La Ecuación de Euler en Procesos de Decisión de Markov Descontados

Gabriel Zacarías Espinoza

Índice general

Capítulo 1

INTRODUCCIÓN

Este trabajo está relacionado con la teoría de Procesos de Decisión de Markov (PDMs) a tiempo discreto (véase [4], [5], [6], [15], [16], y [23]). Los PDMs son procesos que se observan de forma periódica bajo incertidumbre en sus movimientos. Se describen mediante un modelo conocido como Modelo de Control de Markov (MCM) cuyas componentes permiten caracterizar su desarrollo en el transcurso del tiempo.

Un proceso de decisión de Markov se describe de la forma siguiente: un sistema es observado de forma discreta por un controlador en cada instante t ($t = 0, 1, ...$) y en ese momento el controlador selecciona una acción admisible al estado actual que presenta el sistema y, como consecuencia de esto se obtiene una recompensa (o se paga un costo) y, mediante una ley de transición prefijada el sistema transita al siguiente estado.

A la sucesión de acciones que el proceso genera se le llama política. Una manera de evaluar la calidad de la política es mediante un criterio de rendimiento. En este trabajo se considera como criterio de rendimiento la recompensa total descontada con horizonte finito e infinito (véase [15]). El problema de control óptimo (PCO) consiste en encontrar una política que optimice el criterio de rendimiento.

Una manera de solucionar el PCO es mediante Programación Dinámica (PD) iniciada a mediados de los años 50 por Richard Bellman (véase [4]), la cual consiste en llevar el problema de control a otro equivalente, el cual está dado mediante una ecuación funcional conocida como Ecuación de Programación Dinámica (EPD). Dicha ecuación caracteriza a la función de valor óptimo, la cual se aproxima mediante el algoritmo de iteración de valores.

La finalidad del trabajo de tesis consiste en estudiar versiones de la Ecuación de Euler (EE) en tiempo discreto para el problema de control óptimo. La EE es una ecuación funcional que caracteriza a la política óptima o a la función de valor óptimo. La EE tiene sus orígenes en el cálculo de variaciones (véase [22]). Históricamente, Leonhard Euler introdujo un procedimiento matemático general para la investigación sistemática de problemas de variaciones (véase [22]). Los matemáticos consideran este evento como el comienzo de una de las más importantes ramas de las matemáticas, el cálculo de variaciones. Los problemas asociados al cálculo de variaciones consisten en optimizar cantidades donde las variables son curvas, superficies, etcétera, es decir, extremos de funcionales. Haciendo uso de las nociones de derivadas y de ciertos lemas básicos en teoría de optimización se encuentran condiciones suficientes para tener un extremo de una funcional, lo cual nos lleva al estudio de la EE.

Para el objetivo de la tesis se estudian propiedades cualitativas como concavidad, unicidad, continuidad y diferenciabilidad, sobre las funciones de iteración de valores, la política óptima y la función de valor óptimo. Estas propiedades permiten establecer condiciones para obtener la contribución principal de este trabajo, una versión de la EE en el contexto de PDMs. Esta versión permite caracterizar las derivadas de las funciones de iteración de valores de manera recursiva, la cual a diferencia de PD es más fácil de implementar. Este procedimiento está basado en una versión para dinámicas deterministas implementada en [9], [12] y [28]. La EE obtenida también permite usar algoritmos computacionales (véase [14]) para facilitar los cálculos de las funciones de iteración de valores o para aproximar a la política óptima.

El presente trabajo aborda un modelo de consumo, producción e inversión (véase [7], [12], [18], [20] [21] y [26]), el cual es modelado mediante PDMs. Para dicho modelo se proveen condiciones suficientes para que una política sea óptima vía una versión de la EE. Las condiciones suficientes para la EE han sido analizadas por: De la Fuente [12], Levhari [20] y Mirman [21]. La metodología consiste en hacer uso de la EPD y de la diferenciabilidad de la función de valor (véase [10]) junto con condiciones de concavidad (véase [12]).

La organización de la tesis es de la siguiente manera: en el Capítulo 2 se presentan los conceptos básicos de PDMs, el problema de control óptimo y Programación Dinámica. En el Capítulo 3 se estudian propiedades cualitativas sobre el Modelo de control de Markov, en el Capítulo 4 se presenta la contribución principal de esta tesis: una versión de la EE en el contexto de PDMs, junto con ejemplos modelados usando PDMs y resueltos vía este método. En el Capítulo 5 se muestran ejemplos resueltos vía algoritmos computacionales utilizando la EE. Finalmente, se incluye un apéndice sobre resultados establecidos de funciones semicontinuas y cóncavas.

Capítulo 2

PROCESOS DE DECISIÓN DE MARKOV

El objetivo principal en este capítulo es introducir formalmente los conceptos básicos de Procesos de Decisión de Markov, y su interpretación. En la primera sección se presentan los temas siguientes: modelo de control de Markov, políticas de control, criterio de rendimiento y el problema de control óptimo. Consecutivamente, se presenta una herramienta fundamental en el análisis de un problema de control óptimo, conocida como programación dinámica, la cual permite determinar el valor máximo de la función objetivo, así como las estrategias que conforman la política óptima.

2.1. Procesos de Decisión de Markov

Un Proceso de Decisión de Markov (PDM) es utilizado para modelar un sistema que es observado de forma discreta en el tiempo bajo incertidumbre en su movimiento. Para modelar un sistema dinámico por medio de un PDM es necesario caracterizar las siguientes componentes:

Definición 2.1.1 *Un Modelo de Control de Markov (MCM), estacionario, a tiempo discreto, consiste de una quíntupla:*

$$(X, A, \{\, A(x) |\, x \in X\}, Q, r),$$

donde, X y A son espacios de Borel no vacíos, llamados espacio de estados y espacio de acciones (o controles), respectivamente. Para cada $x \in X$, $A(x)$ es un subconjunto medible y no vacío de A conocido como conjunto de acciones *admisibles para el estado $x \in X$. El conjunto $\mathbb{K}$ de parejas de estados-acciones admisibles, está definido por*

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

$Q(\cdot|\cdot)$ es un kérnel estocástico definido en X dado $\mathbb{K}$(véase [15]), llamado ley de transición. Por último $r : \mathbb{K} \to \mathbb{R}$ es una función medible llamada función de recompensa (o costo) en un paso.

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

a) se recibe una recompensa (o se paga un costo) $r(x,a)$ $(c(x,a))$ y,

b) 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 el proceso anteriormente descrito se repite.

Políticas de control. Para introducir el concepto de estrategia o política, considérese un MCM y *al espacio de las historias observadas del proceso hasta el tiempo* t, denotado por $\mathbb{H}_t$, y definido como $\mathbb{H}_0 = X$ y $\mathbb{H}_t = \mathbb{K} \times \mathbb{H}_{t-1}$, para $t = 1, 2, ...$ Un elemento de $\mathbb{H}_t$, llamado t-historia, es un vector de la forma

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

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

Definición 2.1.2 *Una política aleatorizada o simplemente política, es una sucesión $\pi = \{\pi_t\}_{t=0}^{\infty}$ de kérneles estocásticos, donde cada π_t está definida sobre A dado $\mathbb{H}_t$ y satisface que: $\pi_t(A(x_t)|\, h_t) = 1$ para toda $h_t \in \mathbb{H}_t$ y $t = 0, 1, 2, ...$ El conjunto de todas las políticas se denota por Π.*

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

Definición 2.1.3 *Una política $\pi \in \Pi$ es* ***Determinista Markoviana****, si existe una sucesión $\{f_t\} \subset \mathbb{F}$, donde*

$$\mathbb{F} = \{\, f : X \to A | \, f \text{ es medible y } f(x) \in A(x) \text{ para cada } x {\in} X\},$$

con la característica de que $\pi_t(\cdot | h_t)$ está concentrada en $f_t(x_t)$ para cada $h_t \in \mathbb{H}_t$ y $t = 0, 1, 2, ...$

Si existe $f \in \mathbb{F}$ tal que para toda $t = 0, 1, 2, ...$, $f_t = f$, entonces π se le conoce como política ***Determinista Markoviana Estacionaria*** *.*

Se entiende que $\pi(\cdot | h)$ está concentrada en $f(x)$ si, $\pi_t(C | h) = I_C(f(x))$ para cada $C \in \mathcal{B}(A)$. Donde $\mathcal{B}(A)$ denota la σ-álgebra de Borel de A e I_C es la función indicadora sobre C.

Observación 2.1.4 *Existen otras clases de políticas más generales que las anteriores como las Markovianas Aleatorizadas, Markovianas Aleatorizadas Estacionarias y Deterministas cuyas definiciones se pueden consultar en [15].*

Dados $\pi \in \Pi$ y $x_0 = x \in X$, por el Teorema de Ionescu-Tulcea (véase [3]) existe una única medida de probabilidad P_x^π definida en el espacio canónico $(\Omega, \mathcal{F})$, donde $\Omega := (X \times A)^\infty$ y $\mathcal{F}$ es su correspondiente σ-álgebra producto. El proceso estocástico $((\Omega, \mathcal{F}, P_x^\pi), \{x_t\})$ es llamado un *Proceso de Control de Markov a tiempo discreto* o *Proceso de Decisión de Markov (PDM).*

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

$$P_x^\pi(a_t \in C | \, h_t) = \pi_t(C | \, h_t), \tag{2.1}$$

$$P_x^\pi(x_{t+1} \in B | \, h_t, a_t) = Q(B | \, x_t, a_t). \tag{2.2}$$

De acuerdo a (2.2), la distribución del estado x_{t+1} sólo depende de la pareja estado-acción (x_t, a_t), dicha condición es una propiedad tipo Markov. La esperanza con respecto a P_x^π será denotada por E_x^π.

Observación 2.1.5 *En general, en lugar de dar $x_0 = x \in X$, se puede dar una medida de probabilidad ν sobre X, referida como distribución inicial y se cumple que*

$$P_v^{\pi}(x_0 \in B) = \nu(B),$$

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

Criterio de Rendimiento. Cada PDM estará dotado de una función real, llamada función objetivo o criterio de rendimiento, que medirá en algún sentido la calidad de cada política, a través de la sucesión de recompensas que genera.

Considérese un MCM fijo y un conjunto de políticas Π. Se define el criterio de rendimiento **Recompensa (o Costo) total descontada** para $\pi \in \Pi$ y $x \in X$, de la forma

$$v(\pi, x) := E_x^{\pi}\left[\sum_{t=o}^{N} \alpha^t r(x_t, a_t)\right],$$

con $\alpha \in (0,1)$. A α se le conoce como factor de descuento. Cuando $\alpha = 1$, el criterio es conocido como **Recompensa (o Costo) Total Acumulada.** Al entero positivo N se le llama horizonte del problema, el cual representa el número de etapas en el cual el sistema está operando y puede ser finito o infinito.

Definición 2.1.6 *Una política $\pi^* \in \Pi$, es óptima, si para cada $x \in X$*

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

La función V definida para $x \in X$ como

$$V(x) := \sup_{\pi \in \Pi} v(\pi, x),$$

se le llama función de valor óptimo.

El *problema de control óptimo* consiste en determinar a la política óptima.

Observación 2.1.7 *En el caso de costo, la función de valor óptimo se define por*

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

y una política $\pi^* \in \Pi$, es óptima, si para cada $x \in X$

$$v(\pi^*, x) = \inf_{\pi \in \Pi} v(\pi, x).$$

2.2. Programación Dinámica

El tema de esta sección es presentar una herramienta esencial para resolver el problema de control óptimo, conocida como Programación Dinámica. Este procedimiento permite determinar la función de valor óptimo, así como la política óptima. Bajo condiciones adecuadas sobre la función de recompensa y la ley de transición, la función de valor óptimo V se caracteriza mediante una ecuación funcional. Más aún, el conocimiento de V permite obtener una política óptima estacionaria. El criterio de rendimiento a considerar es el de recompensa total descontada, es decir, para $\pi \in \Pi$ y $x \in X$

$$v(\pi, x) := E_x^{\pi}\left[\sum_{t=o}^{N} \alpha^t r(x_t, a_t)\right],$$

donde $N \in \mathbb{N} \cup \{\infty\}$ y el factor de descuento $\alpha \in (0, 1)$. Para entender mejor la Programación Dinámica, se presenta una breve introducción en el contexto de problemas con horizonte finito.

2.2.1. Problemas con Horizonte Finito

Considérese como criterio de rendimiento a la recompensa total esperada con horizonte finito, definida como

$$v_N(\pi, x) := E_x^{\pi}\left[\sum_{t=0}^{N} \alpha^t r(x_t, a_t)\right].$$

Como el objetivo es determinar una política que optimice a $v_N(\pi, x)$. La idea fundamental de Programación Dinámica consiste en dividir el problema de optimizar la recompensa desde el tiempo 0 hasta N, en subproblemas más pequeños, en los que se requiere optimizar la recompensa recibida en un número de periodos menor a N. De manera intuitiva, los subproblemas son más sencillos de resolver en la medida de que involucra menos intervenciones por parte del controlador, que el problema original. Estos subproblemas se caracterizan en el teorema (2.2.1).

Teorema 2.2.1 *Defina para $x \in X$ y $t = 1, 2, ..., N$ a*

$$v_t(x) := \max_{a \in A(x)} \left\{ r(x, a) + \int v_{t-1}(y) Q(dy|\, x, a) \right\}, \tag{2.3}$$

con $v_0(x) = 0$.

Supóngase que estas funciones son medibles y que para cada $t = 1, 2, ..., N$, existe $f_t \in \mathbb{F}$, tal que para $x \in X$

$$v_t(x) = r(x, f_t(x)) + \int v_{t-1}(y) Q(dy|\, x, f_t(x)).$$

Entonces la política determinista de Markov $\pi^* = \{f_1, f_2, ..., f_N\}$ es óptima y la función de valor óptimo V es v_N.

Demostración. Véase [15]. ■

El teorema anterior muestra que v_t es el óptimo del subproblema desde el tiempo 0 a t, es decir,

$$v_t(x) = \sup_{\pi \in \Pi} v_t(\pi, x),$$

para toda $x \in X$ y $t = 0, 1, 2, ..., N$.

