continuación se presentan los problemas que se analizarán en la tesis, después de ello se dan los antecedentes a éstos.

0.3. Problemas

I. Establecer condiciones que garanticen la diferenciabilidad de la función de valores óptimos W y de la política óptima g.

II. Validar la EE y determinar métodos para su solución.

III. Inducir problemas estocásticos a partir de problemas controlados deterministas y analizar las relaciones que guardan las soluciones (i.e., las funciones de valores óptimos y las políticas óptimas) de ambos problemas.

IV. Establecer condiciones que garanticen la diferenciabilidad de la función de valores óptimos V y de la política óptima f.

V. Determinar condiciones que garanticen que la Fórmula de la Envolvente sea válida en problemas estocásticos.

Para estos problemas se darán condiciones generales en las componentes del modelo de control que garanticen su solución, y además se presentarán ejemplos en donde es posible verificar dichas condiciones.

0.4. Antecedentes

En esta sección se da un resumen de los antecedentes sobre las propiedades de diferenciabilidad de la función de valores óptimos y de la política para PCMs descontados. La diferenciabilidad de la función de valores óptimos y de la política óptima es necesaria para extender la metodología del cálculo diferencial en el contexto de Programación Dinámica. Las propiedades de diferenciabilidad son usadas, por ejemplo: para la caracterización de soluciones óptimas, el análisis comparativo, en el desarrollo de métodos numéricos aplicadas a PCMs, y procedimientos de linealización ([34] y [45]). También, en esta sección, se presentan los antecedentes referentes a la Ecuación de Euler, la cual ha sido trabajada principalmente en problemas de control aplicados a Economía (véase [22], [23], [30], [34], [36], [37], [45] y [49]).

La diferenciabilidad de clase C^1 de la función de valores óptimos para el caso determinista fue probada en 1979 por Benveniste & Scheinkman ([7]). A continuación se dan las condiciones que se presentan en su trabajo, para ello considérese un MCM $M = (X, A, \{A(x) : x \in X\}, Q, c)$ fijo donde la ley de transición Q está inducida por una ecuación en diferencias de la forma

$$x_{t+1} = F(x_t, a_t), \tag{3.1}$$

$t = 0, 1, ...,$ donde $F : \mathbb{K} \to X$ es una función medible conocida y $\mathbb{K} = \{(x, a) : x \in X, a \in A(x)\}$.

Hipótesis (BS)

1. $X, A \subset \mathbb{R}^n, n \geq 1$ conjuntos convexos;
2. $A = A(x), x \in X$;
3. $c \in C^2(int(\mathbb{K}), \mathbb{R}), F \in C^2(int(\mathbb{K}), X)$;
4. c y F son funciones convexas en $\mathbb{K}$.
5. $g(x) \in int(X), x \in X$.

Teorema (BS): *Bajo las Hipótesis (BS) se garantiza que W es de clase $C^1(X, R)$.*

La prueba dada por Benveniste & Scheinkman ([7]) está basada en ciertas propiedades de funciones convexas.

Para las próximas dos hipótesis (Hipótesis (S) y (B)) considérese un MCM fijo de la forma $M = (X, A, \{A(x) \mid x \in X\}, Q, c)$, donde Q está inducida por una ecuación en diferencias de la siguiente forma:

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

$t = 0, 1, ...,$ donde $\{\xi_t\}$ es una sucesión de variables aleatorias i.i.d. tomando valores en un espacio de Borel $S \subset \mathbb{R}^k$ ($k \geq 1$ es un entero) con función de densidad Δ. $L : X \times S \to X$ se supone que es una función medible conocida.

En lo referente a la diferenciabilidad de la política óptima, se encuentran los trabajos de Santos ([43] y [44]). Estos trabajos se encargan de modelos de crecimiento económico en sistemas deterministas ([43]) y estocásticos ([44]). (Con respecto a problemas de control aplicados a Economía véase la Observación 2.2.1). Las principales hipótesis en sus modelos son las siguientes:

Hipótesis (S)

1. $\mathbb{K} \subset \mathbb{R}^l \times \mathbb{R}^m$ es convexo y abierto (l y m son enteros positivos);
2. $X = A = A(x), x \in X$;
3. c es acotada y diferenciable, de clase C^2 con primera y segunda derivadas acotadas. Además, existe $\eta > 0$ tal que

$$Z(x,a) := c(x,a) + \eta/2 \left\| a \right\|^2,$$

$(x,a) \in \mathbb{K}$, es convexa.

4. $f(x) \in int(X), x \in X$;
5. $L(\cdot,\cdot,s) \in C^2(\mathbb{K})$, para cada $s \in S$, con derivadas acotadas.

Teorema (S): *Bajo las Hipótesis (S), se tiene que* $V \in C^2(int(X); R)$ *y* $f \in C^1(int(X), A)$.

La idea de la prueba de la diferenciabilidad de la política es la siguiente. Primero, Santos observa que la diferenciabilidad de primer orden de la política óptima, es equivalente, con sus hipótesis, a probar la diferenciabilidad de segundo orden de la función de valores óptimos. Entonces el método de prueba consiste en mostrar que la derivada de la política óptima está definida por la política óptima de un problema de control auxiliar, el cual se obtiene como una expansión cuadrática del problema original a través de la trayectoria óptima.

La diferenciabilidad de orden superior de la función de valores óptimos, V, fue analizada por Blume et.al ([11]). En este artículo se analizan modelos de control con las condiciones siguientes.

Hipótesis (B)

1. $X \subset \mathbb{R}^n, A \subset \mathbb{R}^m$ convexos (n y m son enteros positivos);
2. $A = A(x)$, $x \in X$ compacto;
3. $c \in C^2(int(\mathbb{K}), \mathbb{R})$, y c es acotado y convexo en $\mathbb{K}$;
4. $c(x,\cdot)$ estrictamente convexo para cada $x \in X$;

5. $L(\cdot,\cdot,s) \in C^2(int(\mathbb{K}),X)$, para cada $s \in S$. Además, $L(x,a,s) = G(F(x,a),s)$, donde $G : X \times S \to X$ y $F : \mathbb{K} \to X$ son funciones dadas, diferenciables de orden dos en sus respectivos dominios.

6. $S = S_1 \times S_2$ (S^1 y S^2 son espacios euclidianos tal que la suma de sus dimensiones es igual a k) y existe $\widehat{L} : \mathbb{K} \times S^1 \to X$ tal que $L(x,a,(\widehat{s}_1,\widehat{s}_2)) = \widehat{L}(x,a,\widehat{s}_1)$, $(x,a) \in \mathbb{K}$ y $(\widehat{s}_1,\widehat{s}_2) \in S$.

Teorema (B): *Bajo las Hipótesis (B) se garantiza que* $V \in C^2(int(X);R)$ *y* $f \in C^1(int(X),A)$.

En resumen, las condiciones existentes en la literatura para probar la diferenciabilidad en PCMs están basadas, en general, en hipótesis de convexidad y acotabilidad en las componentes del modelo de control. Además, en los trabajos mencionados no se muestra la aplicabilidad de la diferenciabilidad, por ejemplo, en la solución de problemas de control. En este trabajo se presenta como consecuencia de la diferenciabilidad una fórmula para determinar la solución óptima del problema de control.

Los resultados en la literatura referentes a la Ecuación de Euler, generalmente, son tratados para modelos económicos (véase [12], [22], [30], [34], [36], [37], [46] y [49]). A continuación se presentan las condiciones y el teorema que garantiza la validez de la Ecuación de Euler.

Hipótesis (EE)

1. $X = A = A(x) \subset \mathbb{R}$, $x \in X$, convexos.

2. $A(x)$ es un subconjunto abierto de A para cada $x \in X$

3. La dinámica del sistema está dada por (3.1), se supone que $F \in C^2(int(\mathbb{K}),X)$ y que $F_a(x,a) \neq 0, (x,a) \in \mathbb{K}$.

4. $c \in C^2\left(int(\mathbb{K}),\mathbb{R}\right)$.

5. $W \in C^2(int(X),A), g \in C^1(int(X),A)$.

Teorema (EE): *Bajo las Hipótesis (EE) se garantiza que la Ecuación de Euler es valida, i.e.*

$$c_a(x,g(x)) + \alpha M(F(x,g(x)),g(F(x,g(x)))F_a(x,g(x)) = 0,$$

$x \in X$, donde $M(x,a) = (c_x - c_a F_x F_a^{-1})\,(x,a), (x,a) \in \mathbb{K}$ y g es la política óptima.

En algunos casos la EE es resuelta proponiendo la forma funcional de la política óptima ([23], [36] y [37]).

Finalmente, con respecto al análisis de problemas de control estocásticos vía problemas de control determinista, en el aspecto analizado en la tesis, no se encontraron antecedentes referentes a ello.

Problemas Deterministas Perturbados

0.5. Ecuación de Euler

Este capítulo está basado en el artículo [20]. En el desarrollo de este capítulo se utiliza la notación dada en el Capítulo 2. Por ejemplo, recuérdese que W y g denota la función de valores óptimos y la política óptima. También W_n denotan las funciones de iteración de valores y g_n son los minimizadores (véase (2.12) y 13).

En esta sección, una ecuación funcional, conocida en la literatura de Economía como Ecuación de Euler (EE) ([2], [22], [23], [34], [36], [37], [45], [46], [47], [48], [49]) será deducida. La EE permite dar una caracterización de la derivada de W_n (véase (2.12)). Usando esta información en diferentes situaciones (véase Sección 4.1.1), es posible determinar W_n explícitamente (integrando su derivada) y, después, determinar W tomando el límite de W_n, cuando n tiende a infinito.

Sean R, Y y Z espacios euclidianos.

Para cualquier conjunto $B \subset R$, un punto $x \in B$ es llamado **punto interior** de B, si existe un conjunto abierto U tal que $x \in U \subset B$. El **interior** de B es el conjunto de todos los puntos interiores de B y se denotará por $int(B)$.

Sea $\widehat{R}$ un subconjunto de $R \times Y$. Sea $\theta : \widehat{R} \to Z$ una función y supóngase que $\theta = \theta(\upsilon, \eta)$ es dos veces diferenciable. Las derivadas parciales con respecto a las variables υ y η son denotadas por θ_ν y θ_η, respectivamente. La notación para las segundas derivadas de θ con respecto a υ y η son $\theta_{\upsilon\upsilon}$ y $\theta_{\eta\eta}$, respectivamente.

Además, para $\phi = 1, 2$, se define $C^\phi(R; Z) := \{\theta : R \to Z \mid$las primeras ϕ derivadas de θ existen y son continuas$\}$.

La transpuesta de un vector γ, denotada por γ^{tr} es considerada como un vector vertical.

Consideremos un MCM: $(X,\ A,\ \{A(x) : x \in X\}, Q_F, c)$, fijo (véase 11) y sea

$$G^n(x,a) := c(x,a) + \alpha W_{n-1}(F(x,a)), \tag{16}$$

$n = 1, 2, ...,\ (x,a) \in \mathbb{K}$.

Como es usual, para cada $x \in X$, y $n = 1, 2, ...,$ la **condición de primer orden** para la optimalidad de $G^n(x,\cdot)$ está definida por los puntos $\widehat{a} \in int(A(x))$ que satisfacen: $G^n_a(x,\widehat{a}) = 0$.

Hipótesis 2

a) $c \in C^2(int(\mathbb{K}); \mathbb{R})$.

b) $F \in C^2(int(\mathbb{K}); X)$.

c) Para cada $n = 1, 2, ...,$ y $x \in X$, $G^n_{aa}(x,\cdot)$ es positiva definida (véase la Observación 4.1.2 b)).

d) Para cada $n = 1, 2, ...,\ g_n(x) \in int(A(x))$, para todo $x \in X$ (véase la Observación 4.1.2 a)).

Remark 6 *a) Cuando $A(x) \subset A$, $x \in X$ es abierto, entonces, claramente, $g_n(x) \in int(A(x))$, $x \in X$. Además, en [22] p. 599 y [36] p. 155, para cierta clase de problemas de crecimiento económico, se presentan condiciones las cuales garantizan que los minimizadores g_n se encuentran en el interior del conjunto de restricciones.*

b) Obsérvese que la Hipótesis 2 c) implica que $G^n(x,\cdot)$ es una función estrictamente convexa, para cada $x \in X$ y $n = 1, 2, ...$([22] p.260), y nótese que g_n es única para $n = 1, 2, ...$

Lemma 3 *Bajo la Hipótesis 2, para cada $n = 1, 2...,$ $W_n \in C^2(int(X); \mathbb{R})$ y $g_n \in C^1(int(X); A)$.*

Proof. La demostración se hará por inducción. Considérese $n = 1$. En este caso,

$$W_1(x) = \min_{a \in A(x)} c(x,a),$$

$x \in X$. Como $g_1(x) \in int(A(x))$, $x \in X$, el minimizador g_1 satisface la condición de primer orden para la optimalidad de c, i.e. $c_a(x, g_1(x)) = 0$, $x \in X$. Entonces, usando el Teorema de la Función Implícita (véase Apéndice B) y la Hipótesis 2 a) y 2 c), resulta que $g_1 \in C^1(int(X); A)$. Por otro lado, $W_1(x) = c(x, g_1(x))$, $x \in X$, entonces

$$W_1'(x) = c_x(x, g_1(x)) + c_a(x, g_1(x))g_1'(x), \tag{17}$$

$x \in X$. Usando la condición de primer orden para la optimalidad de c, en (17) se obtiene que $W_1'(x) = c_x(x, g_1(x))$, $x \in X$. Como $c \in C^2(int(\mathbb{K}); \mathbb{R})$ y $g_1 \in C^1(int(X); A)$, $W_1 \in C^2(int(X); \mathbb{R})$. Ahora, supóngase que $W_{n-1} \in C^2(int(X); \mathbb{R})$, para algún entero $n > 1$. Así, combinando este hecho con la Hipótesis 2 a) y 2 b), resulta que $G^n \in C^2(int(\mathbb{K}); \mathbb{R})$. Por otro lado, la condición de primer orden para la optimalidad de G^n está dada por

$$G_a^n(x, g_n(x)) = 0,$$

$x \in X$ (véase (2.13) y (16)). Por la Hipótesis 2 c) y el Teorema de la Función Implícita (véase Apéndice B), se obtiene que $g_n \in C^1(int(X); A)$. Entonces $W_n'(x) = G_x^n(x, g_n(x)) + G_a^n(x, g_n(x))g_n'(x)$, $x \in X$. Usando la condición de primer orden en la última igualdad resulta que $W_n'(x) = G_x^n(x, g_n(x))$, $x \in X$. Como $G^n \in C^2(int(\mathbb{K}); \mathbb{R})$ y $g_n \in C^1(int(X); A)$, se concluye que $W_n \in C^2(int(X); \mathbb{R})$.
■

Hipótesis 3

a) $X, A \subset \mathbb{R}^p$, p es un entero positivo fijo.

b) Existe la inversa de la matriz F_a, la cual será denotada por F_a^{-1}.

c) La función $H(x, a) := (c_x - c_a F_a^{-1} F_x^{tr})(x, a)$, $(x, a) \in \mathbb{K}$ es invertible en la segunda variable, con inversa $H^{-1} : X \times H[\mathbb{K}] \to A$, donde $H[\mathbb{K}] := \{H(x, a) : (x, a) \in \mathbb{K}\}$.