En resumen, se ha calculado para cada tiempo t la recompensa óptima en t-etapas, con esta interpretación de v_t es posible caracterizar a (2.3).

La relación (2.3) también es conocida como Ecuación de Programación Dinámica (EPD) junto con su condición inicial $v_0(x) = 0$ y su nombre es debido a Bellman (véase [4]).

Una consecuencia interesante es que la política óptima, que determina el algoritmo, resulta ser una política determinista Markoviana. De lo anterior se deduce que para seleccionar una acción a_t en el tiempo t solo es necesario conocer al estado actual x_t, es decir, no es indispensable emplear mecanismos aleatorios para seleccionar las acciones.

El teorema (2.2.1) tiene como suposición la existencia de $f_t \in \mathbb{F}$, las cuales maximizan el lado derecho de (2.3) en cada etapa. Esta suposición es referida como condición de selección medible y en algunos problemas se puede verificar directamente, pero desde un punto de vista teórico, es conveniente tener condiciones generales, las cuales se obtienen de los teoremas de selección medible (véase [17] y [25]). El siguiente lema es uno de ellos.

Lema 2.2.2 *Supóngase que para cada $x \in X$, $A(x)$ es un conjunto compacto y $u : \mathbb{K} \to \mathbb{R}$ una función medible. Sea U definida por*

$$U(x) := \sup_{a \in A(x)} u(x, a).$$

a) Si para cada $x \in X$, $u(x, \cdot)$ es semicontinua superiormente (u.s.c) (véase definición (7.1.1) del apéndice) en $A(x)$, entonces existe $f \in \mathbb{F}$ tal que para cada $x \in X$

$$U(x) = u(x, f(x)), \tag{2.4}$$

y U es medible.

b) Si para cada subconjunto abierto C de A, se tiene que el conjunto $\{x \in X|\, A(x) \subset C\}$ es abierto (es decir, la multifunción $x \to A(x) \subset A$ es semicontinua superiormente (u.s.c)), y u es u.s.c y acotada superiormente sobre $\mathbb{K}$, entonces existe $f \in \mathbb{F}$ que satisface (2.4) para todo $x \in X$. Además U es u.s.c y acotada superiormente en X.

Demostración. Véase [17] o [25]. ■

La siguiente condición sobre el modelo junto con el teorema (2.2.4) garantiza la validez del lema anterior, el cual se aplica en el teorema de Programación Dinámica (2.2.1) de manera iterativa para el cálculo de las funciones definidas en (2.3) junto con sus respectivos maximizadores $f_t \in \mathbb{F}$.

Condición 2.2.3 *Para cada $x \in X$*

a) $A(x)$ es compacto;

b) la función de recompensa $r(x, \cdot)$ es u.s.c en $A(x)$;

c) para cada función u medible y acotada sobre X, la función $\overline{u}(x, a) := \int u(y)Q(\,dy|\, x, a)$ definida sobre $\mathbb{K}$, es continua en $A(x)$.

Teorema 2.2.4 *Bajo la condición (2.2.3) se tiene que para cualquier función $u : X \to \mathbb{R}$ u.s.c y acotada superiormente, la función U definida para $x \in X$ como*

$$U(x) := \sup_{a\in A(x)} \left\{ r(x, a) + \int u(y)Q(\,dy|\, x, a) \right\},$$

es medible y existe un selector $f \in \mathbb{F}$ tal que, para toda $x \in X$,

$$U(x) = r(x, f(x)) + \int u(y)Q(\,dy|\, x, f(x)).$$

Además U también es u.s.c. En conclusión, se puede cambiar supremo por máximo.

Demostración. Sean u una función acotada superiormente y u.s.c en X y $x \in X$. Se sabe que la suma de dos funciones u.s.c es de nuevo una función u.s.c (véase [6]). Entonces por el lema (2.2.2) basta probar que la función

$$g(a) := \int u(y)Q(\,dy|\, x, a),$$

es u.s.c en $A(x)$.

Dado que u es acotada superiormente y u.s.c, u es límite de una sucesión $\{u_n\}$ decreciente de funciones ($u_n \downarrow u$) continuas y acotadas (véase proposición (7.1.2) del apéndice). Si $\{a^k\} \subset A(x)$ es una sucesión convergente a $a \in A(x)$, entonces se sigue que para toda n

$$\limsup_{k\to\infty} \int u(y)Q(\,dy|\,x,a^k) \leq \limsup_{k\to\infty} \int u_n(y)Q(\,dy|\,x,a^k),$$

por la condición (2.2.3 (c)) se sabe que

$$\limsup_{k\to\infty} \int u_n(y)Q(\,dy|\,x,a^k) = \int u_n(y)Q(\,dy|\,x,a),$$

implicando que

$$\limsup_{k\to\infty} \int u(y)Q(\,dy|\,x,a^k) \leq \int u_n(y)Q(\,dy|\,x,a),$$

si $n \to \infty$, por el teorema de la convergencia monótona se tiene que

$$\lim_{k\to\infty} \sup \int u(y)Q(\,dy|\,x,a^k) \leq \int u(y)Q(\,dy|\,x,a),$$

es decir, g es u.s.c y esto prueba lo deseado. ■

2.2.2. Problemas con Horizonte Infinito

Esta sección trabaja el problema de control óptimo cuando el horizonte es infinito. El criterio de rendimiento a considerar es el de recompensa total descontada, es decir, para $\pi \in \Pi$ y $x \in X$

$$v(\pi, x) := E_x^\pi \left[\sum_{t=o}^{\infty} \alpha^t r(x_t, a_t) \right].$$

El problema que se presenta es el de aproximar a V mediante problemas con horizonte finito n, cuyas correspondientes funciones de valor óptimo V_n (véase definición 2.2.5, abajo) se podrán determinar de manera recursiva (este método se conoce como aproximación sucesivas). Uno de los resultados principales de esta sección es la convergencia puntual de la sucesión $\{V_n\}$ a V.

Definición 2.2.5 *Las funciones de iteración de valores se definen de la forma siguiente, para $x \in X$ y $n = 1, 2, ...,$*

$$V_n(x) = \max_{a\in A(x)} \left\{ r(x,a) + \alpha \int V_{n-1}(y)Q(\,dy|\,x,a) \right\}, \tag{2.5}$$

con $V_0(x) = 0$.

Consideraciones Preliminares

Sean X un espacio de Borel y

$$\mathbb{B}(X) := \{u : X \to \mathbb{R} |\, u \text{ es una función medible y acotada}\},$$

con la norma del supremo

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

Se sabe que $(\mathbb{B}(X), \|\cdot\|)$ es un espacio de Banach (véase [24]).

Sea $w : X \to [1, \infty)$ una función medible conocida. Si $u : X \to \mathbb{R}$ es una función, se define la *w-norma* por

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

A w se le conoce como función de peso.

Una función $u : X \to \mathbb{R}$ es acotada, si $\|u\| < \infty$ y w-acotada, si $\|u\|_w < \infty$.

Proposición 2.2.6 *Sea*

$$\mathbb{B}_w(X) := \{u : X \to \mathbb{R} |\, u \text{ es una función medible y } w\text{-acotada}\}.$$

Entonces $(\mathbb{B}_w(X), \|\cdot\|_w)$ *es un espacio de Banach y* $\mathbb{B}(X) \subset \mathbb{B}_w(X)$*.*

Demostración. Sea $\{u_n\} \subset \mathbb{B}_w(X)$ una sucesión de Cauchy, entonces, $\{u_n/w\}$ es una sucesión de Cauchy en $\mathbb{B}(X)$. Como $\mathbb{B}(X)$ es un espacio de Banach, existe $u \in \mathbb{B}(X)$ tal que $\|u_n/w - u\| \to 0$ cuando $n \to \infty$. Además si $x \in X$, se tiene que

$$\left|\frac{u_n(x) - u(x)w(x)}{w(x)}\right| = \left|\frac{u_n(x)}{w(x)} - u(x)\right|,$$

esta última relación implica que $\|u_n - uw\|_w \le \|u_n/w - u\|$, entonces cuando $n \to \infty$ se concluye que $\|u_n - uw\|_w \to 0$. Además, $\|uw\|_w = \|u\|$ y por consiguiente $uw \in \mathbb{B}_w(X)$.

Por otro lado, sea $u \in \mathbb{B}(X)$, debido a que para cada $x \in X$, $1 \le w(x)$, entonces $|u(x)|/w(x) \le |u(x)|$, implicando que $\|u\|_w \le \|u\|$, es decir, $u \in \mathbb{B}_w(X)$. ■

Definición 2.2.7 *Sea* (Y, d) *un espacio métrico. Una función* $T : Y \to Y$ *es una contracción modulo* $\gamma \in (0, 1)$ *si para toda* $x, y \in Y$

$$d(T(x), T(y)) \le \gamma d(x, y).$$

Si existe $\widehat{y} \in Y$ *tal que* $T(\widehat{y}) = \widehat{y}$ *entonces* $\widehat{y}$ *se denomina punto fijo de* T*.*

Proposición 2.2.8 *Sean (Y, d) un espacio métrico completo y $T : Y \to Y$ una contracción modulo γ. Entonces T tiene un único punto fijo $\widehat{y}$. Además para cada $y \in Y$ y $n = 0, 1, 2, ...,$*

$$d(T^n(y), \widehat{y}) \leq \gamma^n d(y, \widehat{y}).$$

Demostración. Véase [24]. ■

La siguiente proposición da condiciones para que un operador T sea una contracción en $\mathbb{B}_w(X)$, para ello se requiere que este operador sea monótono, es decir, un operador T sobre $\mathbb{B}_w(X)$ es monótono, si para cualesquiera $u, v \in \mathbb{B}_w(X)$, con $u(x) \leq v(x)$, para cada $x \in X$, se tiene que $T(u)(x) \leq T(v)(x)$.

Proposición 2.2.9 *Sea $T : \mathbb{B}_w(X) \to \mathbb{B}_w(X)$ un operador monótono. Si existe $0 < \gamma < 1$ tal que para cada $u \in \mathbb{B}_w(X)$ y $\beta \in \mathbb{R}$:*

$$T(u + rw) \leq Tu + \gamma\beta w,$$

entonces T es una contracción módulo γ.

Demostración. Sean $u, v \in \mathbb{B}_w(X)$, nótese que $u \leq v + w \|u - v\|_w$ y $v \leq u + w \|u - v\|_w$.

Defínase a $\beta := \|u - v\|_w$, por la monotonía de T se sabe que $Tu \leq T(v + w\beta)$ y $Tv \leq T(u + w\beta)$. Por hipótesis, existe $0 < \gamma < 1$ tal que $T(v + \beta w) \leq Tv + \gamma\beta w$ y $T(u + \beta w) \leq Tu + \gamma\beta w$, implicando que $Tu \leq Tv + \gamma\beta w$ y $Tv \leq Tu + \gamma\beta w$, combinando estas desigualdades se llega a que $-\gamma w \|u - v\|_w \leq Tu - Tv \leq \gamma w \|u - v\|_w$, es decir,

$$|Tu - Tv| \leq \gamma w \|u - v\|_w,$$

lo cual implica que

$$\|Tu - Tu\|_w \leq \gamma \|u - v\|_w.$$

■

Ecuación de Programación Dinámica

A continuación se define al operador de Programación Dinámica, para el cual se verá que la función de valor óptimo V resulta ser un punto fijo de éste.

Definición 2.2.10 *Dada una función $u \in \mathbb{B}_w(X)$ y $\alpha \in (0,1)$, denótese al operador de Programación Dinámica por T_α, el cual está definido de la forma siguiente, para cada $x \in X$*

$$T_\alpha(u)(x) = \sup_{a \in A(x)} \left\{ r(x,a) + \alpha \int u(y) Q(\,dy|\, x, a) \right\},$$

siempre que la integral esté bien definida.

Si para cada $x \in X$, el supremo es alcanzado para alguna $a \in A(x)$, se puede escribir máximo en lugar de supremo.

Nótese que las funciones de iteración de valores (2.5) se pueden reescribir en términos del operador T_α, de la forma siguiente, para $n = 1, 2, ...,$

$$V_n = T_\alpha\left(V_{n-1}\right).$$

Las siguientes condiciones sobre el modelo de control junto con la proposición (2.2.9) permitirán mostrar que el operador de Programación Dinámica T_α es una contracción en $\mathbb{B}_w(X)$. La condición (2.2.11) asegura la existencia de maximizadores medibles para T_α similar a la dada en el caso finito (condición (2.2.3)). También, se considera una función de peso w para imponer una condición de crecimiento sobre la función de recompensa r. Dichas condiciones permitirán mostrar que T_α es un operador sobre $\mathbb{B}_w(X)$.

Condición 2.2.11 *Para cada estado $x \in X$*
a) $A(x)$ es un conjunto compacto;
b) la función de recompensa $r(x,a)$ es u.s.c en a, con $a \in A(x)$;
c) para toda $u \in \mathbb{B}(X)$, la función definida por $\overline{u}(x,a) := \int u(y) Q(\,dy|\, x, a)$ es continua en a, con $a \in A(x)$.

Condición 2.2.12 *Existen constantes no negativas $\overline{c}$ y β, con $1 \leq \beta < 1/\alpha$, y una función de peso w sobre X, tales que, para cada $x \in X$*
a) $\sup_{a \in A(x)} |r(x,a)| \leq \overline{c} w(x)$,
b) $\sup_{a \in A(x)} \int w(y) Q(\,dy|\, x, a) \leq \beta w(x)$.

Condición 2.2.13 *Para cada $x \in X$, la función $\overline{w}(x,a) := \int w(y) Q(\,dy|\, x, a)$ es continua en a, con $a \in A(x)$.*

El siguiente lema es similar al lema (2.2.2) de selección medible, pero para el espacio $\mathbb{B}_w(X)$.

Lema 2.2.14 *Supóngase para cada $x \in X$, $A(x)$ es un conjunto compacto y sea $u : \mathbb{K} \to \mathbb{R}$ una función medible. Defínase a U como*

$$U(x) := \sup_{a \in A(x)} u(x,a).$$

También, supóngase que la multifunción $x \to A(x)$ es u.s.c, para cada $x \in X$, la función $u(x,\cdot)$ es u.s.c en $A(x)$ y

$$\sup_{a \in A(x)} |u(x,a)| \leq kw(x),$$

donde la constante k es positiva y $w(\cdot)$ es una función de peso continua sobre X. Entonces existe un selector $f \in \mathbb{F}$ tal que para todo $x \in X$

$$U(x) = u(x, f(x)), \tag{2.6}$$

y U es medible. Además U es u.s.c y para cada $x \in X$

$$|U(x)| \leq kw(x),$$

es decir, $U \in \mathbb{B}_w(X)$.

Demostración. Aplique el inciso (b) del lema (2.2.2) a la función definida por $\widetilde{u}(x,a) := u(x,a) - kw(x)$ la cual resulta ser u.s.c y acotada superiormente, dado que es $\widetilde{u}(x,a) \leq 0$. ■

Si para cada $x \in X$ el supremo es alcanzado para alguna $a \in A(x)$ entonces se puede escribir máximo en lugar de supremo.

Lema 2.2.15 *Supóngase que las condiciones (2.2.11 (c)) y (2.2.13) se cumplen. Entonces para cada $x \in X$ y $u \in \mathbb{B}_w(X)$ la función $\overline{u}(x,a) := \int u(y) Q(dy|\, x, a)$ es continua en $a \in A(x)$.*

Demostración. Obsérvese que si $u \in \mathbb{B}_w(X)$ entonces para toda $x \in X$, $|u(x)| \leq \|u\|_w w(x)$, lo cual implica que la función $U := u + w\|u\|_w \in \mathbb{B}_w(X)$ y $U \geq 0$.

Sean $x \in X$ y $u \in \mathbb{B}_w(X)$. Como $U \geq 0$, se sabe que existe una sucesión $\{u_n\} \subset \mathbb{B}(X)$ tal que $u_n \to U$ de manera creciente ($u_n \uparrow U$) (véase [1]). Si $\{a^k\} \subset A(x)$ es una sucesión convergente a $a \in A(x)$, se sigue que para toda n

$$\liminf_{k \to \infty} \int U(y) Q(dy|\, x, a^k) \geq \liminf_{k \to \infty} \int u_n(y) Q(dy|\, x, a^k).$$

Por la condición (2.2.11 (c)) se sabe que

$$\liminf_{k\to\infty}\int u_n(y)Q(\,dy|\,x,a^k)=\int u_n(y)Q(\,dy|\,x,a),$$

implicando que

$$\liminf_{k\to\infty}\int U(y)Q(\,dy|\,x,a^k)\geq\int u_n(y)Q(\,dy|\,x,a),$$

si $n\to\infty$ y aplicando el teorema de la convergencia monótona (véase [1]) se obtiene que

$$\liminf_{k\to\infty}\int U(y)Q(\,dy|\,x,a^k)\geq\int U(y)Q(\,dy|\,x,a).$$

Como $\int U(y)Q(\,dy|\,x,a)=\|u\|_w\int w(y)Q(\,dy|\,x,a)+\int u(y)Q(\,dy|\,x,a)$ y por la condición (2.2.13) se sabe que la función $\int w(y)Q(\,dy|\,x,a)$ es continua en $a\in A(x)$, se concluye que

$$\liminf_{k\to\infty}\int u(y)Q(\,dy|\,x,a^k)\geq\int u(y)Q(\,dy|\,x,a). \tag{2.7}$$

Por otra parte, se sabe que

$$\limsup_{k\to\infty}\int u(y)Q(\,dy|\,x,a^k)=-\liminf_{k\to\infty}\left(-\int u(y)Q(\,dy|\,x,a^k)\right),$$

utilizando a (2.7) se obtiene que

$$\int u(y)Q(\,dy|\,x,a)\geq\limsup_{k\to\infty}\int u(y)Q(\,dy|\,x,a^k). \tag{2.8}$$

Finalmente por (2.7) y (2.8) se concluye que $\overline{u}(x,\cdot)$ es una función continua en $a\in A(x)$. ■

Proposición 2.2.16 *Supóngase que las condiciones (2.2.11), (2.2.12) y (2.2.13) se cumplen. Entonces*

a) El operador de Programación Dinámica T_α *es una contracción en* $\mathbb{B}_w(X)$ *con módulo* $\gamma:=\alpha\beta<1$.

b)Para cada $u\in\mathbb{B}_w(X)$ *existe* $f\in\mathbb{F}$, *tal que para cada* $x\in X$

$$T_\alpha u(x)=r(x,f(x))+\alpha\int u(y)Q(\,dy|\,x,f(x)). \tag{2.9}$$

Demostración. Sea $u \in \mathbb{B}_w(X)$. Por el lema (2.2.15) y bajo la condición (2.2.11 (b)), se sabe que la función

$$\overline{u}(x,a) := r(x,a) + \alpha \int u(y) Q(dy|\, x,a),$$

es u.c.s en $a \in A(x)$. De aquí por el lema (2.2.14) se concluye que $T_\alpha(u)$ es una función medible y existe $f \in \mathbb{F}$ tal que (2.9) se satisface. Por otra parte, por la condición (2.2.12) se tiene que para cada $x \in X$ y $a \in A(x)$

$$\begin{aligned} |\overline{u}(x,a)| &\leq \overline{c}w(x) + \alpha \|u\|_w \int w(y) Q(dy|\, x,a) \\ &\leq (\overline{c} + \alpha\beta \|u\|_w) w(x), \end{aligned}$$

lo cual implica que $\|T_\alpha u\|_w < \infty$. Por lo tanto $T_\alpha u \in \mathbb{B}_w(X)$, es decir, T_α es realmente un operador.

Es fácil ver que T_α es un operador monótono, es decir, si $u \leq v$ entonces $T_\alpha u \leq T_\alpha v$. Por definición de $T_\alpha u$ y por la Condición 2.2.12, se tiene que para cualquier $r \in \mathbb{R}$

$$T_\alpha(u + rw)(x) \leq T_\alpha u(x) + r(\alpha\beta) w(x).$$

Finalmente por la Proposición 2.2.9 se concluye que T_α es una contracción en $\mathbb{B}_w(X)$ con módulo $\gamma = \alpha\beta$. ■

El siguiente teorema, conocido como Teorema de Programación Dinámica para horizonte infinito, el cual garantiza que la función de valor óptimo V es un punto fijo del operador de Programación Dinámica T_α, la sucesión de funciones de iteración de valores $\{V_n\}$ converge geométricamente en w-norma a V y también la existencia de una política óptima, la cual resulta ser determinista estacionaria.

Teorema 2.2.17 *Supóngase que las Condiciones 2.2.11, 2.2.12 y 2.2.13 se cumplen y sea $\gamma = \alpha\beta$. Entonces*

a) La función de valor óptimo es un punto fijo del operador de programación dinámica T_α en el espacio $\mathbb{B}_w(X)$, es decir, para $x \in X$

$$V(x) = \max_{a \in A(x)} \left\{ r(x,a) + \alpha \int V(y) Q(dy|\, x,a) \right\} \tag{2.10}$$

y para $n = 1, 2, ...,$

$$\|V_n - V\|_w \leq \frac{\overline{c}\gamma^n}{1-\gamma}.$$

b) Además existe $f \in \mathbb{F}$ tal que para cada $x \in X$

$$V(x) = r(x, f(x)) + \alpha \int V(y) Q(\, dy|\, x, f(x)),$$

y la política determinista estacionaria f es óptima. Recíprocamente, si f es óptima entonces satisface la igualdad anterior.

Demostración. Por la Proposición 2.2.16 y por el Teorema del punto fijo de Banach (Proposición 2.2.8), se sabe que T_α tiene un único punto fijo $\widehat{u}$ en $\mathbb{B}_w(X)$ y si $u \in \mathbb{B}_w(X)$ entonces

$$\|T_\alpha^n u - \widehat{u}\|_w \leq \gamma^n \|u - \widehat{u}\|_w \,.$$

Así para probar el inciso (a), basta mostrar que $V \in \mathbb{B}_w(X)$ y que $\widehat{u} = V$.

Para ello sea $\pi \in \Pi$ y $x_0 = x \in X$. Entonces para toda $t = 0, 1, ...,$

$$E_x^\pi \left[w(x_t) \right] \leq \beta^t w(x), \tag{2.11}$$

y

$$\left| E_x^\pi \left[r(x_t, a_t) \right] \right| \leq \bar{c} \beta^t w(x). \tag{2.12}$$

En efecto, si $t = 0$ se sabe que (2.11) se satisface. Ahora, si $t \geq 1$, por (2.2) y la Condición 2.2.12 se tiene que

$$\begin{aligned} E_x^\pi \left[w(x_t) |\, h_{t-1}, a_{t-1} \right] &= \int w(y) Q(\, dy|\, x_{t-1}, a_{t-1}) \\ &\leq \beta w(x_{t-1}), \end{aligned}$$

y por propiedades de la esperanza condicional se llega a

$$E_x^\pi \left[w(x_t) \right] \leq \beta E_x^\pi \left[w(x_{t-1}) \right],$$

al iterar esta ultima desigualdad se obtiene que (2.11) es verdadera.

Similarmente, por la Condición 2.2.12 observe que para $t = 0, 1, ..., |r(x_t, a_t)| \leq \bar{c} w(x_t)$, y por (2.11) se tiene que (2.12) también se cumple.

Además se sigue de (2.12) y de cálculos directos que

$$|v(\pi, x)| \leq \sum_{t=0}^{\infty} \alpha^t \left| E_x^\pi \left[r(x_t, a_t) \right] \right| \leq \bar{c} w(x)/(1 - \gamma),$$

con $\gamma = \alpha\beta$. Dado que π y x son arbitrarías se concluye que $|V(x)| \leq \bar{c} w(x)/(1-\gamma)$, implicando que $\|V\|_w \leq \bar{c}/(1-\gamma)$ y por lo tanto $V \in \mathbb{B}_w(X)$.

Por otro lado, sean $\pi \in \Pi$, $x \in X$ y $u \in \mathbb{B}_w(X)$. Por definición de w-norma y (2.11), se tiene que

$$E_x^\pi |u(x_t)| \leq \|u\|_w E_x^\pi [w(x_t)] \leq \|u\|_w \beta^t w(x_t),$$

así cuando $t \to \infty$ se obtiene que

$$\lim_{t\to\infty} \alpha^t E_x^\pi [u(x_t)] = 0. \tag{2.13}$$

Ahora como $\widehat{u}$ es punto fijo de T_α y por la Proposición 2.2.16 (b), existe $f \in \mathbb{F}$, tal que para cada $x \in X$

$$T_\alpha \widehat{u}(x) = r(x, f(x)) + \alpha \int \widehat{u}(y) Q(\,dy|\, x, f(x)),$$

al iterar la igualdad anterior se tiene que para toda $n = 1, 2, ...,$

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

si $n \to \infty$, por (2.13) se sigue que

$$\widehat{u}(x) = E_x^f \left[\sum_{t=0}^{\infty} \alpha^t r(x_t, f(x_t)) \right] = v(f, x),$$

y por definición de V se concluye que

$$\widehat{u}(x) \leq V(x). \tag{2.14}$$

Por otra parte, dado que

$$\widehat{u}(x) = T_\alpha \widehat{u}(x) = \sup_{a \in A(x)} \left\{ r(x, a) + \alpha \int \widehat{u}(y) Q(\,dy|\, x, a) \right\},$$

se llega a que para toda $(x, a) \in \mathbb{K}$

$$\widehat{u}(x) \geq r(x, a) + \alpha \int \widehat{u}(y) Q(\,dy|\, x, a),$$

por (2.2) y la desigualdad anterior, se tiene que para $\pi \in \Pi$, $x \in X$ y $t = 0, 1, 2, ...,$

$$\alpha^t E_x^\pi [\, r(x_t, a_t) + \alpha \widehat{u}(x_{t+1}) - \widehat{u}(x_t)|\, h_t, a_t] \leq 0,$$

al tomar esperanza E_x^π y sumar desde $t = 0$ a $n - 1$, se obtiene que

$$\widehat{u}(x) \geq E_x^\pi \left[\sum_{t=0}^{n-1} \alpha^t r(x_t, a_t) \right] + \alpha^n E\left[{}_x^\pi \widehat{u}(x_n) \right].$$

Cuando $n \to \infty$ en la relación anterior, por (2.13) se llega a que $\widehat{u}(x) \geq v(\pi, x)$, y dado que π y x son arbitrarios, entonces

$$\widehat{u}(x) \geq V(x). \tag{2.15}$$

Por (2.14) y (2.15) se concluye que $\widehat{u} = V$.

Finalmente la existencia de $f \in \mathbb{F}$ se sigue de la Proposición 2.2.16 (b). Recíprocamente, si $f \in \mathbb{F}$, entonces por definición

$$v(f, x) = E_x^f \left[\sum_{t=0}^{\infty} \alpha^t r(x_t, a_t) \right],$$

al expandir a v y al aplicar la propiedad de Markov (2.2) se tiene que $v(f(x), x) = c(x, f(x)) + \alpha \int V(y) Q(dy|\, x, f(x))$. ■

2.2.3. Comentarios y Observaciones

En algunas aplicaciones la ley de transición Q es inducida por una ecuación de diferencias estocásticas de la forma siguiente:

$$x_{t+1} = F(x_t, a_t, \xi_t), \tag{2.16}$$

para $t = 1, 2, ...,$ $x_0 = x$ conocido. Donde $\{\xi_t\}$ es una sucesión de variables aleatorias (v.a.'s) independientes e idénticamente distribuidas (i.i.d.) tomando valores en un espacio S y con distribución común μ e independiente del estado x_0 y $F : \mathbb{K} \times S \to X$ es una función medible conocida. En este caso la ley de transición Q está dada por:

$$Q(B|\, x, a) = E\left[I_B(F(x, a, \xi)) \right],$$

$B \in \mathcal{B}(X)$, $(x, a) \in \mathbb{K}$, y E es la esperanza tomada de la variable aleatoria ξ.