Remark 7 *Nótese que la función H en la Hipótesis 3 c) está dada en términos de las derivadas de la función de costo c, de la dinámica del sistema F y de su inversa F^{-1}.*
En el Teorema siguiente, se dará una relación para las derivadas de las funciones de iteración de valores y de los minimizadores en términos de H y

H^{-1}.
En el Ejemplo 4.1.6, H y H^{-1} serán explícitamente calculadas y la Hipótesis 3 será verificada directamente.

Theorem 4 *Bajo la Hipótesis 2 y 3, la derivada de las funciones de valor W_n, $n = 1, 2, ...$, satisfacen la* ***Ecuación de Euler****:*

$$\begin{aligned} & \alpha W'_{n-1}\left(F\left(x, H^{-1}(x, W'_n(x))\right)\right) F_a(x, H^{-1}(x, W'_n(x))) \\ = \; & -c_a(x, H^{-1}(x, W'_n(x))), \end{aligned} \tag{18}$$

$x \in X$.

Proof. Obsérvese que para cada entero n, la condición de primer orden para la optimalidad de G^n es

$$\alpha W'_{n-1}(F(x, g_n(x)))F_a(x, g_n(x)) + c_a(x, g_n(x)) = 0, \tag{19}$$

$x \in X$, $g_n \in \mathbb{F}$ y $n = 1, 2, ...$ Por otro lado, derivando (2.13) tenemos que,

$$\begin{aligned} W'_n(x) \; = \; & c_x(x, g_n(x)) + c_a(x, g_n(x))g'_n(x) \\ & + \alpha W'_{n-1}(F(x, g_n(x))) \left[F_x(x, g_n(x)) + F_a(x, g_n(x))g'_n(x)\right], \end{aligned}$$

$x \in X$. Ahora, usando (19) junto con la definición de H, resulta que $W'_n(x) = H(x, g_n(x))$, $x \in X$. Entonces, por la invertibilidad de H (véase Hipótesis 3 c)) se obtiene que

$$g_n(x) = H^{-1}(x, W'_n(x)), \tag{20}$$

$x \in X$. Sustituyendo (20) en (19), el resultado se sigue. ■

0.5.1. Ejemplos

Example 1 *Sea $\delta \in (0, 1)$ un número fijo. Sean $X = A = [0, 1)$, y $A(0) = \{0\}$, $A(x) = (0, x^\delta]$, $x \in X$, $x \neq 0$. La dinámica del sistema está dada por la ecuación en diferencias $x_{t+1} = x_t^\delta - a_t$, $t = 0, 1, ...$, y la función de recompensa está dada por $r(x, a) = \ln a$, si $x \in (0, 1)$, $a \in (0, x^\delta]$, y $r(0, 0) = -\infty$. Supóngase que $0 < \alpha\delta < 1$.*

Nótese que en este ejemplo $Q(\{x^\delta - a\}\,|x, a) = 1$ si $x \in (0, 1)$, $a \in (0, x^\delta]$, y $Q(\{0\}\,|0, 0) = 1$. Además, para cada $n = 1, 2, ...$, $W_n(0) = -\infty$, $g_n(0) = 0$, $W(0) = -\infty$, y $g(0) = 0$. (Obsérvese que este ejemplo puede ser considerado como un caso de costos no negativos, haciendo $c(x, a) = -r(x, a) = -\ln a \geq 0$, $(x, a) \in \mathbb{K}$).

Lemma 5 *Para el ejemplo 4.1.6,*

a) W_n *es cóncava para* $n = 1, 2, ...;$

b) $W_n \in C^2(int(X); \mathbb{R})$ *y* $g_n \in C^1(int(X); A)$ *para cada* $n = 1, 2, ...;$

c) Las funciones de valor W_n, $n = 2, 3, ...$ *satisfacen la ecuación funcional*

$$\frac{W_n'(x)}{\delta x^{\delta-1}} = \alpha W_{n-1}\left(x^{\delta-1}\frac{W_n'(x)x-\delta}{W_n'(x)}\right), \tag{21}$$

$x \in (0,1)$. *Además,*

$$W_n'(x) = (\delta/x)\sum_{k=0}^{n-1}(\alpha\delta)^k, \tag{22}$$

$n = 1, 2, ..., 0 < x < 1.$

Proof.

a) Similar a la prueba del Lema 6.2 en [16].

b) La prueba será hecha por inducción. Para $n = 1$, resulta que

$$W_1(x) = \max_{a\in(0,x^\delta]} \ln a = \delta \ln x,$$

$x \in (0,1)$. Entonces, $W_1 \in C^2(int(X); \mathbb{R})$ y $g_1 \in C^1(int(X); A)$. Supóngase que $W_n \in C^2(int(X); \mathbb{R})$ y $g_n \in C^1(int(X); A)$, para $n > 1$. Entonces,

$$\begin{aligned} W_{n+1}(x) &= \max_{a\in(0,x^\delta]}\left[\ln a + \alpha W_n(x^\delta - a)\right], \\ &= \max_{a\in(0,x^\delta)}\left[\ln a + \alpha W_n(x^\delta - a)\right], \end{aligned}$$

$x \in (0,1)$, ya que $W_{n+1}(x^\delta - x^\delta) = W_{n+1}(0) = -\infty$. Así, la política óptima $g_{n+1}(x) \in (0, x^\delta)$, $x \in X$. Por otro lado, $G_{aa}^{n+1}(x,a) = -a^{-2} + \alpha W_n''(x^\delta - a) < 0, (x,a) \in \mathbb{K}$, ya que W_n es cóncava para $n = 1, 2, ...$ Entonces el Lema 4.1.7 b) se sigue del Lema 4.1.3.

c) Como $H(x,a) = \left(\delta x^{\delta-1}\right)/a, (x,a) \in \widehat{\mathbb{K}} := \{(x,a) : x \in (0,1), a \in (0,x^{\delta})\}$ y $H^{-1}(x,u) = \left(\delta x^{\delta-1}\right)/u, x \in (0,1)$ y $u \in (\delta/x,\infty)$, la Hipótesis 2 se cumple y, usando el Teorema 4.1.5, (21) se sigue.

Ahora, para $n=1$, se obtiene que

$$W_1(x) = \max_{a \in (0,x^{\delta}]} \ln a = \delta \ln x,$$

$x \in (0,1)$, y $W_1'(x) = \delta/x, x \in (0,1)$. Para $n > 1$, supóngase que $W_{n-1}'(x) = (\delta/x)\sum_{k=0}^{n-2}(\alpha\delta)^k$, $x \in (0,1)$. Entonces, sustituyendo W_{n-1}' en (21), resulta (22).

■

Lemma 6 *Para el Ejemplo 4.1.6, la función de valores óptimos y la política óptima son*

$$W(x) = \frac{\delta}{1-\alpha\delta}\ln x + M, \ \ y \tag{23}$$

$$g(x) = (1-\alpha\delta)x^{\delta}, \tag{24}$$

respectivamente, donde $x \in (0,1)$ *y*

$$M = \frac{1}{1-\alpha}\left[\ln(1-\alpha\delta) + \frac{\alpha\delta}{1-\alpha\delta}\ln(\alpha\delta)\right]. \tag{25}$$

Proof. Por Lema 4.1.7, resulta que

$$W_n(x) = \delta\sum_{k=0}^{n-1}(\alpha\delta)^k \ln x + M_n,$$

$n = 1,2,...,x \in (0,1)$, y M_n es una constante real por determinar. La sucesión $\{M_n\}$ es convergente a un número real M, ya que $W_n(x) \to W(x)$, $x \in (0,1)$, y $\sum_{k=0}^{\infty}(\alpha\delta)^k = 1/(1-\alpha\delta)$. Entonces, (23) se sigue. Sustituyendo (23) en (14), resulta que

$$\frac{\delta}{1-\alpha\delta}\ln x + M = \max_{a\in(0,x^{\delta})}\left[\ln a + \frac{\alpha\delta}{1-\alpha\delta}\ln(x^{\delta}-a) + \alpha M\right], \tag{26}$$

$x \in (0,1)$. La condición de primer orden para la optimalidad del lado derecho de (26) es $1/g(x) = \alpha\delta/\left[(1-\alpha\delta)\left(x^{\delta}-g(x)\right)\right]$, $x \in (0,1)$. Resolviendo la ecuación anterior para g, se obtiene que

$$g(x) = (1-\alpha\delta)x^{\delta},$$

$x \in (0,1)$. Sustituyendo g en (26), se llega a

$$\frac{\delta}{1-\alpha\delta}\ln x + M = \ln(1-\alpha\delta) + \delta \ln x + \frac{\alpha\delta}{1-\alpha\delta}\ln(\alpha\delta) + \frac{\alpha\delta^2}{1-\alpha\delta}\ln x + \alpha M,$$

$x \in (0,1)$. Entonces, $M = \ln(1-\alpha\delta) + \ln(\alpha\delta)\alpha\delta/\left(1-\alpha\delta\right) + \alpha M$. Resolviendo para M, se obtiene (25). ■

0.6. Sistemas Deterministas Perturbados

Considérese un problema determinista con espacio de estados X, espacio de controles A ($X \subseteq \mathbb{R}^p$ y $A \subseteq \mathbb{R}^m$, p, $m \geq 1$ son enteros), con conjunto de acciones admisibles $A(x) \subset A$, $x \in X$. Supóngase que la dinámica del sistema se encuentra determinada por la ecuación en diferencias

$$x_{t+1} = F(x_t, a_t),$$

$t = 0, 1, ...$, donde $F : \mathbb{K} \to X$ es una función medible conocida, y la función de costos $c : \mathbb{K} \to \mathbb{R}$ es una función medible dada.

Considérese un sistema de control estocástico con el mismo: espacio de estados X, espacio de controles A, conjunto de acciones admisibles $A(x)$, $x \in X$, y la función de costo c, pero con la siguiente dinámica:

$$x_{t+1} = L(F(x_t, a_t), \xi_t), \tag{27}$$

$t = 0, 1, ...$, donde $\{\xi_t\}$ es una sucesión de variables aleatorias i.i.d. tomando valores en un espacio de Borel $S \subset \mathbb{R}^k$ ($k \geq 1$ es un entero) con función de densidad Δ. $L : X \times S \to X$ se supone que es una función medible conocida.

Obsérvese que en este caso la ley de transición Q está dada por

$$Q(B\,|x, a) := Q_L(B\,|x, a) = \int I_B(L(F(x,a), s))\Delta(s)ds, \tag{28}$$

$B \in \mathcal{B}(X)$ y $(x, a) \in \mathbb{K}$. También, nótese que los conjuntos $\mathbb{F}$ (conjunto de políticas estacionarias, véase el Capítulo 2) y $\mathbb{M}$ (conjunto de políticas Markovianas, véase el Capítulo 2) son los mismos, tanto para el sistema determinista $(X, A, \{A(x) : x \in X\}, Q_F, c)$ como para el sistema estocástico $(X, A, \{A(x) : x \in X\}, Q_L, c)$.

Un proceso de control de Markov para el cual la ley de transición está dada por (11) será referido como un **sistema de control determinista**. En el caso de que la ley de transición es dada por (27), será referida como un **sistema de control estocástico.**

Sea $(X, A, \{A(x) : x \in X\}, Q_F, c)$ un sistema de control determinista fijo con el correspondiente sistema de control estocástico $(X, A, \{A(x) : x \in X\}, Q_L, c)$. Además, sea W y g la función de valores óptimos y la política óptima para el sistema de control determinista, respectivamente.

Hipótesis 4

a) La función de valores óptimo W y la política óptima g son conocidas para el sistema de control determinista.

b) Existe una función medible no negativa $h : S \to \mathbb{R}$, tal que

$$\int W(y)Q_L(dy\,|x, a) = W(F(x, a)) + \int h(s)\Delta(s)ds,$$

$(x, a) \in \mathbb{K}$.

c) Existen $\vartheta \in [0, 1)$ y $N \geq 0$ tal que

$$\int W(y)Q_L(dy\,|x, a) \leq \vartheta W(x) + N,$$

$(x, a) \in \mathbb{K}$.

Remark 8 *a) La Hipótesis 4 a) ha sido analizada en la primera parte del capítulo. W y g pueden ser obtenidas por medio de la Ecuación de Euler o, directamente, por medio del algoritmo de iteración de valores.*

b) La Hipótesis 4 b) puede ser verificada directamente, cuando W es conocida.

c) Nótese que si c es acotada con cota superior $\widehat{N}$, entonces la Hipótesis 4 c) se cumple, con $\vartheta = 0$ y $N = \widehat{N}/(1 - \alpha)$.