Obsérvese que cuando la dinámica es deterministica, es decir, $x_{t+1} = F(x_t, a_t)$ con $F : \mathbb{K} \to X$, la ley de transición es:

$$Q(B|\, x, a) = I_B(F(x, a)), B \in \mathcal{B}(X).$$

Por el teorema de cambio de variable (véase [1] o [3]), para funciones medibles sobre X, se tiene que

$$E\left[v(x_{t+1})|\, x_t = x, a_t = a\right] = \int v(y) Q(\, dy|\, x, a) = E\left[v(F(x, a, \xi))\right],$$

en el sentido de que si una de las integrales existe, la otra también y son iguales.

Así las funciones de iteración de valores (2.5) toman la forma siguiente

$$V_n(x) = \max_{a \in A(x)} \left\{ r(x, a) + E\left[V_{n-1}(F(x, a, \xi))\right] \right\},$$

para $n \geq 1$ con $V_0 = 0$. Además bajo las Condiciones 2.2.11, 2.2.12 y 2.2.13 la función de valor óptimo satisface:

$$V(x) = \max_{a \in A(x)} \left\{ r(x, a) + E\left[V(F(x, a, \xi))\right] \right\},$$

$x \in X$.

Capítulo 3

PROPIEDADES CUALITATIVAS

En este capítulo se estudian propiedades cualitativas sobre las funciones de iteración de valores, sus correspondientes maximizadores, la función de valor óptimo y la política óptima. A lo largo de este capítulo se supone que $X \subseteq \mathbb{R}^n$, $A \subseteq \mathbb{R}^m$ y para cada $x \in X$, $A(x)$ son conjuntos convexos. También se supondrá que la ley de transición esta inducida por una ecuación en diferencia como la dada en (2.16). Entonces las funciones de iteración de valores presentan la forma siguiente:

$$V_n(x) = \max_{a \in A(x)} \left\{ r(x,a) + \alpha E\left[V_{n-1}(F(x,a,\xi))\right] \right\},$$

con $V_0(x) = 0$, $x \in X$. Además, supóngase para cada $n = 1, 2, ...$, existe $f_n \in \mathbb{F}$ tal que para cada $x \in X$

$$V_n(x) = r(x, f_n(x)) + \alpha E\left[V_{n-1}(F(x, f_n(x), \xi))\right].$$

Sea $G : \mathbb{K} \to \mathbb{R}$ una función medible y supóngase que existe una función medible $h : X \to \mathbb{R}$ tal que

$$G(x,a) \leq h(x),$$

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

$$g(x) := \sup_{a \in A(x)} G(x,a),$$

$x \in X$.

3.1. Concavidad

Para esta sección se necesitan algunas características de la multifunción, $x \to A(x)$. Supóngase que X es dotado de un orden parcial, por ejemplo el orden lexicográfico.

a) La multifunción $x \to A(x)$ es creciente, si para $x, y \in X$ con $x < y$ se tiene que $A(x) \subseteq A(y)$.

b) La multifunción $x \to A(x)$ es convexa, si para cada $x, y \in X$ y $\lambda \in [0,1]$, se tiene que $\lambda a + (1-\lambda)\widetilde{a} \in A(\lambda x + (1-\lambda)y)$, con $a \in A(x)$ y $\widetilde{a} \in A(y)$.

Lema 3.1.1 *Supóngase que G es una función cóncava definida en $\mathbb{K}$ y que existe $f \in \mathbb{F}$, tal que para cada $x \in X$, $g(x) = G(x, f(x))$. Entonces el conjunto*

$$L := \{ h \in \mathbb{F} |\, g(x) = G(x, h(x)) \text{ para toda } x \in X \},$$

es convexo y g es una función cóncava. Además si G es estrictamente cóncava, entonces g también es estrictamente cóncava y L es un conjunto singular, es decir, se garantiza la unicidad del maximizador f.

Demostración. Obsérvese que L es no vació, ya que $f \in \mathbb{F}$. Sean $h_1, h_2 \in L$ y $\lambda \in (0,1)$, entonces para $x \in X$, se tiene que $G(x, h_1(x)) = g(x) = G(x, h_2(x))$. Como $A(x)$ es convexo, $\lambda h_1(x) + (1-\lambda)h_2(x) \in A(x)$ y como G es cóncava, se sigue que

$$\begin{aligned} g(x) &\geq G(x, \lambda h_1(x) + (1-\lambda)h_2(x)) \\ &\geq \lambda G(x, h_1(x)) + (1-\lambda)G(x, h_2(x)) \\ &= g(x), \end{aligned}$$

es decir, para toda $x \in X$, $g(x) = G(x, \lambda h_1(x) + (1-\lambda)h_2(x))$. Por lo tanto $\lambda h_1 + (1-\lambda)h_2 \in L$.

Por otra parte, sean $x, y \in X$ y $\lambda \in [0,1]$. Dado que G es cóncava, implica que para cada $a \in A(x)$ y $\widetilde{a} \in A(y)$

$$\lambda G(x, a) + (1-\lambda)G(y, \widetilde{a}) \leq G(\lambda x + (1-\lambda)y, \lambda a + (1-\lambda)\widetilde{a}),$$

y como $\lambda a + (1-\lambda)\widetilde{a} \in A(\lambda x + (1-\lambda)y)$ se sigue que $\lambda G(x, a) + (1-\lambda)G(y, \widetilde{a}) \leq g(\lambda x + (1-\lambda)y)$. Sea $\widetilde{a}$ fija, entonces para cada $a \in A(x)$, $\lambda G(x, a) \leq g(\lambda x + (1-\lambda)y) - (1-\lambda)G(y, \widetilde{a})$ implicando que

$$\lambda g(x) \leq g(\lambda x + (1-\lambda)y) - (1-\lambda)G(y, \widetilde{a}),$$

equivalentemente,

$$(1-\lambda)G(y,\widetilde{a}) \leq g(\lambda x + (1-\lambda)y) - \lambda g(x).$$

Como $\widetilde{a}$ es arbitraria se obtiene que

$$\lambda g(x) + (1-\lambda)g(y) \leq g(\lambda x + (1-\lambda)y),$$

es decir, g es cóncava.

Supóngase que G es estrictamente cóncava, entonces tomando las desigualdades involucradas como estrictas se concluye que g es estrictamente cóncava. Para probar que L es un conjunto singular, supóngase que existen $h_1, h_2 \in L$, tales que $h_1 \neq h_2$. Se sabe que para $\lambda \in (0,1)$, $\lambda h_1 + (1-\lambda)h_2 \in L$. Por lo tanto para $x \in X$

$$\begin{aligned} g(x) &= G(x, \lambda h_1(x) + (1-\lambda)h_2(x)) \\ &> \lambda G(x, h_1(x)) + (1-\lambda)G(x, h_2(x)) \\ &= g(x), \end{aligned}$$

contradicción. Por lo tanto, L es un conjunto singular. ■

Lema 3.1.2 *Supóngase que $G(\cdot, a)$ es una función (estrictamente) creciente en X para cada $a \in A$ y si la multifunción $x \to A(x)$ es creciente. Entonces la función g es una función (estrictamente) creciente.*

Demostración. Sean $x, y \in X$ tales que $x < y$, entonces para cada $a \in A(x)$ se tiene que $G(x,a) \leq G(y,a)$. Como $A(x) \subseteq A(y)$ se sigue que $G(x,a) \leq g(y)$, para cada $a \in A(x)$ y por lo tanto $g(x) \leq g(y)$.

Ahora, si las desigualdades anteriores se toman estrictas se concluye que g es estrictamente creciente. ■

Lema 3.1.3 *Supóngase que $r(\cdot, a)$ y $F(\cdot, a, s)$ son funciones (estrictamente) crecientes en X para cada $a \in A$ y $s \in S$, y si la multifunción $x \to A(x)$ es creciente. Entonces las funciones de iteración de valores son (estrictamente) crecientes.*

Demostración. La prueba se hará por inducción, para ello sean $x, y \in X$ con $x < y$, fijos. Como $r(\cdot, a)$ es una función creciente y $V_1(x) = \max_{a \in A(x)} r(x,a)$, entonces por el Lema 3.1.2 se concluye que $V_1(x) \leq V_1(y)$.

Supóngase que V_n es una función creciente para $n \geq 1$. Por definición

$$V_{n+1}(x) = \max_{a \in A(x)} \left\{ r(x,a) + \alpha E\left[V_n(F(x,a,\xi))\right] \right\},$$

$x \in X$, dado que $F(x,a,s) \leq F(y,a,s)$ para toda $a \in A$ y $s \in S$. Entonces por hipótesis de inducción se sabe que $V_n(F(x,a,s)) \leq V_n(F(y,a,s))$ para toda $a \in A$ y $s \in S$. Por la monotonía de la esperanza se concluye que $E\left[V_n(F(x,a,\xi))\right] \leq E\left[V_n(F(y,a,\xi))\right]$.

Como la suma de funciones crecientes es creciente, entonces $r(x,a) + \alpha E\left[V_n(F(x,a,\xi))\right]$ es una función creciente, aplicando el Lema 3.1.2 se tiene que V_{n+1} también lo es. Como x y y son arbitrarios se concluye que las funciones V_n son estrictamente crecientes.

Nótese que si las desigualdades involucradas se toman estrictas entonces las funciones V_n son estrictamente crecientes. ■

Lema 3.1.4 *Supóngase que r es estrictamente cóncava y $F(\cdot,s)$ es cóncava en $\mathbb{K}$ para cada $s \in S$ y ambas estrictamente crecientes en X. También, supóngase que la multifunción $x \to A(x)$ es creciente y convexa. Entonces las funciones de iteración de valores son estrictamente cóncavas. Además para cada $n = 1,2,...$, los maximizadores $f_n \in \mathbb{F}$ son únicos.*

Demostración. La prueba se hará por inducción. Como para cada $x \in X$, $V_1(x) = \max\limits_{a\in A(x)} r(x,a)$ y r es estrictamente cóncava en $\mathbb{K}$, por el Lema 3.1.1 se tiene que V_1 es estrictamente cóncava en X y f_1 es único.

Supóngase que para $n \geq 1$, V_n es estrictamente cóncava. Por definición

$$V_{n+1}(x) = \max_{a\in A(x)} \left\{r(x,a) + \alpha E\left[V_n(F(x,a,\xi))\right]\right\},$$

$x \in X$, dado que r es estrictamente cóncava, basta probar que $W(x,a) := E\left[V_n(F(x,a,\xi))\right]$ es una función estrictamente cóncava en $\mathbb{K}$. Para ello sean $(x,a),(y,\widetilde{a}) \in \mathbb{K}$ y $\lambda \in [0,1]$. Dado que $F(\cdot,\cdot,s)$ es cóncava en $\mathbb{K}$ entonces para cada $s \in S$,

$$F(\lambda(x,a) + (1-\lambda)(y,\widetilde{a}),s) \geq \lambda F(x,a,s) + (1-\lambda)F(y,\widetilde{a},s),$$

por el Lema 3.1.3 V_n es creciente, de lo cual se sigue que

$$V_n(F(\lambda(x,a) + (1-\lambda)(y,\widetilde{a}),s)) \geq V_n(\lambda F(x,a,s) + (1-\lambda)F(y,\widetilde{a},s)).$$

y como V_n es estrictamente cóncava y por la desigualdad anterior se obtiene que

$$V_n(F(\lambda(x,a) + (1-\lambda)(y,\widetilde{a}),s)) > \lambda V_n(F(x,a,s)) + (1-\lambda)V_n(F(x,\widetilde{a},s)).$$

Por la monotonía y la linealidad de la esperanza se tiene que

$$H(\lambda(x,a) + (1-\lambda)(y,\widetilde{a})) > \lambda H(x,a) + (1-\lambda)H(y,\widetilde{a}).$$

Finalmente aplicando el Lema 3.1.1 se concluye que V_{n+1} es una función estrictamente cóncava y que f_{n+1} es único. ■

El siguiente lema está basado en [8]. En dicho artículo se trabajó con costos y con dinámicas más generales. Sin embargo se presenta a continuación una versión con recompensas.

Lema 3.1.5 *Bajo las hipótesis del Lema 3.1.4 y bajo la suposición de que las Condiciones: 2.211, 2.2.12 y 2.2.13 son válidas, entonces la función de valor óptimo es estrictamente cóncava y la política óptima es única.*

Demostración. Se sabe por el Lema 3.1.4 que las funciones de iteración de valores V_n son estrictamente cóncavas. Por otro lado, bajo las Condiciones 2.2.11, 2.2.12 y 2.2.13 se cumple el Teorema de Programación Dinámica 2.2.17, por lo cual se tiene que V_n convergen de manera puntual a la función de valor óptimo V y se satisface la ecuación de programación dinámica, es decir, para cada $x \in X$

$$V(x) = \max_{a \in A(x)} \left\{ r(x,a) + \alpha E\left[V(F(x,a,\xi)) \right] \right\}.$$

Si $x, y \in X$ y $\lambda \in [0,1]$, entonces para cada $n \geq 1$ se tiene que

$$V_n(\lambda x + (1-\lambda)y) > \lambda V_n(x) + (1-\lambda)V_n(y),$$

si $n \to \infty$, se concluye que

$$V(\lambda x + (1-\lambda)y) \geq \lambda V(x) + (1-\lambda)V(y),$$

y por lo tanto V es cóncava. De manera análoga se muestra que V es creciente.

Por otra parte, como F y V son cóncavas y ambas crecientes, entonces la función $W(x,a) := E\left[V(F(x,a,s))\right]$ resulta ser cóncava. Finalmente como r es estrictamente cóncava y W es cóncava entonces la función $r + W$ es estrictamente cóncava, por el Lema 3.1.1 se concluye que la función de valor óptimo es estrictamente cóncava y que la política óptima $f \in \mathbb{F}$ es única. ■

3.2. Continuidad

Con relación a la continuidad de la función de valor óptimo en la literatura de PDMs es muy conocida (véase [6], [12] y [15]). En la conclusión del Teorema de Programación Dinámica 2.2.17 se garantiza que la función de valor óptimo $V \in \mathbb{B}_w(X)$. Sin embargo, bajo algunos cambios sobre las Condiciones 2.2.11, 2.2.12 y 2.2.13 se puede verificar que V pertenece a

$$\mathbb{L}_w(X) := \left\{ u : X \to \mathbb{R} |\ u \text{ es u.s.c y } u \in \mathbb{B}_w(X) \right\}.$$

Si la Condición 2.2.11 (c), conocida como continuidad fuerte, es remplazada por la Condición 3.2.1, llamada continuidad débil, es posible garantizar que $V \in \mathbb{L}_w(X)$.

Condición 3.2.1 *Para cada función u continua y acotada en X, la función $\overline{u}(x,a) := \int u(y)Q(dy|\, x,a)$ es continua en $\mathbb{K}$.*

Al considerar la continuidad débil junto con la Condición 2.2.11 (a) y (b), se tiene que las Condiciones 2.2.12 y 2.2.13 se deben reforzar. El cambio que sufren es el siguiente: la función de peso w debe ser una función continua en X y la función $\overline{w}(x,a) := \int w(y)Q(dy|\, x,a)$ continua en $\mathbb{K}$.

Recuerde que bajo las Condiciones 2.2.11, 2.2.12 y 2.2.13, la demostración del Teorema de Programación Dinámica (2.2.17) es la resultante del Lema 2.2.15 y de la Proposición 2.2.16, donde está ultima garantiza que el operador de programación dinámica T_α es una contracción en el espacio de Banach $\mathbb{B}_w(X)$. Con los cambios mencionados en las Condiciones 2.2.12 y 2.2.13, junto con los incisos (a) y (b) de la Condición 2.2.12 y tomando la Condición 3.2.1 en lugar del inciso (c), se mostrará que el Teorema de Programación Dinámica es válido, pero en este caso T_α resulta ser una contracción en $\mathbb{L}_w(X)$.

Lema 3.2.2 *Supóngase que*

a) para cada estado $x \in X$: $A(x)$ es un conjunto compacto, la función de recompensa $r(x,a)$ es u.s.c en a, $a \in A(x)$ y la Condición 3.2.1 se cumple;

b) existen constantes no negativas $\overline{c}$ y β, con $1 \leq \beta < 1/\alpha$, y una función de peso w continua sobre X, tales que para cada $x \in X$

i) $\sup_{A(x)} |r(x,a)| \leq \overline{c}w(x)$;

ii) $\sup_{A(x)} \int w(y)Q(dy|\, x,a) \leq \beta w(x)$;

c) para cada $x \in X$, la función $\overline{w}(x,a) := \int w(y)Q(dy|\, x,a)$ es continua en $\mathbb{K}$.

Entonces el operador de programación dinámica T_α es una contracción en $\mathbb{L}_w(X)$ con módulo $\gamma = \alpha\beta < 1$ y para cada $u \in \mathbb{L}_w(X)$ existe $f \in \mathbb{F}$ tal que para cada $x \in X$

$$T_\alpha u(x) = r(x, f(x)) + \alpha \int u(y)Q(dy|\, x, f(x)).$$

Además, si $\{v_n\} \subseteq \mathbb{L}_w(X)$ es una sucesión convergente en w-norma a V, entonces $V \in \mathbb{L}_w(X)$.

Demostración. Por la suposición (b) nótese que para cada $u \in \mathbb{L}_w(X)$ se tiene que $\left| r(x,a) + \alpha \int u(y)Q(\,dy|\,x,a)\right| \leq (\bar{c} + \alpha k)w(x)$, donde k representa una cota de u con la w-norma. Entonces si se muestra que para cada $u \in \mathbb{L}_w(X)$ la función

$$G(x,a) := r(x,a) + \alpha \int u(y)Q(\,dy|\,x,a),$$

es u.s.c en $A(x)$, por el Lema 2.2.14 se tendrá que $T_\alpha u \in \mathbb{L}_w(X)$ y que existe $f \in \mathbb{F}$, tal que para cada $x \in X$, $T_\alpha u(x) = G(x, f(x))$.

Así basta mostrar que para cada $u \in \mathbb{L}_w(X)$ la función la función $\overline{u}(x,a) := \int u(y)Q(\,dy|\,x,a)$ es continua en $\mathbb{K}$.

En efecto, sea $u \in \mathbb{L}_w(X)$ y considérese a la función $U := u - w\,\|u\|_w$. Como $U \leq 0$ y u.s.c, se sabe que existe una sucesión $\{u^n\}$ de funciones continuas y acotadas, tales que $u^n \downarrow U$ (véase Proposición 7.1.2 del Apéndice). Ahora, sea $\{(x_k, a_k)\} \subset \mathbb{K}$ una sucesión convergente a $(x,a) \in \mathbb{K}$. Entonces como $u^n \downarrow U$, se sigue que para toda n,

$$\limsup_{k\to\infty} \int U(y)Q(\,dy|\,x^k, a^k) \leq \limsup_{k\to\infty} \int u^n(y)Q(\,dy|\,x^k, a^k).$$

Por la Condición 3.2.1 se sabe que

$$\limsup_{k\to\infty} \int u^n(y)Q(\,dy|\,x^k, a^k) = \int u^n(y)Q(\,dy|\,x,a),$$

implicando que

$$\limsup_{k\to\infty} \int U(y)Q(\,dy|\,x^k, a^k) \leq \int u^n(y)Q(\,dy|\,x,a),$$

si $n \to \infty$ y aplicando el Teorema de la convergencia monótona se obtiene que

$$\limsup_{k\to\infty} \int U(y)Q(\,dy|\,x^k, a^k) \leq \int U(y)Q(\,dy|\,x,a).$$

De manera similar a la demostración del Lema 2.2.15 se concluye que

$$\limsup_{k\to\infty} \int u(y)Q(\,dy|\,x^k, a^k) \leq \int u(y)Q(\,dy|\,x,a)$$

es decir, $\overline{u}$ es u.s.c en $\mathbb{K}$. Aplicando a $-u$ en la última desigualdad se concluye que $\overline{u}$ es l.s.c en $\mathbb{K}$ y por tanto $\overline{u}$ es una función continua en $\mathbb{K}$.

Por último, si $\{v_n\} \subset \mathbb{L}_w(X)$ es una sucesión tal que converge en w-norma a V, entonces se sabe que $V \in \mathbb{B}_w(X)$ (dado que $\mathbb{L}_w(X) \subset \mathbb{B}_w(X)$). Obsérvese que para $x \in X$

$$V(x) = [V(x) - v_n(x)] + v_n(x) \leq \|V - v_n\|_w\, w(x) + v_n(x).$$

Sean $x \in X$ fijo y $\{x^k\} \subset X$ una sucesión convergente a x. Debido a que las v_n son u.s.c y w es continua se tiene que

$$\limsup_{k\to\infty} V(x^k) \leq \|V - v_n\|_w \, w(x) + v_n(x).$$

Finalmente si $n \to \infty$ se concluye que $\limsup\limits_{k\to\infty} V(x^k) \leq V(x)$ y por lo tanto V es u.s.c, es decir, $V \in \mathbb{L}_w(X)$. ■

El siguiente lema es más restrictivo que el anterior, debido a que el Lema 2.2.14 también deberá modificarse en el sentido: que la multifunción $x \to A(x)$ sea continua, si u es una función continua y acotada en X y bajo el supuesto de que para cada $x \in X$, $A(x)$ es un conjunto compacto entonces, la función U definida por

$$U(x) = \sup_{a\in A(x)} u(x,a).$$

es continua y acotada en X.

Lema 3.2.3 *Supóngase que*

a) para cada $x \in X$, $A(x)$ es un conjunto compacto;

b) la función de recompensa r es continua y acotada en $\mathbb{K}$;

c) para cada función $u : X \to \mathbb{R}$ continua y acotada, la función $H(x,a) := \int u(y)Q(dy|\, x,a)$, es continua en $\mathbb{K}$.

Entonces, la función de valor óptimo V es continua en X.

Demostración. La demostración es similar a la dada en Lema 3.2.2, basándose en el hecho de que el operador de programación dinámica T_α es una contracción en el espacio de Banach $\mathbb{C}(X)$, donde

$$\mathbb{C}(X) := \{\, u : X \to \mathbb{R} |\; u \text{ es continua y acotada}\}\,,$$

con la noma $\|u\| = \sup\limits_{x\in X} |u(x)|$. ■

Observación 3.2.4 *En el Lema anterior, si se supone que el conjunto de acciones A es compacto. Entonces, es posible mostrar que V es una función continua, bajo ciertas condiciones de continuidad en el modelo.*

También es posible mostrar, bajo el supuesto que X es un conjunto abierto y bajo las hipótesis del Lema 3.1.5, que la función de valor óptimo V es continua (véase Proposición 7.2.1 del Apéndice).

3.3. Diferenciabilidad

Se considera a lo largo de esta sección que la multifunción $x \to A(x)$, es creciente y convexa. También se suponen que se cumplen las Condiciones 2.2.11, 2.2.12 y 2.2.13, dichas condiciones son supuestas para garantizar la validez del Teorema de Programación Dinámica 2.2.17.

Considere la siguiente notación: $C^2(X,Y)$ denota al conjunto de funciones $G : X \to Y$ con derivada de segundo orden continua. G_x, G_a y G_{aa} denotan la derivada de G con respecto x y a, y la segunda derivada con respecto a, respectivamente. El interior del conjunto X se denota por $int(X)$.

Teorema 3.3.1 *(**Envolvente**) Supóngase que $G \in C^2(int(\mathbb{K});\mathbb{R})$ y $G_{aa}(x,\cdot)$ es definida negativa, para cada $x \in X$. Si existe una función $f \in \mathbb{F}$, tal que para cada $x \in X$, $f(x) \in int(A(x))$ y $g(x) = G(x, f(x))$. Entonces se tiene que $f \in C^1(int(X); A)$, $g \in C^2(int(X);\mathbb{R})$ y para $x \in int(X)$*

$$g'(x) = G_x(x, f(x)).$$

Demostración. Sea $x \in int(X)$. Como el maximizador f cumple que $f(x) \in int(A(x))$, por la condición de primer orden se sigue que $G_a(x, f(x)) = 0$.

Debido a que $G_{aa}(x,\cdot)$ es definida negativa, entonces G es estrictamente cóncava (véase [12]) y por lo tanto f es única. Aplicando el teorema de la función implícita (véase [12]), se llega a que $f \in C^1(int(X); A)$.

Por otra parte, como $g(x) = G(x, f(x))$,

$$g'(x) = G_x(x, f(x)) + G_a(x, f(x))f'(x),$$

pero $G_a(x, f(x)) = 0$, lo cual implica que $g'(x) = G_x(x, f(x))$.

Además

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

y como $G \in C^2(int(\mathbb{K});\mathbb{R})$ y $f \in C^1(int(X); A)$, se concluye que $g \in C^2(int(X);\mathbb{R})$.
■

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

$$G^n(x, a) := r(x, a) + \alpha E\left[V_{n-1}\left(F(x, a, \xi)\right)\right].$$

Teorema 3.3.2 *Supóngase que:*

a) $r \in C^2(int(\mathbb{K});\mathbb{R})$, $F(\cdot, s) \in C^2(int(\mathbb{K}); X)$ para cada $s \in S$, ambas crecientes en X, $r_{aa}(x,\cdot)$ es definida negativa y $F_{aa}(x,\cdot)$ semidefinida negativa;

b) $f_n(x) \in int((A(x))$ *para cada* $x \in X$ *y* $n = 1, 2, ...;$

Entonces para cada $n = 0, 1, 2, ...,$ $f_n \in C^1(int(X); A)$, $V_n \in C^2(int(X); \mathbb{R})$ *y para cada* $x \in int(X)$

$$V_n'(x) = G_x^n(x, f_n(x)),$$

donde f_n *son los maximizadores del algoritmo de Programación Dinámica.*

Demostración. La prueba se hará por inducción, para ello sea $x \in int(X)$. Como $V_1(x) = \max_{a \in A(x)} r(x, a)$ y por hipótesis se sabe que $f_1(x) \in int(A(x))$ y que r es de clase C^2 con $r_{aa}(x, a)$ definida negativa, entonces por el Teorema de la envolvente (3.3.1) se tiene que $f_1 \in C^1(int(X); A)$, $V_1 \in C^2(int(X); \mathbb{R})$ y

$$V_1'(x) = r_x(x, f_1(x)) = G_x^1(x, f_1(x)).$$

Además como r es definida negativa, se tiene que es estrictamente cóncava (véase [12]) y por el Lema 3.1.5 se concluye que V_1 es estrictamente cóncava y por ser de clase C^2, resulta ser semidefinida negativa (véase [12]).

Para $n = 2$ se tiene que

$$V_2(x) = \max_{a \in A(x)} G^2(x, a).$$

Como r, F y V_1 son de clase C^2 entonces la función G^2 también lo es. Además dado que V_1 es estrictamente cóncava y de clase C^2 entonces la función $W^2(x, a) := E\left[V_1\left(F(x, a, \xi)\right)\right]$ resulta ser una función cóncava y de clase C^2 y por tanto $W_{aa}^2(x, a)$ es semidefinida negativa.

Dado que la suma de una función definida negativa con una semidefinida negativa es definida negativa y como

$$G_{aa}^2(x, a) = r_{aa}(x, a) + \alpha W_{aa}^2(x, a),$$

$a \in A(x)$, entonces G_{aa}^2 es definida negativa. Debido a que $f_2(x) \in int(A(x))$, por el Teorema de la envolvente 3.3.1 se tiene que $f_2 \in C^1(int(X); A)$, $V_2 \in C^2(int(X); \mathbb{R})$.

Supóngase que $V_{n-1} \in C^2(int(X); \mathbb{R})$ con $n > 2$. Como r, F y V_{n-1} son de clase C^2 entonces la función G^n también lo es. Por el Lema 3.1.5 se sabe que V_{n-1} es estrictamente cóncava, entonces la función $W^n(x, a) := E\left[V_{n-1}\left(F(x, a, \xi)\right)\right]$ resulta ser una función cóncava y de clase C^2 y por lo cual $W_{aa}^n(x, a)$ es semidefinida negativa.