Definition 6 *$M(X)^+$ denota el conjunto de las funciones medibles no negativas definidas en X, y, para cada $U \in M(X)^+$, TU es la función definida en X como*

$$TU(x) := \min_{a \in A(x)} \left[c(x, a) + \alpha \int U(y)Q_L(dy\,|x, a) \right],$$

$x \in X$. T es conocido como el ***Operador de Programación Dinámica*** *asociado a $(X, A, \{A(x) : x \in X\}, Q_L, c)$.*

Definition 7 *Sea $\mathcal{R} \subset M(X)^+$, $\mathcal{R}$ es invariante con respecto al operador de programación dinámica T, si para cada $U \in \mathcal{R}$, $TU \in \mathcal{R}$.*

Lemma 7 *Sea $\mathcal{L} := \{W + \lambda : \lambda \in \mathbb{R}\}$. Entonces bajo la Hipótesis 4, $\mathcal{L}$ es invariante con respecto al operador de programación dinámica T.*

Proof. Sea $U \in \mathcal{L}$. Entonces, $U(x) = W(x) + \lambda$, $x \in X$ para algún $\lambda \in \mathbb{R}$. Se sigue que

$$
\begin{aligned}
TU(x) &= \min_{a\in A(x)} \left[c(x,a) + \alpha \int U(y) Q_L(dy\,|x,a) \right] \\
&= \min_{a\in A(x)} \left[c(x,a) + \alpha E[U\left(L(F(x,a),\xi)\right)] \right], \\
&= \min_{a\in A(x)} \left[c(x,a) + \alpha E[W\left(L(F(x,a),\xi)\right) + \lambda] \right], \\
&= \min_{a\in A(x)} \left[c(x,a) + \alpha W(F(x,a)) \right] + \overline{\lambda}, \\
&= W(x) + \overline{\lambda},
\end{aligned}
$$

$x \in X$, donde $\overline{\lambda} = \alpha\left(E\left[h(\xi)\right] + \lambda\right)$, y

$$
E[h(\xi)] = \int h(s)\Delta(s)ds.
$$

Por lo tanto, $TU \in \mathcal{L}$. ■

Lemma 8 *Bajo las Hipótesis 4, $T\overline{W} = \overline{W}$, donde*

$$
\overline{W}(x) := W(x) + \frac{\alpha}{1-\alpha} E\left[h(\xi)\right],
$$

$x \in X$.

Proof. Nótese que en la prueba del Lema anterior, $TU = U$ si y solo si $\lambda = \left[\alpha/(1-\alpha)\right] E\left[h(\xi)\right]$. Entonces $T\overline{W} = \overline{W}$. Esto concluye la prueba del Lema 4.2.5. ■

Theorem 9 *Bajo la Hipótesis 4, la función de valores óptimos del sistema de control estocástico es $\overline{W}$, i.e. $V = \overline{W}$. Además, $V(x) = v(g,x)$, $x \in X$ (así, g es una política óptima para el sistema de control estocástico).*

Proof. Como $T\overline{W} = \overline{W}$ y g es una política óptima para el problema determinista, resulta que

$$\overline{W}(x) = c(x, g(x)) + \alpha \int \overline{W}(y) Q_L(dy\,|x, g(x))\,,$$

$x \in X$. Iterando la última igualdad, resulta que

$$\overline{W}(x) = E_x^g \left[\sum_{t=0}^{n-1} \alpha^t c(x_t, a_t) \right] + \alpha^n E_x^g \overline{W}(x_n),$$

para todo $n \geq 1$, $x \in X$. Como $\overline{W}$ es no negativa,

$$\overline{W}(x) \geq E_x^g \left[\sum_{t=0}^{n-1} \alpha^t c(x_t, a_t) \right],$$

$n \geq 1$, $x \in X$. Entonces, cuando $n \to \infty$ en la última desigualdad, se obtiene que

$$\overline{W}(x) \geq v(g, x), \tag{29}$$

$x \in X$. Por lo tanto, $\overline{W} \geq V$.

Por otro lado, sea $\pi \in \Pi$ y $x \in X$. Entonces usando la propiedad tipo Markov (véase [33] p.p. 19 y 20) y que $T\overline{W} = \overline{W}$, resulta que para cada $t = 0, 1, ...,$

$$\begin{aligned}
E_x^\pi[\alpha^{t+1}\overline{W}(x_{t+1})\,|h_t, a_t] &= \alpha^t \left[\alpha \int_X \overline{W}(y) Q_L(dy\,|x_t, a_t) \right], \\
&= \alpha^t \left[c(x_t, a_t) + \alpha \int_X \overline{W}(y) Q_L(dy\,|x_t, a_t) - c(x_t, a_t) \right], \\
&\geq \alpha^t \left[\overline{W}(x_t) - c(x_t, a_t) \right].
\end{aligned}$$

Así, $\alpha^t c(x_t, a_t) \geq -E_x^\pi \left[\alpha^{t+1}\overline{W}(x_{t+1}) - \alpha^t \overline{W}(x_t) \middle|\, h_t, a_t \right]$. Tomando esperanzas y sumando sobre $t = 0, ..., n-1$, se obtiene que

$$E_x^\pi \left[\sum_{t=0}^{n-1} \alpha^t c(x_t, a_t) \right] \geq \overline{W}(x) - \alpha^n E_x^\pi \left[\overline{W}(x_n) \right],$$

$x \in X$. Como $\overline{W}(x) = W(x) + [\alpha/(1-\alpha)]\, E\,[h(\xi)]$, $x \in X$,

$$E_x^\pi \left[\sum_{t=0}^{n-1} \alpha^t c(x_t, a_t) \right] \geq \overline{W}(x) - \alpha^n E_x^\pi \left[W(x_n) \right] - \frac{\alpha^{n+1}}{1-\alpha} E\left[h(\xi) \right], \tag{30}$$

$x \in X$. Por otro lado, usando la Hipótesis 4 c), resulta que

$$\begin{aligned} E_x^\pi[W(x_n)\,|h_{n-1}, a_{n-1}] &= \int W(y) Q_L(dy\,|x_{n-1}, a_{n-1}), \\ &\leq \vartheta W(x_{n-1}) + N. \end{aligned}$$

Por lo tanto, $E_x^\pi\left[W(x_n)\right] \leq \vartheta E_x^\pi\left[W(x_{n-1})\right] + N$. Iterando esta desigualdad se llega a que

$$E_x^\pi\left[W(x_n)\right] \leq \vartheta^n W(x) + (1 + \vartheta + ... + \vartheta^{n-1})N.$$

Entonces se obtiene por (30) que

$$E_x^\pi\left[\sum_{t=0}^{n-1} \alpha^t c(x_t, a_t)\right] \geq \overline{W}(x) - (\alpha\vartheta)^n W(x) - \alpha^n N \sum_{k=0}^{n-1} \vartheta^k - \frac{\alpha^{n+1}}{1-\alpha} E\left[h(\xi)\right], \tag{31}$$

$x \in X$. Finalmente, dejando n tender a ∞ en (31)

$$E_x^\pi\left[\sum_{t=0}^{\infty} \alpha^t c(x_t, a_t)\right] \geq \overline{W}(x), \tag{32}$$

$x \in X$ (recuérdese que $\vartheta \in [0, 1)$ y obsérvese que $0 \leq \alpha\vartheta < 1$).

Ahora, de (32) y ya que π y x son arbitrarios, $V \geq \overline{W}$.

Por lo tanto, $\overline{W} = V$. Ahora, combinando (5) y (29) se sigue que $V(x) = v(g, x)$, $x \in X$. Por lo tanto g es óptima para el sistema de control estocástico. Esto concluye la prueba del Teorema 4.2.6. ■

0.6.1. Ejemplos

Example 2 *Considérese el modelo determinista trabajado en el Ejemplo 4.1.6 con la siguiente dinámica:* $x_{t+1} = \left(x_t^\delta - a_t\right)\xi_t$, $t = 0, 1, ...,$ *donde* $\{\xi_t\}$ *es una sucesión de variables aleatorias i.i.d. tomando valores en* $S = (0, \infty)$. *Sea* $\kappa := E(\ln \xi)$, *donde* ξ *es un elemento genérico de la sucesión* $\{\xi_t\}$. *Supóngase que* κ *es finito.*

Lemma 10 *Para el Ejemplo 4.2.7, la función de valores óptimos* V *y la política óptima* f *están dadas por*

$$V(x) = \frac{\delta}{1-\alpha\delta}\ln x + \frac{1}{1-\alpha}\left[\ln(1-\alpha\delta) + \frac{\alpha\delta}{1-\alpha\delta}\ln(\alpha\delta)\right] + \frac{\alpha\delta}{(1-\alpha\delta)(1-\alpha)}\kappa,$$

y $f(x) = (1-\alpha\delta)\,x^\delta$, $x \in (0, \infty)$, *respectivamente.*

Proof. *En este caso, usando la solución determinista (véase Ejemplo 4.1.6), resulta que*

$$\begin{aligned} E\left[W(\left(x^{\delta}-a\right)\xi)\right] &= E\left[\frac{\delta}{1-\alpha\delta}\ln\left[\left(x^{\delta}-a\right)\xi\right]+M\right], \\ &= \frac{\delta}{1-\alpha\delta}\ln\left(x^{\delta}-a\right)+M+\frac{\delta}{1-\alpha\delta}\kappa, \\ &= W(x^{\delta}-a)+E\left[h(\xi)\right], \end{aligned}$$

$(x,a)\in\mathbb{K}$ *y* $h(s)=[\delta/(1-\alpha\delta)]\ln s,\ s\in S$ *. La Hipótesis 4 c) puede ser probada en la forma siguiente:*

$$\begin{aligned} \int W((x^{\delta}-a)s)\Delta(s)ds &= \int\left(\frac{\delta}{1-\alpha\delta}\ln\left[\left(x^{\delta}-a\right)s\right]+M\right)\Delta(s)ds, \\ &= \frac{\delta}{1-\alpha\delta}\ln\left(x^{\delta}-a\right)+M+\frac{\delta}{1-\alpha\delta}\kappa, \\ &\leq \frac{\delta^{2}}{1-\alpha\delta}\ln x+M+\frac{\delta}{1-\alpha\delta}\kappa, \\ &= \delta\left[\frac{\delta}{1-\alpha\delta}\ln x+M\right]+\frac{\delta}{1-\alpha\delta}\kappa+M(1-\delta), \end{aligned}$$

$(x,a)\in\widehat{\mathbb{K}}$*. La última desigualdad es debido a que* $0<x^{\delta}-a\leq x^{\delta}$ *y* $0<\delta<1$*. Por lo tanto, Hipótesis 4 se cumple con* $\vartheta=\delta$ *y* $N=\delta\kappa/(1-\alpha\delta)+M(1-\delta)$*.* ■

Ahora, el siguiente ejemplo ilustra que la teoría desarrollada se cumple para modelos no convexos.

Example 3 *Sea* ζ *un número positivo fijo. Sea* $X=(0,\zeta)$*,* $A=A(x)=[-\pi,\pi/2],x\in X$*. La función de costo y la dinámica del sistema están dadas por,* $c(x,a)=x+\sin a+1$*, y* $x_{t+1}=\theta x_{t}$*,* $t=0,1,2,...$ *respectivamente, donde* $0<\theta<1$*.*

Lemma 11 *Para el Ejemplo 4.2.9, las funciones de iteración de valores* $W_{n},n=1,2,...$*, están dadas por*

$$W_{n}(x)=(1+\alpha\theta+...+(\alpha\theta)^{n-1})x, \tag{33}$$

$x\in X$*.*

Proof. *La prueba será hecha por inducción. Para* $n = 1$*,*

$$W_1(x) = \min_{a \in A(x)} [x + \sin a + 1] = x,$$

$x \in X$*. Supóngase que para* $n > 1$*,* $W_n(x) = (1 + \alpha\theta + ... + (\alpha\theta)^{n-1})x$*. Como el mínimo de la función* $\eta(a) = \sin a$*,* $a \in [-\pi, \pi/2]$ *es únicamente alcanzado en* $a^* = -\pi/2$*, entonces* $g_{n+1}(x) = -\pi/2$ *y*

$$\begin{aligned} W_{n+1}(x) &= \min_{a \in A(x)} [x + \sin a + 1 + \alpha W_n(\theta x)], \\ &= \min_{a \in A(x)} \left[x + \sin a + 1 + \alpha\theta(1 + \alpha\theta + ... + (\alpha\theta)^{n-1})x\right], \\ &= (1 + \alpha\theta + ... + (\alpha\theta)^n)x, \end{aligned}$$

$x \in X$*. Esto completa la prueba del Lema 4.2.10.* ■

Lemma 12 *La función de valores óptimos y la política óptima están dadas por* $W(x) = x/(1 - \alpha\theta)$*,* $g(x) = -\pi/2$*,* $x \in X$*, respectivamente.*

Proof. *Usando el Lema 4.2.10 (y el hecho que* $0 < \alpha\theta < 1$*), cuando n tiende a* ∞ *en (33), se obtiene que* $W(x) = [1/(1 - \alpha\theta)]\, x$*,* $x \in X$*. Finalmente, sustituyendo* W *en (14), resulta que* $g(x) = -\pi/2$*,* $x \in X$*.* ■

Ahora, supóngase que la dinámica del sistema es perturbada en la forma siguiente: $x_{t+1} = \theta x_t + \xi_t$*,* $t = 0, 1, ...$*, donde* $\{\xi_t\}$ *es una sucesión de variables aleatorias i.i.d. tomando valores en* $S = (0, (1 - \theta)\zeta)$*. Se supone que estos elementos tienen media finita* μ*.*

Lemma 13 *Para la versión estocástica del Ejemplo 4.2.9, la función de valores óptimos* V *y la política óptima* f *están dadas por*

$$V(x) = \frac{1}{1 - \alpha\theta} x + \frac{\alpha}{1 - \alpha}\mu$$

y $f(x) = -\pi/2$*,* $x \in X$*, respectivamente.*

Proof. *Usando la solución determinista, se tiene que*

$$\begin{aligned} E[W(\theta x + \xi)] &= E\left[\frac{\theta x}{1 - \alpha\theta} + \xi\right], \\ &= W(\theta x) + E[h(\xi)], \end{aligned}$$

$x \in X$ *y* $h(s) = s$*,* $s \in S$*. Nótese que, en este caso, la función de costos es acotada. Entonces la Hipótesis 4 c) se satisface. Por lo tanto, por el Teorema 4.2.6, el resultado se sigue.* ■

0.7. Conclusiones del Capítulo

Recuérdese que W denota la función de valores óptimos y g la política óptima, correspondientes al problema de control óptimo determinista. De igual forma $W_n, n \geq 0$ denotan las funciones de iteración de valores, y $g_n, n \geq 0$ denotan los minimizadores correspondientes.