Nuevamente dado que la suma de una función definida negativa con una semidefinida negativa es definida negativa y como

$$G_{aa}^n(x, a) = r_{aa}(x, a) + \alpha W_{aa}^n(x, a),$$

$a \in A(x)$, entonces $G^n_{aa}(x,a)$ es definida negativa, con un razonamiento similar se concluye que $f_n \in C^1(int(X);A)$, $V_n \in C^2(int(X);\mathbb{R})$. ■

Sea $G : \mathbb{K} \to \mathbb{R}$ definida para cada $(x,a) \in \mathbb{K}$

$$G(x,a) := r(x,a) + \alpha \int V(y)Q(\,dy|\,x,a).$$

Lema 3.3.3 *Supóngase que*

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

b) para cada $s \in S$, $F(\cdot,\cdot,s) \in C^2(int(\mathbb{K});X)$*. Además* F *tiene inversa en la tercera variable, denotada por* $R : \mathbb{K}\times X \to S$*, tal que* $R(\cdot,\cdot,s) \in C^2(int(\mathbb{K}\times X);S)$ *y para toda* $s \in S$ $|\det R_s(\cdot,\cdot,s)| \in C^2(int(\mathbb{K});\mathbb{R})$*, donde* $\det R_s$ *denota el determinante de la matriz* R_s*;*

c) S *es un conjunto abierto y* $\Delta \in C^2(S;\mathbb{R})$*;*

Entonces, $G \in C^2(int(\mathbb{K});\mathbb{R})$*.*

Demostración. Nótese primero que para cada subconjunto medible B de X y cada $(x,a) \in \mathbb{K}$,

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

Entonces, por el Teorema de Cambio de Variable (véase [1]), se tiene que

$$Q(\,B|\,x,a) = \int \Delta(R(x,a,u))\,|\det R_s(x,a,u)|\,du,$$

por lo que G se expresa como

$$G(x,a) = r(x,a) + \alpha \int V(u)\Delta(R(x,a,u))\,|\det R_s(x,a,u)|\,du,$$

y el resultado se sigue al derivar a G. ■

Observación 3.3.4 *En la conclusión del Teorema anterior es necesario realizar el intercambio entre derivada e integral. Para garantizar la diferenciabilidad de segundo orden para* $\int \hat{U}(x,a,u)du$*, con respecto a* x *y* a*, donde*

$$\hat{U}(x,a,u) = V(u)\Delta(R(x,a,u))\left|\det R_{(3)}(x,a,u)\right|,$$

$(x, a, u) \in int(\mathbb{K}\times X)$. *En la práctica esta condición puede verificarse cuando las derivadas de* $\hat{U}$ *están acotadas (véase [1]), es decir, para cada* $(x, a, u) \in int(\mathbb{K}\times X)$, $\left|\hat{U}_x(x, a, u)\right| \leq \widehat{g}(u, a)$, $\left|\hat{U}_a(x, a, u)\right| \leq \widehat{h}(u, x)$, $\left|\hat{U}_{xx}(x, a, u)\right| \leq \widehat{l}(u, a)$, $\left|\hat{U}_{aa}(x, a, u)\right| \leq \widehat{m}(u, a)$ *y* $\left|\hat{U}_{xa}(x, a, u)\right| \leq \widehat{n}(u, a)$, *para algunas funciones* $\widehat{g}$, $\widehat{h}$, $\widehat{l}$, $\widehat{m}$ *y* $\widehat{n}$ *integrables con respecto a* μ.

Teorema 3.3.5 *Supóngase que para cada* $x \in X$*:* $G_{aa}(x, \cdot)$ *es definida negativa y* $f(x) \in int(A(x))$. *Entonces* $f \in C^1(int(X); A)$ *y* $V \in C^2(int(X); \mathbb{R})$.

Demostración. Esto es una consecuencia directa del Teorema de la envolvente 3.3.1. ■

Capítulo 4

LA ECUACIÓN DE EULER

En este capítulo se presenta una versión de la Ecuación de Euler en el contexto de PDMs, la cual permite calcular las derivadas de las funciones de iteración de valores de manera iterativa. A diferencia de Programación Dinámica resulta ser más fácil de implementar debido a que no se involucran de manera directa el cálculo de los maximizadores f_n de las funciones de iteración de valores.

4.1. La Ecuación de Euler

En el desarrollo subsecuente de esta sección se considera que $X \subseteq \mathbb{R}^n$, $A \subseteq \mathbb{R}^m$ y para cada $x \in X$, $A(x)$ son conjuntos convexos. También se considera que la multifunción $x \to A(x)$, es creciente y convexa. Además, suponga que la ley de transición es inducida por la siguiente ecuación en diferencias

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

donde, $\{\xi_t\}$ es una sucesión de v.a.'s i.i.d. tomando valores en un espacio $S \subset \mathbb{R}^k$ con densidad común Δ e independiente del estado inicial $x_0 = x \in X$; se supone que $L : X' \times S \to X$ es una función medible conocida, con $X' \subset \mathbb{R}^m$ y $F : \mathbb{K} \to X'$ es una función medible conocida.

Entonces las funciones de iteración de valores son de la forma

$$V_n(x) = \max_{a \in A(x)} \left\{ r(x,a) + \alpha E\left[V_{n-1}(L\left(F(x,a), \xi\right))\right] \right\},$$

con $V_0(x) = 0$, $x \in X$ y $n \geq 1$.

Además supóngase que para cada $n = 0, 1, ...$, existe $f_n \in \mathbb{F}$ tal que para cada $x \in X$

$$V_n(x) = r(x, f_n(x)) + \alpha E\left[V_{n-1}(L\left(F(x, f_n(x)), \xi\right))\right].$$

Se define para $(x, a) \in \mathbb{K}$ y $n = 1, 2, ...,$ la función

$$G^n(x, a) := r(x, a) + \alpha E\left[V_{n-1}(L\left(F(x, a), \xi\right))\right].$$

Teorema 4.1.1 *Supóngase que*

a) $r \in C^2(int(\mathbb{K}); \mathbb{R})$, creciente en X y $r_{aa}(x, \cdot)$ definida negativa. $L(\cdot, s) \in C^2(int(X' \times S); X)$, creciente en X' para cada $s \in S$ y $L_{uu}(\cdot, s)$ definida negativa. $F \in C^2(int(\mathbb{K}); X')$, creciente en X y $F_{aa}(\cdot)$ semidefinida negativa, además $F_a(x, a)$ invertible;

b) $f_n(x) \in int((A(x))$ para cada $x \in X$ y $n = 0, 1, ...$;

c)La función definida para $(x, a) \in \mathbb{K}$ como

$$H(x, a) := \left[r_x - r_a F_a^{-1} F_x\right](x, a),$$

es invertible en la segunda variable, la cual será denotada por h.

Entonces, para $n = 1, 2, ...,$ $f_n \in C^1(int(X); A)$, $V_n \in C^2(int(X); \mathbb{R})$. Además V_n satisface la Ecuación de Euler dada por

$$(r_a F_a^{-1})(x, h(x, V_n'(x))) = \begin{array}{c} -\alpha E[V_{n-1}'(L(F(x, h(x, V_n'(x))), \xi)) \\ L_u(F(x, h(x, V_n'(x))), \xi))] \end{array}, \tag{4.1}$$

para cada $x \in int(X)$, donde $u = F(x, a)$ para $a \in int(A(x))$.

Demostración. Como $F(\cdot, a)$ y $L(\cdot, s)$ son crecientes en X y X', respectivamente, la función $M(x, a) := L(F(x, a), s)$ es creciente en X. Además M es de clase C^2 en $\mathbb{K}$, ya que L y F lo son. También dado que L_{uu} y F_{aa} son semidefinidas negativas, L y F resultan ser cóncavas y por lo tanto M también lo es. En resumen M es de clase C^2 y cóncava, implicando que M_{aa} es semidefinida negativa.

Por lo anterior se tiene que r y G satisfacen las hipótesis del Teorema 3.3.2 y por lo tanto, se tiene que para cada $n = 1, 2, ...,$ $V_n \in C^2(int(X); \mathbb{R})$, $f_n \in C^1(int(X); A)$ y

$$V_n'(x) = G_x^n(x, f_n(x)). \tag{4.2}$$

Por otro lado, sea $x \in int(X)$, como

$$V_1(x) = \max_{a \in A(x)} \left\{r(x, a)\right\},$$

teniendo en cuenta que $f_1(x) \in int((A(x))$, por la condición de primer orden se tiene que $G_a^1(x, f_1(x)) = r_a(x, f_1(x)) = 0$ y por lo tanto V_1 satisface la ecuación (4.1).

Ahora para $n > 1$, como en la demostración del Teorema 3.2.2 se sabe que $G^n \in C^2(int(\mathbb{K}); \mathbb{R})$ y dado que $f_n(x) \in int((A(x))$, por la condición de primer orden se sigue que $G_a^n(x, f_n(x)) = 0$. Debido a que $F_a(x, a)$ es invertible se tiene que

$$-r_a F_a^{-1}(x, f_n(x)) = \alpha E\left[V'_{n-1}\left(L\left(F(x, f_n(x)), s\right)\right) L_u(F(x, f_n(x)), s)\right], \quad (4.3)$$

donde $u = F(x, a)$.

Dado que $V_n(x) = G^n(x, f_n(x))$ y por (4.2), se obtiene que

$$V'_n(x) = r_x(x, f_n(x)) + \alpha E\left[V'_{n-1}\left(L\left(F(x, f_n(x)), s\right)\right) L_u(F(x, f_n(x)), s)\right] F_x(x, f_n(x)).$$

Al sustituir (4.3) en esta ultima igualdad se llega a la relación siguiente

$$\begin{aligned} V'_n(x) &= \left[r_x - r_a F_a^{-1} F_x\right](x, f_n(x)) \\ &= H(x, f_n(x)). \end{aligned}$$

por la invertibilidad de H se concluye que

$$f_n(x) = h(x, V'_n(x)). \quad (4.4)$$

Finalmente al sustituir (4.4) en (4.3) se tiene que (4.1) es válida. ■

4.2. Un Ejemplo Determinista

Considere un individuo que tiene un capital inicial x y desea planificar su consumo en los siguientes $N+1$ periodos. En cada periodo $t = 0, 1, ..., N$ debe decidir qué parte a_t del capital disponible dedicar al consumo. Sus decisiones le producen una utilidad que depende de a_t y su objetivo es diseñar un plan que optimice la utilidad total. Supóngase que la utilidad está dada por

$$u(a) = \gamma \log(a),$$

donde $\gamma > 0$, y en este contexto se interpreta como un incremento del consumo.

En este caso, el espacio de estado es igual al de acciones y es $X = A = [0, \infty)$. Como su acción es la cantidad que dedicará al consumo en cada periodo, supóngase que dicho monto no puede superar a x, y por lo tanto $A(x) = [0, x]$. La función de recompensa está dada por $r(x, a) = \gamma \log(a)$. El problema descrito se analizará en los casos siguientes: cuando el individuo tiene empleo y cuando no lo tiene.

4.2.1. Caso sin Empleo

En el caso del que individuo no tenga empleo la dinámica del proceso está dada por

$$x_{t+1} = x_t - a_t,$$

lo que significa que su capital en el periodo $t+1$ es el capital que tenía en el tiempo t menos lo asignado al consumo en ese tiempo. Finalmente dado $x \geq 0$ se desea maximizar la utilidad total

$$\sum_{t=0}^{N} \gamma \log(a_t).$$

donde la recompensa terminal es la función constante 0.

Lema 4.2.1 *Para $n \geq 1$ las funciones de iteración de valores están dadas por*

$$V_n(x) = \max_{a \in [0,x]} \left\{ \gamma \log(a) + V_{n-1}(x-a) \right\},$$

y se satisface que para cada $x \in (0, \infty)$

$$V_n'(x) = \frac{n\gamma}{x}.$$

Además $f_n(x) \in (0, x)$ y

$$f_n(x) = \frac{x}{n}.$$

Demostración. La prueba se hará por inducción. Dado que $V_0(x) = 0$ y debido a que la función logaritmo es creciente, entonces $V_1(x) = \gamma \log(x)$, por lo tanto, si $x \in (0, \infty)$ se sigue que

$$V_1'(x) = \frac{\gamma}{x},$$

y $f_1(x) = x$. Además $V_1''(x) = -\frac{\gamma}{x^2} < 0$ y es continua en $x \in (0, \infty)$. Así para $x \in (0, \infty) : V_1'(x) > 0$, $V_1''(x) < 0$ y $\lim_{x \to 0} V_1'(x) = \infty$.

Supóngase que para $n > 1$ y para cada $x \in (0, \infty)$,

$$V_{n-1}'(x) = \frac{(n-1)\gamma}{x}.$$

En este caso obsérvese que $V_{n-1} \in C^2((0, \infty))$ y $\lim_{x \to 0} V_{n-1}'(x) = \infty$.

Sea $x > 0$ y defina a $G^n(x,a) := \gamma \log(a) + V_{n-1}(x-a)$. Si $a \in (0,x)$, entonces

$$G^n_a(x,a) = \frac{\gamma}{a} - V'_{n-1}(x-a).$$

Además se tiene que $\lim_{a\to 0} G^n_a(x,a) = +\infty$ y $\lim_{a\to x} G^n_a(x,a) = \frac{\gamma}{x} - \lim_{a\to x} V'_{n-1}(x-a)$. Como

$$\lim_{a\to x} V'_{n-1}(x-a) = \lim_{y\to 0} V'_{n-1}(y) = \infty.$$

Entonces $\lim_{a\to x} G^n_a(x,a) = -\infty$. Por el teorema del valor intermedio se obtiene que existe $\overline{a} := f_n(x) \in (0,x)$ tal que $G^n_a(x, f_n(x)) = 0$. Además este punto es único dado que $G^n(x,\cdot)$ es estrictamente cóncava en a. Así

$$\frac{\gamma}{f_n(x)} = V'_{n-1}(x - f_n(x)). \tag{4.5}$$

Nótese que

$$G^n_{aa}(x,a) = -\frac{\gamma}{a^2} - \frac{(n-1)\gamma}{(x-a)^2},$$

para $a \in (0,x)$. De lo anterior se sigue que G^n es de clase C^2, $G^n_{aa}(x,a) < 0$ y $f_n(x) \in (0,x)$. Por el Teorema de la envolvente 3.3.1 se tiene que $V'_n(x) = G^n_x(x, f_n(x))$, es decir,

$$V'_n(x) = V'_{n-1}(x - f_n(x)). \tag{4.6}$$

Como $V'_n(x) > 0$, por (4.5) y (4.6) se sigue que $V'_n(x) = \frac{\gamma}{f_n(x)}$, equivalentemente,

$$f_n(x) = \frac{\gamma}{V'_n(x)}, \tag{4.7}$$

sustituyendo (4.7) en (4.6), se tiene que V_n satisface la ecuación de Euler

$$V'_n(x) = V'_{n-1}\left(x - \frac{\gamma}{V'_n(x)}\right). \tag{4.8}$$

Por hipótesis, $V'_{n-1}(x) = \frac{(n-1)\gamma}{x}$, así la ecuación (4.8) toma la forma

$$V'_n(x) = \frac{(n-1)\gamma}{x - \frac{\gamma}{V'_n(x)}},$$

equivalentemente,

$$V'_n(x) = \frac{n\gamma}{x}. \tag{4.9}$$

Finalmente sustituyendo (4.9) en (4.7) se llega a que

$$f_n(x) = \frac{x}{n}.$$

■

Corolario 4.2.2 *Para $x > 0$ y $n \geq 1$,*

$$V_n(x) = n\gamma \log\left(\frac{x}{n}\right).$$

Demostración. Por el lema anterior se sabe que para $n \geq 1$ y $x > 0$, $V_n'(x) = \frac{n\gamma}{x}$ por consiguiente

$$V_n(x) = n\gamma \log(x) + c_n.$$

Además se sabe que el maximizador de V_n es $f_n(x) = \frac{x}{n}$, y ya que $V_n(x) = \max_{a \in [0,x]} \{\gamma \log(a) + V_{n-1}(x-a)\}$, se tiene que

$$n\gamma \log(x) + c_n = \gamma \log\left(\frac{x}{n}\right) + (n-1)\gamma \log\left(x - \frac{x}{n}\right) + c_{n-1},$$

simplificando la relación anterior se llega a que

$$c_n = \gamma\left[(n-1)\log(n-1) - n\log(n)\right] + c_{n-1},$$

dado que $V_1(x) = \gamma \log(x)$ se obtiene que $c_1 = 0$, al iterar la relación anterior se tiene que para $n \geq 1$, $c_n = -n\gamma \log(n)$. Por lo tanto $V_n(x) = n\gamma \log\left(\frac{x}{n}\right)$. ■

Corolario 4.2.3 *La función de valor óptimo está definida de la forma siguiente: para $x \in (0, \infty)$*

$$V_N(x) = N\gamma \log\left(\frac{x}{N}\right),$$

y $V_N(0) = -\infty$. Además la política óptima está determinada por

$$f_n(x) = \frac{x}{n}.$$

para $n = 1, 2, ..., N$, y $f_0(x) = 0$.

Demostración. Por el Teorema de Programación Dinámica (2.2.1) se sabe que la función de valor óptimo se alcanza en la última etapa y la política óptima es la sucesión de maximizadores que se obtienen. Así utilizando el Corolario 4.2.2 el resultado se sigue. ■

Se interpreta la decisión en el tiempo n de la forma siguiente, se asigna la cantidad $\frac{x}{n}$ al consumo en el tiempo n.

4.2.2. Caso con Empleo

Ahora supóngase que el individuo tiene un ingreso adicional fijo, por ejemplo que trabaja y gana una cantidad $w > 0$. Entonces la dinámica del sistema está dada por

$$x_{t+1} = x_t - a_t + w.$$

Nuevamente se desea maximizar la utilidad total, donde la utilidad está dada como en caso sin empleo.

En este caso se tienen las siguientes posibilidades, la primera es cuando la riqueza x es mayor o igual a lo que gana, es decir, $x > w$, en el cual se verá que los maximizadores son interiores. Para el segundo caso $x \leq w$, el maximizador se alcanzará en un extremo.

Lema 4.2.4 *Supóngase que $x > w$. Entonces para $n \geq 1$, las funciones de iteración de valores están dadas por*

$$V_n(x) = \max_{a \in [0,x]} \left\{ \gamma \log(a) + V_{n-1}(x - a + w) \right\},$$

son diferenciables y

$$V_n'(x) = \frac{n\gamma}{x + (n-1)w}.$$

Además $f_n(x) \in (0, x)$ y

$$f_n(x) = \frac{x + (n-1)w}{n}.$$

Si $x \leq w$ entonces $f_n(x) = x$.

Demostración. Supóngase que $x > w$. Como $V_0(x) = 0$ y la función logaritmo es creciente, se sigue que $V_1(x) = \gamma \log(x)$ y

$$V_1'(x) = \frac{\gamma}{x}.$$

Además, $V_1''(x) = -\frac{\gamma}{x^2} < 0$, y es continua para en $x \in (0, \infty)$. Por lo que se concluye que para $x \in (0, \infty)$: $V_1'(x) > 0$, $V_1''(x) < 0$ y $\lim_{x \to 0} V_1'(x) = \infty$.

Supóngase que para $n > 1$

$$V_{n-1}'(x) = \frac{(n-1)\gamma}{x + (n-2)w},$$

y nótese que $V_{n-1}'(x) > 0$ y $V_{n-1}''(x) < 0$.

Sea $G^n(x,a) = \gamma \log(a) + V_{n-1}(x-a+w)$. Si $a \in (0,x)$ entonces

$$G^n_a(x,a) = \frac{\gamma}{a} - V'_{n-1}(x-a+w).$$

Además, se tiene que $\lim_{a\to 0} G^n_a(x,a) = \infty$ y $\lim_{a\to x-} G^n_a(x,a) = \frac{\gamma}{x} - V'_{n-1}(w) = \frac{\gamma}{x} - \frac{\gamma}{w} < 0$, dado que $x > w$.

Por el teorema del valor intermedio se tiene que existe $\overline{a} := f_n(x) \in (0,x)$ tal que $G^n_a(x, f_n(x)) = 0$. Además este punto es único dado que $G^n(x,\cdot)$ es estrictamente cóncava en a. Así

$$\frac{\gamma}{f_n(x)} = V'_{n-1}(x - f_n(x) + w). \tag{4.10}$$

Nótese que

$$G^n_{aa}(x,a) = -\frac{\gamma}{a^2} - \frac{(n-1)\gamma}{(x-a+(n-2)w)^2},$$

$a \in (0,x)$. De lo anterior se sigue que G^n_{aa} es de clase C^2, $G^n_{aa}(x,a) < 0$ y $f_n(x) \in (0,x)$. Por el Teorema de le envolvente 3.3.1 se tiene que $V'_n(x) = G^n_x(x, f_n(x))$, es decir,

$$V'_n(x) = V'_{n-1}(x - f_n(x) + w). \tag{4.11}$$

y nótese que $V'_n(x) > 0$. Por (4.10) y (4.11) se sigue que $V'_n(x) = \frac{\gamma}{f_n(x)}$, equivalentemente

$$f_n(x) = \frac{\gamma}{V'_n(x)}. \tag{4.12}$$

Sustituyendo (4.12) en (4.11), se tiene que V_n satisface la ecuación de Euler

$$V'_n(x) = V'_{n-1}\left(x - \frac{\gamma}{V'_n(x)} + w\right). \tag{4.13}$$

Por hipótesis, $V'_{n-1}(x) = \frac{(n-1)\gamma}{x+(n-2)w}$, así la ecuación (4.13) toma la forma

$$V'_n(x) = \frac{(n-1)\gamma}{x - \frac{\gamma}{V'_n(x)} + (n-1)w},$$

equivalentemente,

$$V'_n(x) = \frac{n\gamma}{x+(n-1)w}. \tag{4.14}$$

Finalmente sustituyendo (4.14) en (4.12) se tiene que

$$f_n(x) = \frac{x+(n-1)w}{n}.$$

En el caso de que $x \leq w$ se tiene que $x < \frac{x+(n-1)w}{n}$, lo cual implica que el maximizador obtenido en el caso anterior no es admisible, es decir, no está en $A(x) = [0, x]$. En este caso, dado que $G_a^n > 0$, $n \geq 1$, entonces G^n es una función creciente para $n \geq 1$ y por lo tanto alcanza su máximo en x, es decir $f_n(x) = x$ en $A(x)$, para $n \geq 1$. ■

Corolario 4.2.5 *Para $x > w$ se tiene que*

$$V_n(x) = n\gamma \log\left(\frac{x + (n-1)w}{n}\right),$$

y para $x \leq w$

$$V_n(x) = \gamma \log(x) + \gamma \log((n-1)w).$$

Demostración. Si $x > w$, por el lema anterior sabemos que $V_n'(x) = \frac{n\gamma}{x+(n-1)w}$ entonces

$$V_n(x) = n\gamma \log(x + (n-1)w) + c_n,$$

donde c_n se encuentra de manera similar al caso sin empleo (véase Corolario 4.2.2). En este caso se obtiene que $c_n = n\gamma \log(\frac{1}{n})$ y el resultado se sigue.

Por otro lado si $x \leq w$, para toda $n \geq 1$, $f_n(x) = x$, así al sustituir en las funciones de iteración de valores se obtiene que

$$V_n(x) = \gamma \log(x) + \gamma \log((n-1)w).$$

■

En resumen se tiene el resultado siguiente.

Corolario 4.2.6 *La función de valor óptimo está dada por*

$$V_N(x) = \begin{cases} N\gamma \log\left(\frac{x+(N-1)w}{N}\right) & si\ x > w \\ \gamma \log(x) + \gamma \log((N-1)w) & si\ x \leq w \end{cases}$$

y la política óptima está determinada por

$$f_n(x) = \begin{cases} \frac{x+(n-1)w}{n} & si\ x > w \\ x & si\ x \leq w \end{cases}$$

para $n = 1, 2, ..., N$ con $f_0(x) = 0$.

Se concluye que la decisión en el tiempo n es de la forma siguiente:

a) Si $x \leq w$ entonces se decide gastar toda la riqueza;

b) si $x > w$ entonces se decide gastar solo una parte de ella, para ser más específico, $\frac{x+(n-1)w}{n}$.

4.3. Un ejemplo Estocástico

Supóngase que el espacio de estados es $X = [0, \infty)$, el de acciones es $A = [0, \pi]$, y para cada $x \in X$, $A(x) = A$. También supóngase que la dinámica del sistema está dada por

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

donde $w > 0$ y $\{\xi_t\}$ es una sucesión de v.a.'s i.i.d., tal que para toda t, ξ_t tiene soporte finito (por ejemplo, Bernoulli, Binomial, por mencionar algunas). Además considérese que la función de recompensa está dada por

$$r(x, a) = sen(a),$$

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

En este caso se desea maximizar la recompensa total esperada en $N + 1$ periodos, es decir,

$$E_x^\pi \left[\sum_{t=0}^{N} \alpha^t sen(a_t) \right].$$

Lema 4.3.1 *Para $n \geq 1$, las funciones de iteración de valores dadas por*

$$V_n(x) = \max_{a \in [0,\pi]} \{sen(a) + \alpha E[V_{n-1}(x + a + w\xi)]\},$$

son diferenciables y satisfacen

$$V_n'(x) = 0.$$

Demostración. Sea $x \in X$ fijo. Dado que $V_1(x) = \max_{a \in [0,x]} \{sen(a)\}$, entonces $f_1(x) = \frac{\pi}{2} \in (0, \pi) = int(A(x))$ maximiza la función entre llaves y por consiguiente $V_1(x) = 1$, lo cual implica que $V_1'(x) = 0$. Además noté que $|V_1(x)| \leq 1$.

Bajo la hipótesis inductiva de que V_{n-1} es de clase C^2 y $|V_n(x)| \leq 1$. Entonces la función $G^n(x, a) = sen(a) + \alpha E[V_{n-1}(x + a + w\xi)]$, es diferenciable, en esta caso el intercambio entre derivada e integral es válido ya que el soporte de la v.a es finito. Así, para $a \in (0, \pi)$ se tiene que

$$G_a^n(x, a) = \cos(a) + \alpha E\left[V_{n-1}'(x + a + w\xi)\right],$$

es más, si $\overline{a} \in (0, \pi)$ es maximiza a $G^n(x, \cdot)$, entonces por la condición de primer orden se llega a

$$-\cos(\overline{a}) = \alpha E\left[V_{n-1}'(x + \overline{a} + w\xi)\right], \tag{4.15}$$

para esta misma $\overline{a}$, por el Teorema de la envolvente se sabe que $V_n'(x) = G_x^n(x, \overline{a})$, es decir,

$$V_n'(x) = \alpha E\left[V_{n-1}'(x + \overline{a} + w\xi)\right], \tag{4.16}$$

por la igualdad (4.15) y por la hipótesis $|V_n(x)| \leq 1$, se obtiene que

$$\overline{a} = \arccos(-V_n'(x)). \tag{4.17}$$

Así sustituyendo (4.17) en (4.16) se tiene la siguiente ecuación de Euler,

$$V_n'(x) = \alpha E\left[V_{n-1}'(x + \arccos(-V_n'(x)) + w\xi)\right]. \tag{4.18}$$

Finalmente dado que $V_1'(x) = 0$, al ir iterando a (4.18) el resultado se sigue. ■

Corolario 4.3.2 *Para $n \geq 1$ las funciones de iteración de valores son*

$$V_n(x) = \frac{1 - \alpha^n}{1 - \alpha}.$$

Además $f_n(x) \in (0, \pi)$ y

$$f_n(x) = \frac{\pi}{2},$$

donde f_n es el maximizador de V_n.

Demostración. Por el Lema anterior se sabe que $V_{n-1}'(x) = 0$, entonces $V_{n-1}(x) = k_{n-1}$, implicando que $V_n(x) = \max_{a \in [0,\pi]} \{sen(a) + k_{n-1}\}$. Al maximizar a la función entre llaves se tiene que $f_n(x) = \frac{\pi}{2}$ y por lo tanto $f_n(x) \in (0, \pi)$. Además al sustituir f_n en la función entre llaves y recordando que $V_1(x) = 1$, se tiene que $V_n(x) = \frac{1-\alpha^n}{1-\alpha}$. ■

4.4. Un ejemplo Lineal Cuadrático

El siguiente ejemplo es muy común en la literatura de PDMs (véase [5] y [15], por ejemplo), además es utilizado en diversas áreas como: ingeniería ([13]), radiobiología ([27]), economía ([19] y [30]), entre otras. En la presentación que se hace acontinuación se caracteriza la política óptima mediante una ecuación de Euler, dicha ecuación permite aproximarla de manera numérica (véase Capítulo 5).