En este capítulo se presentaron condiciones en el modelo de control para garantizar la diferenciabilidad de clase $C^2(int(X); \mathbb{R})$ de W y la diferenciabilidad de clase $C^1(int(X); A)$ de g. Dichas condiciones de diferenciabilidad fueron determinadas dando condiciones en el modelo de control y la prueba de la diferenciabilidad, se basó en el Teorema de la Función Implícita. Una vez que se tuvieron las propiedades de diferenciabilidad en W y g, se determinó una ecuación funcional, conocida como Ecuación de Euler (EE), la cual se dió en términos de las derivadas de las funciones de iteración de valores. La EE fue ilustrada en un ejemplo de crecimiento económico (a continuación se dará un bosquejo del método usado para su solución).

El procedimiento de solución que se usó, de forma general, es el siguiente:

(a) comprobar la diferenciabilidad de las componentes del modelo de control;

(b) determinar la EE a partir de la Ecuación de Programación Dinámica;

(c) la EE se resuelve de manera iterada, iniciando con $W_0' = 0$;

(d) utilizando la información del paso anterior, se obtiene la función de valores óptimos W, tomando el límite de las funciones de iteración de valores;

(e) por último, para determinar la política óptima se usa la Ecuación de Programación Dinámica.

Después del estudio de la diferenciabilidad en modelos deterministas, se pasa al análisis de problemas de control estocástico inducidos por problemas deterministas. Se presentan condiciones estructurales en el modelo de control, las cuales garantizan la conexión entre ambos problemas. Una de las condiciones presentada es la suposición de que el problema determinista inicial tiene una solución, y ésta es usada para generar la solución del problema estocástico.

También, se presentaron algunos ejemplos que muestran el uso del procedimiento descrito en el párrafo anterior, que en general se puede resumir en los puntos siguientes:

(a) dado un problema de control determinista, su solución puede ser obtenida, ya sea por programación dinámica o por algún otro método iterativo o variacional como, por ejemplo, la Ecuación de Euler;

(b) el problema determinista es perturbado, de tal forma que se cumplan las Suposiciones 4;

(c) hacemos uso del Teorema 4.2.6 y determinamos la solución del problema estocástico inducido.

Diferenciabilidad

0.8. Fórmula de la Envolvente

Este capítulo está basado en el artículo [21].

Sean p y m enteros positivos fijos. Sea $X \subseteq \mathbb{R}^p$, $A \subseteq \mathbb{R}^m$, con A convexo, y $A(x) \subset A, x \in X$ son conjuntos medibles y no vacíos (con respecto a la $\sigma-$álgebra de Borel de A).

Sea $G : \mathbb{K} \to \mathbb{R}$ una función medible donde $\mathbb{K} := \{(x,a) : x \in X, a \in A(x)\}$. Supóngase que existe una función medible $\widehat{H} : X \to \mathbb{R}$ tal que

$$G(x,a) \geq \widehat{H}(x),$$

para todo $x \in X$ y $a \in A(x)$. Defínase $g : X \to \mathbb{R}$ por

$$g(x) = \inf_{a \in A(x)} G(x,a), \tag{34}$$

$x \in X$.

Hipótesis 5

a) $G \in C^2(int(\mathbb{K}); \mathbb{R})$, además, $G_{aa}(x,\cdot)$ es positiva definida, para cada $x \in X$.

b) Existe una función $h : X \to A$ tal que $h(x) \in int\ (A(x))$ y $g(x) = G(x,h(x))$, para cada $x \in X$.

Remark 9 *a) Como es usual, para cada $x \in X$, la condición de primer orden para la optimalidad de $G(x,\cdot)$ está definida por medio de los puntos $\widehat{a} \in int(A(x))$ los cuales satisfacen $G_a(x,\widehat{a}) = 0$.*

b) *Evidentemente, en la Hipótesis 5 (b), si $A(x)$ es abierto para cada $x \in X$, entonces el minimizador $h(x)$ pertenece al interior de $A(x)$, $x \in X$. Además, obsérvese que la Hipótesis 5 (b) permite remplazar "inf" por "min" in (34).*

c) *Obsérvese que la Hipótesis 5 (a) implica que $G(x,\cdot)$ es una función estrictamente convexa, para cada $x \in X$ (véase [22] p. 260).*

Supóngase que las Hipótesis 5 se satisfacen. Entonces $h \in C^1(int(X); A)$, $g \in C^2(int(X); \mathbb{R})$ y la siguiente fórmula es válida:

$$g'(x) = G_x(x, h(x)), \tag{35}$$

$x \in int(X)$.

Proof. Sea $x \in int(X)$ fijo. Como el minimizador $h(x)$ pertenece al interior de $A(x)$, entonces usando la condición de primer orden para la optimalidad de G, resulta que $G_a(x, h(x)) = 0$.

La Hipótesis 5 a) implica que $G_{aa}(x,\cdot)$ tiene una inversa G_{aa}^{-1}, y que h es único (esta es una consecuencia de la convexidad estricta de $G(x,\cdot)$). Entonces usando el Teorema de la Función Implícita (véase Apéndice B), resulta que $h \in C^1(int(X); A)$ y

$$h'(x) = -G_{ax}(x, h(x))G_{aa}^{-1}(x, h(x)).$$

Por otro lado, de la Hipótesis 5 b), se sigue que

$$g'(x) = G_x(x, h(x)) + G_a(x, h(x))h'(x).$$

Como $G_a(x, h(x)) = 0$, entonces $g'(x) = G_x(x, h(x))$, i.e. (35) se cumple. Consecuentemente,

$$g''(x) = G_{xx}(x, h(x)) + G_{xa}(x, h(x))h'(x).$$

Por la Hipótesis 5 a) y ya que $h \in C^1(int(X); A)$, y x es arbitrario, se sigue que $g \in C^2(int(X); \mathbb{R})$. ■

0.8.1. Ejemplos

Example 4 *Sea $\delta \in (0,1)$ un número real fijo. Tómese $X = A = [0,1)$, y $A(0) = \{0\}$, $A(x) = (0, x^\delta]$, $x \in X$, $x \neq 0$. La dinámica del sistema está* *dada por la ecuación en diferencias:* $x_{t+1} = x_t^\delta - a_t$, $t = 0, 1, ...,$ *y la función de recompensa está dada por* $r(x, a) = \ln a$, *si* $x \in (0, 1)$, $a \in (0, x^\delta]$, *y* $r(0, 0) = -\infty$.

Nótese que $Q(\{x^\delta - a\}\,|x, a) = 1$, *si* $x \in (0, 1), a \in (0, x^\delta]$, *y* $Q(\{0\}\,|0, 0) = 1$; *además, para cada* $n = 1, 2, ...,$ $V_n(0) = -\infty$, $f_n(0) = 0$, $V(0) = -\infty$, *y* $f(0) = 0$.

El resto de la solución al ejemplo se da a continuación.

Lemma 14 *Para cada* $n = 1, 2, ...,$

$$V_n'(x) = \frac{\delta}{x}\sum_{k=0}^{n-1}(\alpha\delta)^k, \tag{36}$$

$0 < x < 1$.

Proof. La prueba será hecha por inducción. Para $n = 1$, directamente se obtiene que

$$V_1(x) = \max_{a\in(0,x^\delta]} \ln a = \delta \ln x,$$

$0 < x < 1$.

Así, $V_1'(x) = \delta/x, 0 < x < 1$.

Ahora supóngase que para algún $n > 1$, V_n' satisface (36). Obsérvese que, para cada $x \in (0, 1)$,

$$V_{n+1}(x) = \max_{a\in(0,x^\delta]} \left[\ln a + \alpha V_n(x^\delta - a)\right] = \max_{a\in(0,x^\delta)} \left[\ln a + \alpha V_n(x^\delta - a)\right],$$

$x \in (0, 1)$, porque $V_n(x^\delta - x^\delta) = V_n(0) = -\infty$.

Sea $\widehat{\mathbb{K}} := \{(x, a) : x \in (0, 1), a \in (0, x^\delta)\}$ y

$$G^{n+1}(x, a) = \ln a + \alpha V_n(x^\delta - a),$$

$(x, a) \in \widehat{\mathbb{K}}$. Nótese que $G^{n+1} \in C^2(\widehat{\mathbb{K}}; \mathbb{R})$, y algunos cálculos adicionales permiten obtener que

$$G_{aa}^{n+1}(x, a) = -(a^{-2} + \alpha\delta\sum_{k=0}^{n-1}(\alpha\delta)^k(x^\delta - a)^{-2}) < 0,$$

$(x, a) \in \widehat{\mathbb{K}}$.

Ahora, usando el Teorema 5.1.2, se sigue que

$$V'_{n+1}(x) = \alpha V'_n(x^\delta - f_{n+1}(x))\delta x^{\delta-1}, \tag{37}$$

$x \in (0,1)$. Por otro lado, la condición de primer orden para la optimalidad de G^{n+1} permite obtener:

$$\frac{1}{f_{n+1}(x)} = \alpha V'_n(x^\delta - f_{n+1}(x)), \tag{38}$$

$x \in (0,1)$. Sustituyendo (36) en (38), se obtiene que

$$f_{n+1}(x) = x^\delta / \sum_{k=0}^{n} (\alpha\delta)^k, \tag{39}$$

$x \in (0,1)$. Finalmente, sustituyendo (36) y (39) en (37), resulta que

$$V'_{n+1}(x) = (\delta/x) \sum_{k=0}^{n} (\alpha\delta)^k,$$

$x \in (0,1)$. Esto termina la prueba del Lema 5.1.4. ■

Corollary 15 *Para el Ejemplo 5.1.3, se tiene que*

$$V(x) = [\delta/(1-\alpha\delta)] \ln x + M, \text{ y}$$

$$f(x) = (1-\alpha\delta)x^\delta,$$

$x \in (0,1)$, *donde*

$$M = \frac{1}{\delta - 1}\left[\ln(1-\alpha\delta) + \frac{\alpha\delta}{1-\alpha\delta}\ln(\alpha\delta)\right].$$

Proof. Usando el Lema anterior y su prueba, se sigue que

$$V_{n-1}(x) = \delta K_{n-1} \ln x + M_{n-1}, \; V_n(x) = \delta K_n \ln x + M_n, \text{ y} \tag{40}$$

$$f_n(x) = x^\delta / K_n, \tag{41}$$

$x \in (0,1)$, donde $M_{n-1}, M_n \in \mathbb{R}$, $K_{n-1} = \sum_{k=0}^{n-2} (\alpha\delta)^k$ y $K_n = \sum_{k=0}^{n-1} (\alpha\delta)^k$, $n = 2, \ldots$.

Sustituyendo (40) y (41) en (7), se llega a que

$$M_n - \alpha M_{n-1} = -\ln K_n + \alpha\delta K_{n-1}\ln(\frac{K_n - 1}{K_n}), \tag{42}$$

$n = 1, 2....$ Como $V_n(x) \to V(x)$, para cada $x \in X$ y $K_n \to 1/(1-\alpha\delta)$, $n \to \infty$, (nótese que $0 < \alpha\delta < 1$), entonces de (40) se sigue que $\{M_n\}$ es convergente. Sea $M := \lim_{n\to\infty} M_n$. Entonces dejando n tender a infinito en (42), resulta que

$$M = \frac{1}{1-\alpha}\left[\ln(1-\alpha\delta) + \frac{\alpha\delta}{1-\alpha\delta}\ln(\alpha\delta)\right].$$

Por lo tanto, $V(x) = [\delta/(1-\alpha\delta)]\ln x + M$, $x \in (0,1)$, y entonces de (8) se sigue que $f(x) = (1-\alpha\delta)x^{\delta}$, $x \in (0,1)$. ■

Ahora será presentado un problema de consumo-inversión. En [25] pp. 155-159 (véase también [23] y [36]) se garantiza el procedimiento de programación dinámica es válido para este ejemplo.

Sea $X = A = [0, \infty)$ y $A(x) = [0, x], x \in X$. La recompensa y la dinámica del sistema están dadas por

$$\begin{aligned} r(x,a) &= \frac{b}{\gamma}(x-a)^{\gamma}, \\ x_{t+1} &= a_t\xi_t, \end{aligned}$$

$t = 0, 1, ...,$ $x \in X$ y $a \in A(x)$, donde $b > 0$ y $0 < \gamma < 1$. Se supone que $S = [0, \infty)$, $\{\xi_t\}$ es una sucesión de variables aleatorias i.i.d., y si ξ es un elemento genérico de $\{\xi_t\}$, $m_\gamma := E[\xi^\gamma] < \infty$, con $0 < \alpha m_\gamma < 1$; también supóngase que Δ denota la función de densidad de ξ y $\widehat{\delta} := (\alpha m_\gamma)^{1/(\gamma-1)}$.

Nótese que, para $n = 1, 2, ...,$ $V_n(0) = 0$, $f_n(0) = 0$, $V(0) = 0$ y $f(0) = 0$.

Para $n = 1, 2, ...,$ se tiene que

$$V_n'(x) = \left(\frac{\widehat{\delta}^{n-1}(1-\widehat{\delta})}{1-\widehat{\delta}^{n}}\right)^{\gamma-1} bx^{\gamma-1}, \tag{43}$$

$x \in (0, \infty)$.

La prueba será hecha por inducción. Para $n = 1$,

$$V_1(x) = \max_{a\in[0,x]} \frac{b}{\gamma}(x-a)^{\gamma} = \frac{b}{\gamma}x^{\gamma},$$

$x \in (0,\infty)$. Entonces $V_1'(x) = bx^{\gamma-1}$, $x \in (0,\infty)$. Supóngase que para algún $n > 1$, $V_n'(x) = K_n b x^{\gamma-1}$, donde $K_n = \left(\widehat{\delta}^{n-1}(1-\widehat{\delta})/(1-\widehat{\delta}^n)\right)^{\gamma-1}$. Sea

$$\begin{aligned} G^{n+1}(x,a) &= \frac{b}{\gamma}(x-a)^\gamma + \alpha \int V_n(as)\Delta(s)ds \\ &= \frac{b}{\gamma}\left[(x-a)^\gamma + \alpha m_\gamma K_n a^\gamma\right], \end{aligned} \tag{44}$$

$(x,a) \in \mathbb{K}$, donde V_n fue obtenida integrando (43). Obsérvese que $f_{n+1}(x) \in (0,x), x \in (0,\infty)$. De hecho, para cada $x \in (0,\infty)$ considérese la función $G_a^{n+1}(x,\cdot)$. Como $G_a^{n+1}(x,0) = -bx^{\gamma-1} < 0$ y

$$G_a^{n+1}(x,x) = \alpha b K_n x^{\gamma-1} m_\gamma > 0,$$

entonces, por el Teorema del Valor Medio (véase Apéndice B) se tiene que $f_{n+1}(x) \in (0,x)$. Así, para cada $x \in (0,\infty)$,

$$V_{n+1}(x) = \max_{a \in [0,x]} G^{n+1}(x,a) = \max_{a \in (0,x)} G^{n+1}(x,a).$$

Sea $\overline{\mathbb{K}} = \{(x,a) : x \in (0,\infty), a \in (0,x)\}$. Entonces de (44), se sigue que $G^{n+1} \in C^2(\overline{\mathbb{K}};\mathbb{R})$; y que

$$G_{aa}^{n+1}(x,a) = b(\gamma-1)\left[(x-a)^{\gamma-2} + K_n \alpha m_\gamma a^{\gamma-2}\right] < 0,$$

$(x,a) \in \overline{\mathbb{K}}$.

Por otro lado, del Teorema 5.1.2, se tiene que $V_{n+1} \in C^2(\overline{\mathbb{K}};\mathbb{R})$, y

$$V_{n+1}'(x) = b(x - f_{n+1}(x))^{\gamma-1}, \tag{45}$$

$x \in (0,\infty)$. Usando la condición de primer orden para la optimalidad de G^{n+1}, nos lleva a la siguiente relación

$$(x - f_{n+1}(x))^{\gamma-1} = \alpha m_\gamma K_n f_{n+1}(x)^{\gamma-1},$$

$x \in (0,\infty)$, o equivalentemente,

$$f_{n+1}(x) = \frac{x}{1 + (\alpha m_\gamma K_n)^{\frac{1}{\gamma-1}}}, \tag{46}$$

$x \in (0, \infty)$. Finalmente, sustituyendo (46) en (45), resulta que

$$V'_{n+1}(x) = \left(\frac{\widehat{\delta}^n (1 - \widehat{\delta})}{1 - \widehat{\delta}^{n+1}} \right)^{\gamma - 1} b x^{\gamma - 1},$$

$x \in (0, \infty)$. Esto completa la prueba del Lema 5.1.7.

Para el Ejemplo 5.1.6, se tiene que

$$V(x) = \left(\frac{\widehat{\delta} - 1}{\widehat{\delta}} \right)^{\gamma - 1} \frac{b}{\gamma} x^{\gamma}, \text{ y} \tag{47}$$

$$f(x) = \frac{x}{\widehat{\delta}}, \tag{48}$$

$x \in (0, \infty)$.

Integrando (43), resulta que

$$V_n(x) = \left(\frac{\widehat{\delta}^{n-1} (1 - \widehat{\delta})}{1 - \widehat{\delta}^{n}} \right)^{\gamma - 1} \frac{b}{\gamma} x^{\gamma},$$

$x \in (0, \infty)$ (tomando la constante de integración igual a cero ya que $V_n(0) = 0$, para $n = 1, 2, ...$). Entonces haciendo n tender a infinito en la última igualdad conduce a (47). Sustituyendo (47) en (8), se sigue que $f(x) = x/\widehat{\delta}$, $x \in (0, \infty)$.

Obsérvese que la solución obtenida en (47) y (48) satisface:

$$\begin{aligned} \lim_{n\to\infty} \alpha^n E_x^f \left[V(x_n) \right] &= \lim_{n\to\infty} \alpha^n \left(\frac{\widehat{\delta} - 1}{\widehat{\delta}} \right)^{\gamma - 1} \frac{b}{\gamma} \frac{x^{\gamma}}{\widehat{\delta}^{\gamma n}} m_{\gamma}^n \\ &= x^{\gamma} \frac{b(\widehat{\delta} - 1)^{\gamma - 1}}{\gamma \widehat{\delta}^{\gamma - 1}} \lim_{n\to\infty} (\alpha m_{\gamma})^{n/(1-\gamma)} = 0. \end{aligned}$$

Ahora será considerado un sistema lineal con costo cuadrático. Este ejemplo es una variante cercana del problema lineal cuadrático estándar (LQ) ([9]). Obviamente, el problema LQ estándar puede ser resuelto usando la misma técnica.

Example 5 *Sea $X = A = A(x) = S = \mathbb{R}, x \in X$. La función de costo y la dinámica del sistema están dadas por*

$$\begin{aligned} c(x,a) &= x^2 + a^2 + ax, \\ x_{t+1} &= x_t + a_t + \xi_t, \end{aligned}$$

$t = 0, 1, 2...., (x, a) \in \mathbb{K}$.

Las variables aleatorias ξ_t, $t = 0, 1, 2, ...,$ se suponen que son i.i.d. con densidad continua Δ, con media igual a cero, y varianza finita. Sea $\sigma^2 := E(\xi^2)$, donde ξ es un elemento genérico de la sucesión $\{\xi_t\}$.

Remark 10 *En este ejemplo, el costo c es continuo. Además, c es inf-compacto porque para $x \in X$ y $l \in \mathbb{R}$: $\{a \in \mathbb{R} : c(x,a) \leq l\} = \emptyset$ si $l < \frac{3}{4}x^2$, y $\{a \in \mathbb{R} : c(x,a) \leq l\} = \left[-\sqrt{l - \frac{3}{4}x^2} - \frac{x}{2}, \sqrt{l - \frac{3}{4}x^2} - \frac{x}{2}\right]$ si $l \geq \frac{3}{4}x^2$; en ambos casos el conjunto $\{a \in \mathbb{R} : c(x,a) \leq l\}$ es compacto. La ley de transición está dada por $Q(B \mid x, a) = \int I_B(x + a + s)\Delta(s)\, ds$, $x \in X$, $a \in \mathbb{R}$, $B \in \mathcal{B}(X)$, entonces usando el Teorema de Cambio de Variable (véase Apéndice B), puede obtenerse que*

$$Q(B \mid x, a) = \int_B \Delta(u - x - a)\, du,$$

i.e., $\Delta(\cdot - x - a)$ es una densidad para $Q(\cdot \mid x, a)$ con respecto a la medida de Lebesgue en $\mathbb{R}$. Como $\Delta(\cdot)$ es continua, y tomando en cuanta el Lema 2.3 en [16], resulta que Q es fuertemente continua. Tomando $g(x) = -x$, $x \in X$, es simple verificar que $v(g, x) < \infty$, para todo $x \in X$. Por lo tanto, las Hipótesis 1 se cumplen. Así (6) y (8) son válidas.

Lemma 16 *Para cada $n = 1, 2, ...,$*

$$V_n'(x) = 2K_n x, \tag{49}$$

$x \in X$, donde

$$K_n = \frac{3 + 4\alpha K_{n-1}}{4 + 4\alpha K_{n-1}},$$

con $K_0 = 0$.

Proof. Para $n = 1$, es fácil probar que

$$V_1(x) = \min_{a \in \mathbb{R}} \left[x^2 + a^2 + ax\right] = \frac{3}{4}x^2,$$

$x \in X$. Entonces $V_1'(x) = 2(3/4)x$, $x \in X$.

Para $n > 1$, supóngase (49). Integrando (49), resulta que $V_n(x) = K_n x^2 + L_n$, donde L_n es un número real. Entonces,

$$G^{n+1}(x, a) = x^2 + a^2 + ax + \alpha K_n[(x + a)^2 + \sigma^2] + \alpha L_n,$$

$(x, a) \in \mathbb{K}$. Así, $G^{n+1} \in C^2(\mathbb{K}; \mathbb{R})$ y un cálculo directo permite obtener que $G^{n+1}_{aa}(x, a) = 2(1 + \alpha K_n) > 0$, $(x, a) \in \mathbb{K}$.

Usando el Teorema 5.1.2, se tiene que $V_{n+1} \in C^2(\mathbb{K}; \mathbb{R})$, y

$$V_{n+1}'(x) = 2x + f_{n+1}(x) + 2\alpha K_n(x + f_{n+1}(x)), \tag{50}$$

$x \in X$. La condición de primer orden para la optimalidad de G^{n+1} lleva a que

$$2f_{n+1}(x) + x + 2\alpha K_n(x + f_{n+1}(x)) = 0, \tag{51}$$

$x \in X$. Entonces de (51), resulta que

$$f_{n+1}(x) = -x \left(\frac{1 + 2\alpha K_n}{2 + 2\alpha K_n} \right), \tag{52}$$

$x \in X$, y sustituyendo (52) en (50), se sigue que

$$V_{n+1}'(x) = 2K_{n+1}x,$$

$x \in X$, donde

$$K_{n+1} = (3 + 4\alpha K_n)/(4 + 4\alpha K_n).$$

■

Lemma 17 *a) La sucesión de números reales $\{K_n\}$ obtenida en el Lema anterior, es convergente a K, donde K debe satisfacer la siguiente versión de la ecuación de Ricatti,*

$$4\alpha K^2 + (4 - 4\alpha)K - 3 = 0.$$

b) La función de valores óptimos y la política óptima son

$$V(x) = Kx^2 + L, \tag{53}$$

$$f(x) = -\left(\frac{1 + 2\alpha K}{2 + 2\alpha K} \right) x, \tag{54}$$

respectivamente, donde $x \in X$, y $L = [\alpha K/(1 - \alpha)]\, \sigma^2$.

Proof.

a) Para determinar el límite de $\{K_n\}$, el argumento usado por Dynkin y Yushkevich ([25] pp. 77-78) será usado. El punto principal es que $g(K_{n-1}) = K_n$, donde g es la transformación lineal fraccional,

$$g(z) = \frac{3 + 4\alpha z}{4 + 4\alpha z}, \tag{55}$$

$z \in \mathbb{C}$, donde $\mathbb{C}$ es el conjunto de números complejos. Los puntos fijos de g, i.e. los número complejos z tales que $z = g(z)$, son las raíces de la ecuación

$$4\alpha z^2 + (4 - 4\alpha)z - 3 = 0. \tag{56}$$

Algunos cálculos adicionales permiten probar que la ecuación (56) tiene dos raíces reales: una raíz positiva denotada por z_1 y una raíz negativa denotada por z_2. Entonces usando (55) y (56), resulta que

$$\begin{aligned} g(z) - z_j &= \frac{3 + 4z\alpha - 4z_j - 4z\alpha z_j}{4(1 + z\alpha)}, \\ &= \frac{\alpha(1 - z_j)(z - z_j)}{1 + \alpha z} \end{aligned}$$

$j = 1, 2$, $z \in \mathbb{C}$. Por lo tanto,

$$\frac{g(z) - z_1}{g(z) - z_2} = \lambda \frac{z - z_1}{z - z_2},$$

donde

$$\lambda = \frac{1 - z_1}{1 - z_2}.$$

Como $g(K_{n-1}) = K_n$, se obtiene que

$$\frac{K_n - z_1}{K_n - z_2} = \lambda \frac{K_{n-1} - z_1}{K_{n-1} - z_2}, \tag{57}$$

$n = 1, 2, ...$ Iterando (57), es posible llegar a la relación siguiente:

$$\frac{K_n - z_1}{K_n - z_2} = \lambda^n \frac{z_1}{z_2},$$

$n = 1, 2, \ldots.$ Entonces

$$K_n = \frac{z_1 z_2 (1 - \lambda^n)}{z_2 - \lambda^n z_1},$$

$n = 1, 2, \ldots$ Como $z_1 > z_2$, entonces $|\lambda| < 1$. Cuando $n \to \infty$ en la última igualdad, se obtiene que

$$K := \lim_{n \to \infty} K_n = z_1.$$

b) Usando la información del inciso anterior, resulta que

$$V_{n-1}(x) = K_{n-1} x^2 + L_{n-1},\ V_n(x) = K_n x^2 + L_n, \text{ y} \tag{58}$$

$$f_n(x) = -x \left(\frac{1 + 2\alpha K_{n-1}}{2 + 2\alpha K_{n-1}} \right), \tag{59}$$

donde $x \in X$, L_{n-1} y $L_n \in \mathbb{R}$, $n = 1, 2, \ldots.$ Sustituyendo (58) y (59) en (7), se sigue que

$$L_n - \alpha L_{n-1} = \alpha K_{n-1} \sigma^2, \tag{60}$$

$n = 1, 2, \ldots.$ Como la sucesión $\{K_n\}$ es convergente, la sucesión $\{L_n\}$ también es convergente. Sea $L := \lim_{n \to \infty} L_n$. Entonces, cuando $n \to \infty$ en (60), se obtiene que

$$L = \frac{\alpha K}{1 - \alpha} \sigma^2.$$

Por lo tanto, (53) y (54) se siguen. Esto termina la prueba del Lema 5.1.13.

■

0.9. Diferenciabilidad de PCMs

Sea $(X,\ A,\ \{A(x) : x \in X\}, Q, c)$ un modelo de control de Markov fijo. Supóngase que la ley de transición está dada por la ecuación en diferencias:

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

donde $t = 0, 1 \ldots$, $F : \mathbb{K} \times S \to X$ es una función medible; $\{x_t\}$ y $\{a_t\}$ son las sucesiones de estados y controles, respectivamente, y $\{\xi_t\}$ es una sucesión de variables aleatorias i.i.d. tomando valores en un espacio de Borel $S \subseteq \mathbb{R}^p$ (recuérdese que X es también un subconjunto no vacío de $\mathbb{R}^p$). Se supone también que S es un conjunto abierto. Sea ξ un elemento genérico de la sucesión $\{\xi_t\}$. La densidad de ξ es designada por Δ.

Sea $G^\infty : \mathbb{K} \to \mathbb{R}$ la función definida por

$$G^\infty(x,a) := c(x,a) + \alpha \int V(y) Q\left(dy|\, x,a\right),$$

$(x,a) \in \mathbb{K}$.

Hipótesis 6

a) $c \in C^2(int(\mathbb{K}); \mathbb{R})$.

b) $F(\cdot,\cdot,s) \in C^2(int(\mathbb{K}); \mathbb{X})$, para cada $s \in S$. Además, F tiene una inversa R en la tercera variable, $R : \mathbb{K} \times X \to S$ tal que $R \in C^2(int(\mathbb{K} \times X); S)$.

c) $\Delta \in C^2(S; \mathbb{R})$.

d) El intercambio entre derivadas e integrales es válido (véase Observación 5.2.2, más adelante)

Lemma 18 *Bajo las Hipótesis 6, $G^\infty \in C^2(int(\mathbb{K}); \mathbb{R})$.*

Proof. Nótese que la Hipótesis 6 permite expresar el kérnel estocástico en la forma siguiente: para B subconjunto medible de X y $(x,a) \in \mathbb{K}$,

$$\begin{aligned} Q\left(B|\, x,a\right) &= \mu\left\{s \in S|\, F(x,a,s) \in B\right\} = \mu\left\{s \in S|\, s \in R(x,a,B)\right\}, \\ &= \int_{R(x,a,B)} d\mu(u) = \int_{R(x,a,B)} \Delta(u) du. \end{aligned}$$

Entonces por el Teorema de Cambio de Variable (véase Apéndice B), resulta que:

$$Q\left(B|\, x,a\right) = \int_B \Delta(R(x,a,u)) \left|R_u(x,a,u)\right| du, \tag{61}$$

donde $|R_u|$ denota el determinante Jacobiano de R. De (61), G^∞ puede ser expresado como

$$G^\infty(x,a) = c(x,a) + \alpha \int V(u) \Delta(R(x,a,u)) \left|R_u(x,a,u)\right| du,$$

$(x, a) \in \mathbb{K}$. Ahora, usando la diferenciabilidad de la función de costo c, la densidad Δ y el determinante Jacobiano de R, y el intercambio entre derivada e integral, el resultado se sigue. ■

Remark 11 *En el Lema 5.2.1, la Hipótesis 6 d) fue usada para garantizar la diferenciabilidad de segundo orden de la integral $\int \widehat{U}(x, a, u)du$, con respecto a x o a, donde*

$$\widehat{U}(x, a, u) := V(u)\Delta(R(x, a, u)) \left|R_u(x, a, u)\right|,$$

$(x, a, u) \in int(\mathbb{K} \times X)$. Esta condición puede ser verificada en la práctica cuando las derivadas $\widehat{U}$ pueden acotarse, en el siguiente sentido: para $(x, a, u) \in int(\mathbb{K} \times X)$, $\left|\widehat{U}_x(x, a, u)\right| \leq \widehat{g}(u, a)$, $\left|\widehat{U}_a(x, a, u)\right| \leq \widehat{h}(u, x)$, $\left|\widehat{U}_{xx}(x, a, u)\right| \leq \widehat{l}(u, a)$, $\left|\widehat{U}_{aa}(x, a, u)\right| \leq \widehat{m}(u, a)$ y $\left|\widehat{U}_{xa}(x, a, u)\right| \leq \widehat{n}(u, a)$, para funciones integrables $\widehat{g}$, $\widehat{h}$, $\widehat{l}$, $\widehat{m}$ y $\widehat{n}$ (véase Teorema 20.4, p. 147 en [1]).

Hipótesis 7

a) $G^{\infty}(x, \cdot)$ es positiva definida, para cada $x \in X$.

b) $f(x) \in int(A(x))$, para cada $x \in X$.

Theorem 19 *Supóngase que las Hipótesis 6 y 7 se cumplen. Entonces $f \in C^1(int(X); A)$ y $V \in C^2(int(X); \mathbb{R})$.*

Proof. Usando el hecho que $G^{\infty} \in C^2(int(\mathbb{K}); \mathbb{R})$ (véase Lema 5.2.1), Hipótesis 7, y el Teorema 5.1.2, se sigue el Teorema 5.2.3. ■

Example 6 *Sea $X = A = S = \mathbb{R}$, y $A(x) = [x-1, x+1]$, $x \in X$. La función de costo y la dinámica del sistema están dadas por*

$$c(x, a) = (x - a)^2 + T,$$

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

$t = 0, 1, 2, ...,$ $(x, a) \in \mathbb{K}$ y T es un número real fijo. Las variables aleatorias ξ_t, $t = 1, 2, ...$ se suponen que son variables aleatorias i.i.d. con densidad normal,

$$\Delta(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2},$$

$y \in \mathbb{R}$.

Lemma 20 *Existe una única política estacionaria f tal que $f(x) \in int(A(x))$, $x \in X$.*

Proof. La prueba para la existencia de una política estacionaria f es similar a la dada en el Ejemplo 5.1.10. Evidentemente este ejemplo satisface la condición C2 de [16]. Consecuentemente la política óptima f es única. Por otro lado, f es interior, porque: $v(g,x) = T/(1-\alpha)$, donde $g(x) = x, x \in X$ y en la frontera de $A(x)$, $v(g_1, x) = v(g_2, x) = (T+1)/(1-\alpha) > v(g,x)$, donde $g_1(x) = x - 1$ y $g_2(x) = x + 1$, $x \in X$. ■

Remark 12 *Obsérvese que para este ejemplo, en particular, $0 \leq V(x) \leq v(g,x) = T/(1-\alpha)$, $x \in X$, donde $g(x) = x, x \in X$.*

Lemma 21 *$G^{\infty} \in C^2(int(\mathbb{K}); \mathbb{R})$ y $G^{\infty}_{aa}(x, \cdot) > 0$, para cada $x \in X$.*

Proof. Sea $x \in X$, fijo y

$$H(a) = \int V(a+s)\Delta(s)ds,$$

$a \in A(x)$. Entonces por el Lema 6.2 en [16] es posible obtener que H es convexo en $A(x)$. Por otro lado, usando el Teorema de Cambio de Variable (véase Apéndice B), resulta que $H(a) = \int V(u)\Delta(u-a)du$, $a \in A(x)$. Para garantizar el intercambio entre derivada e integral, considere las relaciones siguientes:

$$\begin{aligned} \int V(u)\Delta(u-a)du &= \int \frac{V(u)}{\sqrt{2\pi}} e^{-\frac{1}{2}(u-a)^2} du \\ &= \int M(u,a)du, \end{aligned}$$

donde

$$M(u,a) := \frac{V(u)}{\sqrt{2\pi}} e^{-\frac{1}{2}(u-a)^2},$$

$u \in \mathbb{R}$ y $a \in A(x)$. Así, $M_a(u,a) = \dfrac{V(u)}{\sqrt{2\pi}} e^{-\frac{1}{2}(u-a)^2}(u-a)$, $u \in \mathbb{R}$ y $a \in A(x)$.

Como $x-1 \leq a \leq x+1$ y usando la observación 5.2.6 los casos siguientes son posibles para determinar cotas para M_a.
Case 1. Supóngase que $u-a \geq 0$. Entonces resulta que,

$$\begin{aligned} |M_a(u,a)| &\leq \frac{T}{\sqrt{2\pi}(1-\alpha)} e^{-\frac{1}{2}(u-a)^2}(u-x+1), \\ &\leq \theta_1(u,x), \end{aligned}$$

donde

$$\theta_1(u,x) := \frac{T(u-x+1)}{\sqrt{2\pi}(1-\alpha)} \left[e^{\frac{-1}{2}u^2+u(x+1)} I_{[0,\infty)}(u) + e^{\frac{-1}{2}u^2+u(x-1)} I_{(-\infty,0]}(u) \right],$$

y $I_{[\cdot]}$ denota la función indicadora del conjunto $[\cdot]$.

Case 2. Ahora, supóngase que $u-a<0$. Es posible obtener que

$$\begin{aligned} |M_a(u,a)| &\leq \frac{T}{\sqrt{2\pi}\,(1-\alpha)} e^{-\dfrac{1}{2}(u-a)^2}(x+1-u), \\ &\leq \theta_2(u,x), \end{aligned}$$

donde

$$\theta_2(u,x) := \frac{T(x+1-u)}{\sqrt{2\pi}(1-\alpha)} \left[e^{\frac{-1}{2}u^2+u(x+1)} I_{[0,\infty)}(u) + e^{\frac{-1}{2}u^2+u(x-1)} I_{(-\infty,0]}(u) \right],$$

Como x fue arbitrario, se sigue que $\theta = \theta_1 + \theta_2$, es una cota para $|M_a|$. Además, es fácil verificar que esta cota es integrable. La cota para M_{aa} puede ser obtenida en forma similar. Entonces $H \in C^2(A;\mathbb{R})$. Como $c \in C^2(\mathbb{K};\mathbb{R})$, resulta que $G^\infty \in C^2(int(\mathbb{K});\mathbb{R})$. Además, $G^\infty_{aa} > 0$, ya que $c_{aa} > 0$ y $H_{aa} \geq 0$ (esto es una consecuencia de la convexidad y la diferenciabilidad de segundo orden de H). ■

Lemma 22 *$V \in C^2(X;\mathbb{R})$ y $f \in C^1(X;A)$.*

Proof. La prueba es una consecuencia del Teorema 5.2.3. ■

0.10. La Diferenciabilidad en Modelos no Convexos

0.10.1. La Fórmula de la Envolvente

Referentes a los problemas de optimización trabajados en Sección 5.1 y en la sección anterior, considérese las hipótesis siguientes:

Hipótesis 8

a) Existe una única función $h : X \to A$ tal que $h(x) \in int(A(x))$ y $g(x) = G(x, h(x))$, para cada $x \in X$.

b) $G \in C^2(int(\mathbb{K}); \mathbb{R})$ y el determinante de $G_{aa}(x, h(x))$ es diferente de cero, para cada $x \in X$.

Remark 13 *Obsérvese que la Hipótesis 8 es más débil que la Hipótesis 7 a).*

Hipótesis 9

a) f es única, y $f(x) \in int(A(x))$, para cada $x \in X$.

b) $G^\infty \in C^2(int(\mathbb{K}); \mathbb{R})$ y el determinante de $G^\infty_{aa}(x, f(x))$ es diferente de cero, para cada $x \in X$.

Remark 14 *La unicidad de f requerida en la Hipótesis 9 a) ha sido analizada en [16] para PCMs descontados.*

Ahora los Teoremas 5.3.3 y 5.3.4 serán establecidos. Estos son versiones no cóncavas de los Teoremas 5.1.2 y 5.2.3. Las demostraciones de los Teoremas 5.3.3 y 5.3.4 son similares a las pruebas de los Teoremas 5.1.2 y 5.2.3, respectivamente.

Theorem 23 *Supóngase que las Hipótesis 8 se cumplen. Entonces $h \in C^1(int(X); A)$, $g \in C^2(int(X); \mathbb{R})$ y la siguiente fórmula:*

$$g'(x) = G_x(x, h(x)),$$

$x \in int(X)$, es válida.

Theorem 24 *Supóngase que las Hipótesis 8 y 9 se cumplen. Entonces $f \in C^1(int(X); A)$ y $V \in C^2(int(X); \mathbb{R})$.*

0.10.2. Ejemplos

El ejemplo siguiente ilustra el uso del Teorema 5.3.3. Nótese que la función de costo del ejemplo es no cóncava.

Example 7 *Sea $X = S = \mathbb{R}$, $A = A(x) = [-\pi, \pi]$, para cada $x \in X$. Supóngase que la función de costo y la dinámica están dadas por*

$$\begin{aligned} c(x,a) &= x^2 + \sin a + 1, \\ x_{t+1} &= x_t + \xi_t, \end{aligned}$$

para $t = 0, 1, ...,$ y $\{\xi_t\}$ es una sucesión de variables aleatorias i.i.d. con densidad continua Δ, con media igual a cero, y varianza finita. Sea $\widehat{\sigma}^2 := E(\xi^2)$, donde ξ es un elemento genérico de la sucesión $\{\xi_t\}$.

Remark 15 *La existencia de una política óptima es similar a la dada en ejemplos anteriores (ver Ejemplo 5.1.10).*

Lemma 25 *Para cada $n = 1, 2, ...,$*

$$V_n'(x) = 2\left(1 + \alpha + \alpha^2 + ... + \alpha^{n-1}\right) x,$$

$x \in X$.

Proof. Para $n = 1$,

$$V_1(x) = \min_{a \in [-\pi,\pi]} \left[x^2 + \sin(a) + 1\right] = x^2,$$

$x \in X$. Entonces $V_1'(x) = 2x$, $x \in X$. Supóngase para $n > 1$, que

$$V_n'(x) = 2\left(1 + \alpha + \alpha^2 + ... + \alpha^{n-1}\right) x, \tag{62}$$

$x \in X$. Integrando (62), se obtiene que $V_n(x) = x^2(1+\alpha+\alpha^2+...+\alpha^{n-1})+\widehat{L}_n$, donde $x \in X$ y $\widehat{L}_n$ es un número real constante. Entonces

$$\begin{aligned} G^{n+1}(x,a) &= x^2 + \sin a + 1 + \alpha \int V_n(x+s)\Delta(s)ds, \qquad (63) \\ &= x^2 + \sin(a) + 1 + \alpha(1 + \alpha + ... + \alpha^{n-1})(x^2 + \widehat{\sigma}^2) + \alpha L_n, \end{aligned}$$

$(x,a) \in \mathbb{K}$. Entonces $G^{n+1} \in C^2(X;\mathbb{R})$. También, el minimizador f_{n+1} de G^{n+1} es único debido al hecho que el mínimo de la función $\eta(a) = \sin a$,

$a \in (-\pi, \pi)$ se alcanza únicamente en $a = -\pi/2$ (véase (63)); i.e., $f_{n+1}(x) = -\pi/2 \in (-\pi, \pi)$, $x \in X$. Entonces, usando el Teorema 5.3.3, resulta que

$$V'_{n+1}(x) = (1 + \alpha + ... + \alpha^n)2x,$$

$x \in X$. ■

Corollary 26 *Para este ejemplo,* $V(x) = [1/(1-\alpha)]\, x + \alpha\widehat{\sigma}/(1-\alpha)^2$*,* $f(x) = -\frac{\pi}{2}$*,* $x \in X$*.*

Proof. Como para cada $n = 1, 2, ...,$ $V_n(x) = M_n x^2 + \widehat{L}_n$, $x \in X$, donde $M_n = 1 + \alpha + \alpha^2 + ... + \alpha^{n-1}$ y $\widehat{L}_n \in \mathbb{R}$ entonces sustituyendo V_n, V_{n-1} y $f_n = -\pi/2$ in (7), se obtiene que

$$\widehat{L}_n - \alpha \widehat{L}_{n-1} = \widehat{\sigma}^2(\alpha + ... + \alpha^{n-1}) \tag{64}$$

Dejando $n \to \infty$ en (64) se tiene que $\lim_{n\to\infty} \widehat{L}_n = \alpha\widehat{\sigma}^2/(1-\alpha)^2$. Por lo tanto,

$$V(x) = \frac{1}{1-\alpha}x + \frac{\alpha\widehat{\sigma}^2}{(1-\alpha)^2}, \tag{65}$$

$x \in X$. Finalmente, sustituyendo (65) en (8), resulta que $f(x) = -\pi/2, x \in X$.
■

Ahora se presenta una aplicación de Teorema 5.3.4.

Example 8 *Sea* $X = S = \mathbb{R}^2$*,* $A = A(\overline{x}) = [\pi/4, 3\pi/2] \times (-1, 1), \overline{x} \in X$ *and*

$$c(\overline{x}, \overline{a}) = \cos a + \frac{1}{20}a^2 + b^2 + 1, \tag{66}$$

$\overline{x} = (x, y) \in X, \overline{a} = (a, b) \in A(\overline{x})$*. La dinámica del sistema está dada por*

$$x_{t+1} = (\xi_t, \rho_t),$$

$t = 0, 1, ...$ *Los vectores aleatorios* (ξ_t, ρ_t)*,* $t = 0, 1, ...$ *se supone que son i.i.d. con densidad* $\Delta \in C^2(in(S); \mathbb{R})$*.*

Remark 16 *Obsérvese que c dada en (66) no es cóncava. De hecho, el determinante de* $c_{\overline{a}\overline{a}}(\overline{x}, \overline{a})$*, denotado por* $det(c_{\overline{a}\overline{a}}(\overline{x}, \overline{a}))$*, es igual a* $2(1/10 - \cos a_1)$*,* $\overline{x} \in X$*,* $\overline{a} \in A(\overline{x})$*. Entonces* $det(c_{\overline{a}\overline{a}}(\overline{x}, \overline{a})) > 0$*, si* $a > \arccos(1/10) = 1{,}47062...$*, y* $det(c_{\overline{a}\overline{a}}(\overline{x}, \overline{a})) < 0$*, si* $a < \arccos(1/10) = 1{,}47062....$

Lemma 27 *El Ejemplo 5.3.9 satisface (7) y (6).*

Proof. Evidentemente, la función de costo es continua y $c(\overline{x}, \overline{a}) \geq 0$, $\overline{x} \in X$, $\overline{a} \in A(\overline{x})$. Obsérvese que la función de costo es inf-compacta, como una consecuencia del hecho que $A(\overline{x})$ es compacto. Como (x_{t+1}, y_{t+1}) es independiente de estados y controles, se sigue que Q es fuertemente continua. Además, $v(\pi, \overline{x}) < \infty$ para todo $\pi \in \Pi$ y $\overline{x} \in X$, ya que la función de costo es acotada. ■

Lemma 28 *La política óptima f es única.*

Proof. En este caso, $G^{\infty}(\overline{x}, \overline{a}) = \cos a + \frac{1}{20}a^2 + b^2 + 1 + \alpha E[V(\xi, \rho)]$, $(\overline{x}, \overline{a}) \in \mathbb{K}$ y (ξ, ρ) es un elemento genérico de la sucesión $\{(\xi_t, \rho_t)\}$. Entonces la condición de primer orden para la optimalidad de $G^{\infty}(x, \cdot)$ está dada por

$$(-\sin a + \frac{1}{10}a, 2b) = (0, 0).$$

Así, $\sin a_1 - \frac{1}{10}a_1 = 0$ y $a_2 = 0$. Nótese que $\sin 3\pi/5 - 3\pi/50 > 0$ y $\sin \pi - \pi/10 < 0$, entonces por el Teorema del Valor Intermedio (véase Apéndice B) existe $a_1^* \in (3\pi/5, \pi)$, tal que $\sin a^* - \frac{1}{10}a^* = 0$. Para probar la unicidad de f, sea $w(a) = \sin a - \frac{1}{10}a$, $a \in [\pi/4, 3\pi/2]$ y supóngase que existe $d \in (\pi/4, 3\pi/2)$, con $w(d) = 0$ y $a^* < d$ (observe que el caso $a^* > d$ es similar). Entonces usando el Teorema de Rolle (véase Apéndice B) en cada uno de los intervalos $[0, a^*]$ y $[a^*, d]$, se obtiene que existen $i, j \in (0, 3\pi/2)$, tal que $w'(i) = w'(j) = 0$, pero esto no es posible ya que $w'(a) = \cos a - \frac{1}{10} = 0$, $a \in [\pi/4, 3\pi/2]$, si y sólo si, $a = \arccos(1/10) = 1{,}47062....$ Por lo tanto, $f(\overline{x}) = (a^*, 0) \in [\pi/4, 3\pi/2] \times [-1, 1]$, $\overline{x} \in X$ es única ■

Lemma 29 *Para el Ejemplo 5.3.9, $V \in C^2(X; \mathbb{R})$ y $f \in C^1(X; A)$.*

Proof. Nótese que $det(G^{\infty}_{aa}(\overline{x}, \overline{a})) > 0$, para $\overline{x} \in X$, $a \in (1{,}47062..., (3\pi)/2)$, $b \in [-1, 1]$, $\overline{x} = (x, y)$ y $\overline{a} = (a, b)$. Entonces las Hipótesis 8 y 9 se cumplen y el Lema 5.3.13 se sigue. ■

0.11. Conclusiones del Capítulo

En este capítulo se analizó el problema de la diferenciabilidad de la función de valores óptimos y del minimizador, correspondientes a un problema de optimización general. También, se dedujo una relación funcional para la derivada de la función de valores óptimos, conocida en la literatura de Economía como Fórmula de la Envolvente. Dicha fórmula se usó para resolver algunos problemas de la teoría de control óptimo. Lo anterior se hizo de forma directa, en el algoritmo de iteración de valores. El procedimiento se puede resumir en los puntos siguientes:

(Recuérdese que V denota la función de valores óptimos y f denota la política óptima, correspondientes al problema de control óptimo estocástico).

(a) Se aplica el Teorema 5.1.2 al algoritmo de iteraciones de valores.

(b) Con la aplicación del Teorema 5.1.2, se obtiene una relación de recurrencia para las derivadas de las funciones de iteración de valores. Integrando dichas relaciones, se obtienen las funciones de valores óptimos correspondientes.

(c) Para determinar V, tomamos el límite de las funciones de iteración de valores.

(d) Por último, para determinar f, sustituimos V en la Ecuación de Programación Dinámica.

Finalmente en este capítulo se presentó el análisis del problema de la diferenciabilidad para problemas de control estocástico con horizonte infinito, se dan condiciones estructurales sobre el modelo de control para garantizar la diferenciabilidad de clase C^2 de V y la diferenciabilidad de clase C^1 de f. También, se proporcionó un ejemplo que cumple las hipótesis impuestas al modelo, y se presentó una sección para el estudio de problemas no convexos.

Conclusiones y Problemas Abiertos

En la tesis se trató con la teoría de Procesos de Control de Markov. En ella se trabajaron problemas a tiempo discreto, con horizonte infinito. El criterio de rendimiento que se utilizó, para evaluar la calidad de las políticas admisibles fue el de costo total descontado. En el trabajo se usa la notación siguiente. Para problemas estocásticos la función de valores óptimos es denotada por V y la correspondiente política óptima por f. Para el problema de control determinista se denota la función de valores óptimos por W y la política óptima por g. De forma análoga se denotan las funciones de iteración de valores V_n y W_n, $n = 1, 2, ...$, y los minimizadores f_n y g_n, $n = 1, 2, ...$, para cada problema.

0.12. Conclusiones

La teoría de PCMs es presentada en el Capítulo 2, dando sólo los resultados necesarios para el desarrollo subsiguiente de la tesis. También se proporcionan las referencias para justificar los resultados principales de esta teoría. En el Capítulo 3 se dan, de manera general, los problemas que se trabajan en la tesis; además se presenta una sección donde se proveen los antecedentes referentes a los problemas que se trabajaron en ésta.

En los capítulos siguientes se presentó la contribución que se hace a la teoría de PCMs descontados.

En el Capítulo 4 se presentó, para PCMs deterministas, una ecuación funcional que caracteriza a la solución del problema de control óptimo (i.e., la función de valores óptimos y la política óptima), conocida como Ecuación de Euler. La EE se dió en términos de las derivadas de las funciones de iteración de valores, aunque es posible dar una versión en términos de la política óptima. La ventaja de plantear la EE en términos de las funciones de iteración de valores es que permite resolverla de manera inductiva. Para determinar la EE se usó la diferenciabilidad de primer orden de la política óptima, y la diferenciabilidad de segundo orden de la función de valores óptimos.

El Capítulo 5 presenta el tema de la diferenciabilidad en PCMs descontados. En dicho capítulo se dan condiciones para garantizar la diferenciabilidad de las funciones de iteración de valores y de los minimizadores correspondientes, después se procede a caracterizar la diferenciabilidad para la función de valores óptimos y la política óptima.

En resumen los resultados obtenidos en la tesis, son los siguientes:

(a) La diferenciabilidad de orden dos para V_n, V, W_n y $W, n \geq 1$.

(b) La diferenciabilidad de primer orden de f_n, f, g_n y $g, n \geq 1$.

Lo anterior se hizo suponiendo, principalmente, la diferenciabilidad de la función de costo y de la dinámica del sistema. Además, para determinar la diferenciabilidad de V, se usó el Teorema de Cambio de Variable (véase Apéndice B) y el Teorema de la Función Implícita (véase Apéndice B).

(c) Se determinaron condiciones para vincular problemas controlados estocásticos con problemas controlados deterministas. Para garantizar lo anterior, la principal condición fue suponer que para el problema determinista se conocen la función de valores óptimos W y la política óptima g. Para garantizar dicha condición se propone determinar a W y a g mediante el uso de la Ecuación de Euler, pero en general se puede obtener por cualquier otro procedimiento como, por ejemplo, mediante el uso directo de la EPD.

Las consecuencias que se tienen después de haber hecho un análisis de los resultados obtenidos son las siguientes:

(a) Se obtuvo una ecuación funcional (conocida como Ecuación de Euler) que caracteriza a las funciones de valores óptimos W_n. Además, se propuso un nuevo método iterativo para resolver dicha ecuación.

(b) A partir de la solución de un problema controlado determinista (donde se conoce a W y a g) se mostró que es posible inducir un problema controlado estocástico, el cual se resuelve utilizando la solución del determinista.

(c) Al aplicar la diferenciabilidad al algoritmo de iteración de valores se obtuvo una fórmula que caracteriza a la derivada de las funciones de iteración de valores V_n en términos de los minimizadores f_n. Dicha relación es posible usarla para resolver el problema de control óptimo, a partir de V_n'.

0.13. Problemas Abiertos

Los problemas que se consideran como consecuencia del trabajo, y se espera continuar su análisis, son los siguientes.

En referencia a la Ecuación de Euler (EE), tenemos:

(a) Para un MCM M donde la ecuación en diferencias presenta la forma (27), determinar la EE en términos de la política óptima f, o la función de valores óptimos V, y proponer un método de solución. Una opción para resolverla es utilizando un método iterativo, dando la EE en términos de las funciones de iteración de valores y seguir el procedimiento que se analizó en el Capítulo 2, para problemas deterministas.

(b) Encontrar procedimientos numéricos (programables) que den una solución aproximada de la EE.

(c) Determinar condiciones suficientes para caracterizar a la política óptima, a partir de la EE. Es decir, si una política estacionaria f satisface la EE entonces es óptima.

(d) En general, la Ecuación de Euler se ha trabajado para el caso descontado, tanto en versiones deterministas como estocásticas, salvo en el artículo de Brock & Mirman ([12]), en donde se analiza el problema de recompensa total, para problemas de crecimiento económico. Lo que se propone es estudiar la Ecuación de Euler bajo otros criterios de rendimiento, i.e. costo total y costo promedio. En el caso de recompensa promedio, la metodología con la cual se planea trabajar este problema consiste en revisar los artículos referentes a la Ecuación de Euler para el caso descontado (véase Aoki [2], Brock & Mirman [12], De la Fuente [22], Duffie [23], Levhari & Srinivasan [36], Mirman [37] y Stokey & Lucas [49]) y además el artículo Montes-de-Oca ([38]), en el cual se trabaja el caso promedio, vía el caso descontado.

En lo que respecta al estudio de la diferenciabilidad en PCMs consideramos trabajar en un futuro la siguiente clase de problemas.

(a) Estudiar la diferenciabilidad de la solución óptima para los criterios como son el costo total esperado y el costo promedio esperado.

(b) Hacer un estudio del Principio del Mínimo (Máximo) (véase [4], [5], [6], [13], [14], [15], [26], [27], [28], [29], [39] y [40]) vía la diferenciabilidad de la función de valores óptimos. Esto ya ha sido hecho en algunos casos específicos, véase, por ejemplo, Clarke ([15]) y Pontryagin et.al. ([40]).

(c) También vincular el Principio del Mínimo (Máximo) con la EE, para que juntos hagan posible un estudio más simple de problemas de control óptimo.

Apéndice: Notación

A : espacio de acciones.
$A(x)$: espacio de acciones admisibles en el estado del sistema x.
X : espacio de estados.
α : factor de descuento.
x_t : estado del sistema en el tiempo de observación t.
a_t : acción aplicada al sistema en el tiempo de observación t.
$\mathbb{K}$: conjunto de parejas estado-acción admisible.
$P(x_{t+1} \in B | x_t = x, a_t = a)$: ley de transición del sistema.
$c(x_t, a_t)$: costo recibido de x_t y a_t.
$\mathbb{H}_t$: espacio de historias admisibles hasta el tiempo t.
Π : conjunto de políticas de control.
$\mathbb{F}$: conjunto de políticas estacionarias.
$\mathbb{M}$: Conjunto de Políticas Markovianas.
$v(\pi, x)$: función objetivo, costo total descontado.
$V(\cdot)$: función de valores óptimos.
$V_n(\cdot)$: funciones de iteración de valores, $n \geq 0$.
$w(\pi, x)$: función objetivo, costo total descontado determinista.
$W(\cdot)$: función de valores óptimos determinista.
θ_ν : derivada parcial de θ con respecto a ν.
$\theta_{\nu\nu}$: segunda derivada parcial de θ con respecto a ν.
$C^\phi(\tilde{N}; Z)$: conjunto de funciones con derivada continua hasta de orden ϕ definidas en un conjunto $\tilde{N}$ y con valores en un conjunto Z.
$int(B)$: conjunto de puntos interiores de B.
$I_B(\cdot)$: función indicadora de B.

Apéndice: Resultados Auxiliares

Theorem 30 *(*Teorema del Valor Intermedio*) Sean $A \subset \mathbb{R}^n$ y $f : A \to \mathbb{R}$ continua. Supóngase que $K \subset A$ es conexo y $x, y \in K$. Para cada número $c \in \mathbb{R}$ tal que $f(x) \leq c \leq f(y)$, existe un punto $z \in K$ tal que $f(z) = c$.*

Proof. Véase Teorema 4.23 en Rudin [42] p. 100. ■

Theorem 31 *(*Teorema de Rolle*) Supóngase que $f : [a, b] \to \mathbb{R}$ es continua, que la derivada existe en el intervalo abierto y que $f(a) = f(b) = 0$. Entonces existe un punto c en (a, b) tal que $f'(c) = 0$.*

Proof. Véase Teorema 5.12 en Rudin [42] p. 116. ■

Theorem 32 *(*Teorema del Valor Medio*) Supóngase que $f : [a, b] \to \mathbb{R}$ es una función continua y que tiene una derivada en el intervalo abierto (a, b). Entonces existe un punto c en (a, b) tal que*

$$f(b) - f(a) = f'(c)(b - a).$$

Proof. Véase Teorema 5.9 en Rudin [42] p. 114. ■

Theorem 33 *(*Teorema de la Función Implícita*) Sean $A \subset \mathbb{R}^p \times \mathbb{R}^m$ un conjunto abierto y $F : A \to \mathbb{R}^m$ una función de clase $C^p(A; \mathbb{R}^m)$ (i.e. F tiene p derivadas continuas). Supóngase que $(x_0, y_0) \in A$ y $F(x_0, y_0) = 0$. Sea*

$$J = \begin{vmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ . & . & . \\ . & . & . \\ . & . & . \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{vmatrix}$$

evaluado en (x_0, y_0)*, donde* $F = (F_1, ..., F_m)$*. Supóngase que* $J \neq 0$*. Entonces existe una vecindad abierta* $U \subset \mathbb{R}^n$ *de* x_0 *y una vecindad* Λ *de* y_0 *y una única función* $f : U \to \Lambda$ *tal que*

$$F(x, f(x)) = 0,$$

para todo $x \in U$*. Además,* F *es de clase* C^p*.*

Proof. Véase Teorema 9.28 en Rudin [42] p. 242. ■

Theorem 34 *(*Teorema de Cambio de Variable*) Sea* T *una función continuamente diferenciable definida en conjunto abierto* U *con valores sobre un conjunto* V*. Supóngase que* T *es inyectiva y que su jacobiano,* $J(x)$*, es distinto de cero para todo* $x \in X$*. Si* f *es no negativa, entonces*

$$\int_U f(T(x)) \, |J(x)| \, dx = \int_V f(y) dy.$$

Proof. Véase el Teorema 17.2 en Billingsley [10] p. 225. ■

Bibliografía

[1] C. D. Aliprantis and O. Burkinshaw, Principles of Real Analysis. Academic Press, Inc (1998).

[2] M. Aoki, Optimization of Stochastic Systems. Academic Press Inc (1989).

[3] A. Araujo and J. A. Scheinkman, Smoothness, comparative dynamics, and the turnpike property. Econometrica, V. 45 N. 3, (1977), 601-620.

[4] A. Araujo and J. A. Scheinkman, Maximum principle and transversality condition for concave infinite horizon economic models. Journal of Economic Theory, V. 30 N. 1, (1983), 1-16.

[5] V. I. Arkin and I. V. Evstigneev, Stochastic Models of Control and Economic Dynamics. Academic Press Inc (1987).

[6] V. I. Arkin and L. M. Krechetov, A Stochastic maximum principle in control problems with discrete time. Lectures Notes in Mathematics, In Proceedings of the third Japan-USSR Symposium on Probability Theory, N 550, Springer-Verlag, (1976), 692-712.

[7] L. M. Benveniste and J. A. Scheinkman, On the differentiability of the value function in dynamic models of economics. Econometrica, V. 47 N. 3, (1979), 727-732.

[8] R. Bellman, Dynamic Programming. Dover (2003).

[9] D. P. Bertsekas, Dynamic Programming: Deterministic and Stochastic Models. Prentice-Hall, Inc., New Jersey (1987).

[10] P. Billingsley, Probability and Measure. A Wiley-Interscience Publications, JohnWiley & Sons (1995).

[11] L. Blume, D. Easley and M. O'Hara, Characterization of optimal plans for stochastic dynamic programs. Journal of Economic Theory, V. 28 N. 2, (1982), 221-234.

[12] W. A. Brock and L. J. Mirman, Optimal economic growth and uncertainty: The no discounted case. International Economic Review, V. 14 N. 3, (1973), 560-573.

[13] M. D. Canon, C. D. Cullum and E. Pollak, Constrained minimization problems in finite-dimensional spaces. SIAM Journal on Control and Optimization, V. 4 N.3, (1966), 528-547.

[14] M. D. Canon, C. D. Cullum and E. Pollak, Theory of Optimal Control and Mathematical Programming. McGraw-Hill (1970).

[15] F. H. Clarke and R. B. Vinter, The relationship between the maximum principle and dynamic programming. SIAM Journal on Control and Optimization, V. 25 N. 5, (1987), 1291-1311.

[16] D. Cruz-Suárez, R. Montes-de-Oca and F. Salem-Silva, Conditions for the uniqueness of optimal policies of discounted Markov decision processes. Mathematical Methods of Operations Research, V. 60 N. 3, (2004), 415-436.

[17] H. Cruz-Suárez, Análisis de problemas de control estocástico vía problemas de control determinista. Memorias de la Tercera Conferencia Iberoamericana en Sistemas, Cibernética e Informática CISCI'04, Orlando, Fl., V. 3, (2004), 66-71.

[18] H. Cruz-Suárez, La diferenciabilidad en procesos de decisión de Markov. Memorias de las Jornadas del Posgrado de la División de Ciencias Básicas e Ingeniería, Posgrado en Matemáticas, (2004), 167-170.

[19] H. Cruz-Suárez, Procesos de decisión de Markov descontados: soluciones óptimas mediante problemas de control determinista diferenciables, Revista Iberoamericana de Sistemas Cibernética e Informática, V. 2 N. 1, (2005). (Disponible en la página web: http://www.iiisci.org/journal/risci/.)

[20] H. Cruz-Suárez and R. Montes-de-Oca, Discounted Markov control processes induced by deterministic systems. Aceptado para publicación en Kybernetika (Prague), (2006).

[21] H. Cruz-Suárez and R. Montes-de-Oca, An envelope theorem and some applications to discounted Markov decision processes. Aceptado para publicación en Mathematical Methods of Operations Research, (2006).

[22] A. De la Fuente, Mathematical Methods and Models for Economists. Cambridge, University Press, New York (2000).

[23] D. Duffie, Security Markets. Academic Press, Inc, Boston (1988).

[24] J. Durán, On dynamic programming with unbounded returns. Economic Theory, V. 15 N. 2, (2000), 339-352.

[25] E. B. Dynkin and A. A. Yushkevich, Controlled Markov Processes. Springer-Verlag, New York (1979).

[26] L. T. Fan and C. S. Wang, The Discrete Maximum Principle. Wiley, New York (1974).

[27] H. Halkin, A maximum principle of Pontryagin type for systems described by nonlinear difference equations. SIAM Journal on Control and Optimization, V. 4 N. 1, (1966), 90-111.

[28] H. Halkin, Necessary conditions for optimal control problems with infinite horizons. Econometrica, V. 42 N. 2, (1974), 267-272.

[29] A. Haurie and S. P. Sethi, Decision and forecast horizons, agreeable plans and the maximum principle for infinite horizon control problems. Operations Research Letters, V. 3 N. 5, (1985), 261-265.

[30] B. Heer and A. Maussner, Dynamic General Equilibrium Modelling: Computational Method and Application, Springer-Verlag, Berlin (2005).

[31] O. Hernández-Lerma, Adaptive Markov Control Processes. Springer-Verlag, New York (1989).

[32] O. Hernández-Lerma and J. B. Lasserre, Value iteration and rolling plans for Markov control processes with unbounded rewards. Journal of Mathematical Analysis and Applications, V. 177 N. 1, (1993), 38-55.

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

[34] K. L. Judd, Numerical Methods in Economics. The MIT Press (1998).

[35] C. Le Van and L. Morhaim, Optimal growth models with bounded or unbounded returns: a unifying approach. Journal of Economic Theory, V. 105 N. 1, (2002), 158-187.

[36] D. Levhari and T. N. Srinivasan, Optimal savings under uncertainty. Review of Economic Studies, V. 36 N. 106, (1969), 153-164.

[37] L. J. Mirman, Dynamic models of fishing: a heuristic approach. Control Theory in Mathematical Economics, Editors Pan-Tai Liu & JG Sutinen, Marcel Dekker Inc., New York (1979), 39-73.

[38] R. Montes-de-Oca, The average cost optimality equation for Markov control processes on Borel spaces. Systems & Control Letters, V. 22 N. 5, (1994), 351-357.

[39] T. N. Pham, Feedback under control-dependent Markov disturbances: A discrete maximum principle. Journal of Mathematical Analysis and Applications, V. 197 N. 2, (1996), 325-340.

[40] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, The Mathematical Theory of Optimal Processes. Interscience Publishers (1962).

[41] J. L. Rincón-Zapatero and C. Rodríguez-Palmero, Existence and uniqueness of solutions to the Bellman equation in the unbounded case. Econometrica, V. 71 N. 5, (2003), 1519-1555.

[42] W. Rudin, Principles of Mathematical Analysis. Third Edition, McGraw Hill, New York (1987).

[43] M. S. Santos, Smoothness of the policy function in discrete time economic models. Econometrica, V. 59 N. 5, (1991), 1365-1382.

[44] M. S. Santos, Smooth dynamics and computation in models of economic growth. Journal of Economic Dynamics and Control, V. 18 N. 3-4, (1994), 879-895.

[45] M. S. Santos, Numerical solution of dynamic economic models. Handbook of Macroeconomic, Editors J. B. Taylor and M. Woodford, North Holland, Amsterdam (1999).

[46] M. S. Santos, Accuracy of numerical solutions using the Euler equation residuals. Econometrica, V. 68 N. 6, (2000), 1377-1402.

[47] M. S. Santos and J. Vigo, Error bounds for a numerical solution for dynamic economic models. Applied Mathematics Letters, V. 9 N. 4, (1996), 41-45.

[48] M. S. Santos and J. Vigo, Analysis of numerical dynamic programming algorithm applied to economic models. Econometrica, V. 66 N. 2, (1998), 409-426.

[49] N. L. Stokey and R. E. Lucas, Recursive Methods in Economic Dynamics. Harvard University Press Massachusetts (1989).

Casa abierta al tiempo
UNIVERSIDAD AUTÓNOMA METROPOLITANA

ACTA DE DISERTACIÓN PÚBLICA

No. 00004

LA DIFERENCIABILIDAD EN PROCESOS DE CONTROL DE MARKOV DESCONTADOS

En México, D.F., se presentaron a las 11:00 horas del día 8 del mes de diciembre del año 2006 en la Unidad Iztapalapa de la Universidad Autónoma Metropolitana, los suscritos miembros del jurado:

DR. ROLANDO CAVAZOS CADENA

DR. JUAN GONZALEZ HERNANDEZ

DR. EVGUENI GORDIENKO

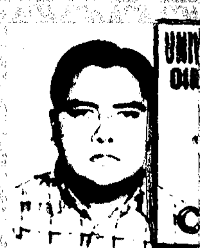
DR. ANDREI NOVIKOV

DR. JOSE RAUL MONTES DE OCA MACHORRO

Bajo la Presidencia del primero y con carácter de Secretario el último, se reunieron a la presentación de la Disertación Pública cuya denominación aparece al margen, para la obtención del grado de:

DOCTOR EN CIENCIAS (MATEMÁTICAS)

DE: HUGO ADAN CRUZ SUAREZ



UNIVERSIDAD AUTONOMA METROPOLITANA
DIRECCION DE SISTEMAS ESCOLARES
Casa abierta al tiempo

y de acuerdo con el artículo 78 fracción IV del Reglamento de Estudios Superiores de la Universidad Autónoma Metropolitana, los miembros del jurado resolvieron:

Aprobar

Acto continuo, el presidente del jurado comunicó al interesado el resultado de la evaluación y, en caso aprobatorio, le fue tomada la protesta.

HUGO ADAN CRUZ SUAREZ
FIRMA DEL ALUMNO

REVISÓ

LIC. JULIO CESAR DE LARA ISASSI
DIRECTOR DE SISTEMAS ESCOLARES

DIRECTORA DE LA DIVISIÓN DE CBI	PRESIDENTE	VOCAL
DRA. VERONICA MEDINA BANUELOS	DR. ROLANDO CAVAZOS CADENA	DR. JUAN GONZALEZ HERNANDEZ

VOCAL	VOCAL	SECRETARIO
DR. EVGUENI GORDIENKO	DR. ANDREI NOVIKOV	DR. JOSE RAUL MONTES DE OCA MACHORRO