E problema Lineal Cuadrático (LQ) consiste de un sistema lineal con un costo cuadrático y es uno de los problemas de control más usados en ingeniería, economía y muchos otros campos. Considérese un sistema definido por la siguiente ecuación en diferencias

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

para $t = 0, 1, ...$ Donde γ y β son constantes reales. La función de costo está dada por

$$c(x, a) = qx^2 + ra^2,$$

donde q y r son constantes reales tales que $q \geq 0$ y $r > 0$. La sucesión $\{\xi_t\}$ de v.a's son i.i.d tomando valores en $S = \mathbb{R}$, con función de densidad continua Δ, independientes del estado inicial $x_0 \in X$, con media 0 y varianza finita σ^2, es decir, $E(\xi) = 0$ y $E(\xi^2) = \sigma^2 < +\infty$. Considere a $X = A = A(x) = \mathbb{R}$. Se desea minimizar a

$$v(\pi, x) = E_x^\pi \left[\sum_{t=0}^{\infty} \alpha^t \left(qx_t^2 + ra_t^2 \right) \right],$$

para $\pi \in \Pi$, $x \in X$.

En este caso las funciones de iteración de valores tendrán la forma siguiente:

$$V_n(x) = \min_{a \in A(x)} \left\{ qx^2 + ra^2 + \alpha E \left[V_{n-1}(\gamma x + \beta a + \xi) \right] \right\},$$

con $v_0(x) = 0$.

Lema 4.4.1 *Para este ejemplo es válido el Teorema de Programación Dinámica.*

Demostración. Véase [8] y [15]. ∎

Lema 4.4.2 *Para cada $x \in X$ y $n = 1, 2, ...,$ las funciones de iteración de valores satisfacen la ecuación de Euler*

$$V_n'(x) = 2qx + \alpha\gamma E \left[V_{n-1}' \left(\left(\gamma + \frac{\beta^2 q}{r\gamma} \right) x - \frac{\beta^2}{2r\gamma} V_n'(x) + \xi \right) \right]. \tag{4.19}$$

Demostración. La prueba de este Lema se sigue de manera análoga a los ejemplos anteriores. Para una prueba de dicho resultado se puede consultar en [29]. ∎

Corolario 4.4.3 *La función de valor óptimo y la política óptima para el problema LQ están dadas por*

$$V(x) = Kx^2 + T \ \ y \ f(x) = -\left(\frac{\alpha\beta\gamma K}{r + \alpha\beta^2 K}\right)x,$$

donde

$$T = \frac{\alpha}{1-\alpha}K\sigma^2,$$

y K es la solución positiva de la siguiente ecuación

$$K = q + \frac{\alpha\gamma^2 rK}{r + \alpha\beta^2 K},$$

la cual es conocida como ecuación de Ricatti.

Demostración. Es una consecuencia del Lema anterior. ■

Lema 4.4.4 *La política óptima satisface la siguiente ecuación de Euler, para $x \in X$*

$$\alpha q\gamma x + \left(\frac{\alpha q\beta^2 + r}{\beta}\right) f(x) = \frac{\alpha r\gamma}{\beta} E\left[f\left(\gamma x + \beta f(x) + \xi\right)\right]. \tag{4.20}$$

Demostración. Dado el Corolario anterior se sabe que la función de valor óptimo V es diferenciable. Sea $x \in X$ y considere a la función definida como

$$g(a) := qx^2 + ra^2 + \alpha E\left[V(\gamma x + \beta a + \xi)\right],$$

con $a \in A(x)$. Dado que la política óptima f minimiza a g, entonces por la condición de primer orden se tiene que

$$-2\frac{r}{\beta}f(x) = \alpha E\left[V'(\gamma x + \beta af(x) + \xi)\right]. \tag{4.21}$$

Además V satisface la fórmula de la envolvente

$$V'(x) = 2qx + \gamma\alpha E\left[V'(\gamma x + \beta af(x) + \xi)\right],$$

sustituyendo (4.21) en la última relación se llega a que

$$V'(x) = 2\left(qx - \frac{r}{\beta}f(x)\right), \tag{4.22}$$

entonces

$$V'(\gamma x + \beta af(x) + \xi) = 2\left(q\left(\gamma x + \beta af(x) + \xi\right) - \frac{r}{\beta}f(\gamma x + \beta af(x) + \xi)\right). \tag{4.23}$$

Finalmente sustituyendo (4.23) en (4.21) se sigue que (4.20) es válida. ■

4.5. Un ejemplo de consumo y producción

Uno de los problemas clásicos en economía es el de consumo y producción. Se desea maximizar una utilidad a lo largo de la vida, por lo que se busca determinar en base a esto, el consumo y el ahorro en cada instante de tiempo y controlar de forma dinámica que parte de su presupuesto iría destinada al consumo y que parte a la producción. Para ser más precisos, se desea asignar el capital presente x_t entre una parte de la producción y el consumo en cada periodo. Entonces el objetivo es maximizar a

$$v(\pi, x) = E\left[\sum_{t=0}^{\infty} \alpha^t U(a_t)\right],$$

donde $U : [0, \infty) \to \mathbb{R}$ es una función conocida como función de utilidad y se supone $U \in C^2((0, \infty); \mathbb{R})$, creciente, estrictamente cóncava y U' es invertible con inversa H.

Supóngase que la conexión entre las decisiones de inversión y capital acumulado esta dado por

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

para $t = 0, 1, 2, ...$, con $x_0 = x \in X = [0, \infty)$ conocido, de modo que el capital en el periodo $t+1$ sea proporcional al consumo. Bajo el supuesto que h es una función positiva, de clase C^2 y estrictamente cóncava sobre X, con derivada positiva y $h(0) = 0$. Supóngase que tanto U y h satisfacen las condiciones siguientes

$$U'(0) = h'(0) = \infty \text{ y } U'(\infty) = h'(\infty) = 0, \tag{4.24}$$

(en la literatura de Economía estas condiciones se conocen como condiciones de Inada, véase [12]).

Además, $\{\xi_t\}$ es una sucesión de v.a's i.i.d tomando valores en $S = [0, \infty)$, independientes al capital inicial $x_0 = x$ y con densidad Δ la cual se supone que es de clase C^2 en $int(S)$. Supóngase que el conjunto de acciones admisibles esta dado por $A(x) = [0, h(x)]$, con $x \in X$.

También, se considera que existen una función continua de peso w y constantes $c, \beta > 0$, tales que $1 \leq \beta < 1/\alpha$ y para cada $x \in X$

$$\sup_{a \in A(x)} |U(a)| \leq cw(x), \tag{4.25}$$

$$\sup_{a \in A(x)} E(w(\xi_t(h(x_t) - a_t)) \leq \beta w(x). \tag{4.26}$$

En casos particulares donde la dinámica y la utilidad son conocidas las Condiciones (4.25) y (4.26) pueden ser verificadas a partir de las características del modelo de control de Markov.

En este caso las funciones de iteración de valores tendrán la forma siguiente:

$$V_n(x) = \max_{a \in [0,h(x)]} \{U(a) + \alpha E[V_{n-1}(\xi(h(x) - a))]\},$$

con $V_0(x) = 0$.

Para $n \geq 1$, la función G^n usada en el Capítulo 3, presenta la forma siguiente

$$G^n(x, a) = U(a) + \alpha E[V_{n-1}(\xi(h(x) - a))],$$

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

Lema 4.5.1 *Para este ejemplo es válido el Teorema de Programación Dinámica 2.2.17.*

Demostración. Es fácil ver que $A(x)$ es compacto para cada $x \in X$. Dado que $U \in C^2((0,\infty);\mathbb{R})$, se tiene que U es continua en $A(x) = [0, h(x)]$.

Sea v una función continua y acotada definida en $X = [0,\infty)$, entonces

$$\overline{v}(x, a) = E[v(\xi(h(x) - a))] = \int \overline{v}(s(h(x) - a))\Delta(s)ds.$$

Aplicando el siguiente cambio de variable $u = s(h(x)-a)$, a la integral anterior, se tiene que para $a \neq h(x)$

$$\overline{v}(x, a) = \int \overline{v}(u)\Delta\left(\frac{u}{h(x) - a}\right)\frac{du}{h(x) - a},$$

la cual resulta continua dado que la densidad Δ lo es. Por otro lado, sea $l = h(x) - a$, entonces $\lim_{l\to 0} \overline{v}(x, a) = \overline{v}(x, h(x)) = v(0)$, por consiguiente la función $\overline{v}$ es continua y la Condición 3.2.1 se verifica.

Además, por (4.25) y (4.26) se tiene que las Condiciones de crecimiento (2.2.12) y (2.2.13) se satisfacen y por lo tanto el algoritmo de programación dinámica se verifica. ■

Lema 4.5.2 *La función de valor óptimo V es estrictamente cóncava y la política óptima f es única.*

Demostración. Sean $x, y \in X = [0,\infty)$ y $\lambda \in [0, 1]$. Si $a \in A(x) = [0, h(x)]$ y $\overline{a} \in A(y) = [0, h(y)]$, entonces $0 \leq \lambda a \leq \lambda h(x)$ y $0 \leq (1 - \lambda)\overline{a} \leq (1 - \lambda)h(y)$, implicando que $0 \leq \lambda a + (1 - \lambda)\overline{a} \leq \lambda h(x) + (1 - \lambda)h(y)$, como h es cóncava se concluye que $\lambda a + (1 - \lambda)\overline{a} \in A(h(\lambda x + (1 - \lambda)y)) = [0, h(\lambda x +$

$(1-\lambda)y)]$. Además si $x < y$, es fácil ver que $A(x) \subset A(y)$. Por lo tanto la multifunción $x \to A(x)$ es convexa y creciente.

Por otro lado, como h es estrictamente cóncava, entonces la función $F(x, a, s) = s(h(x) - a)$ es cóncava en $\mathbb{K}$ para cada $s \in S$ y como la función de utilidad U se supone estrictamente cóncava entonces por el Lema 3.1.5 se concluye que V es estrictamente cóncava y la política óptima f es única. ∎

Lema 4.5.3 *Para las funciones de iteración de valores existen sus maximizadores f_n y para $n \geq 2$ con $x > 0$, $f_n(x) \in (0, h(x))$. Además $V_n \in C^2(int(X); \mathbb{R})$ y $f_n \in C^1(int(X); A)$.*

Demostración. Sea $x > 0$. Como, $V_1(x) = \max\limits_{a\in[0,h(x)]} \{U(a)\}$ y U es creciente se tiene que $V_1(x) = U(h(x))$ y $f_1(x) = h(x)$. Además se sabe que U y h son de clase C^2 y cóncavas, por lo que V_1 es de clase C^2, $V_1'(x) = U'(h(x))h'(x)$ y $V_1''(x) < 0$. En este caso nótese que $V_1'(0) = \infty$.

Así por programación dinámica es válida la siguiente relación

$$V_2(x) = \max_{a\in[0,h(x)]} \{U(a) + \alpha E [V_1(\xi(h(x) - a))]\}.$$

En este caso para $a \in (0, h(x))$, $G_a^2(x, a) = U'(a) - \alpha E [V_1'(\xi(h(x) - a))\xi]$. Además por (4.24) se tiene que $G_a^n(x, 0) = +\infty$ y $G_a^n(x, h(x)) = -\infty$, entonces por el teorema del valor intermedio se sigue que existe $\overline{a} \in (0, h(x))$ tal que $G_a^2(x, \overline{a}) = 0$.

Como $G_{aa}^2(x, a) = U''(a) + \alpha E \left[V_1''(\xi(h(x) - a))\xi^2\right]$ se sigue que G^2 es de clase C^2. También se sabe que $U''(\cdot)$ y $V_1''(\cdot)$ son estrictamente cóncavas y por lo tanto $G_{aa}^2(x, a) < 0$ para $x > 0$ y $a \in (0, h(x))$ y $\xi > 0$, bajo el supuesto de que $\Pr(\xi = 0) = 0$. Por consiguiente G^2 es estrictamente cóncava y por lo tanto $\overline{a} := f_2(x)$ es único.

Por otro lado debido a que $f_2(x) \in (0, h(x))$ y $V_2(x) = G^2(x, f_2(x))$ pertenece a C^2 y es estrictamente cóncava (por lo cual $V_2''(x) < 0$), entonces por el Teorema de la envolvente se tiene que $V_2 \in C^2(int(X); \mathbb{R})$ y $f_2 \in C^1(int(X); A)$. Además, $V_2'(0) = \infty$ (debido a (4.24)).

Ahora, para $n > 2$ supóngase que $V_{n-1} \in C^2(int(X); \mathbb{R})$, $V_{n-1}''(x) < 0$ y $V_{n-1}'(0) = \infty$.

Debido a la definición de la función de iteración de valores

$$V_n(x) = \max_{a\in[0,h(x)]} \{U(a) + \alpha E [V_{n-1}(\xi(h(x) - a))]\}.$$

De manera análoga al argumento utilizado para el caso $n = 2$ se puede concluir que $f_n(x) \in (0, h(x))$, G^n es estrictamente cóncava y de clase C^2. Entonces por el Teorema de la envolvente se garantiza que $V_n \in C^2(int(X);\mathbb{R})$ y $f_n \in C^1(int(X);A)$. ■

Lema 4.5.4 *Las funciones de iteración de valores satisfacen la ecuación de Euler*

$$V_n'(x) = \alpha E\left\{V_{n-1}'\left[\xi\left(h(x) - H\left(\frac{V_n'(x)}{h'(x)}\right)\right)\right]\xi\right\}h'(x), \tag{4.27}$$

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

Demostración. Sea $x > 0$. Por el Lema 4.5.3 se sabe que $G^2(x,a) = U(a) + \alpha E\left[V_1(\xi(h(x)-a))\right]$, $a \in A(x)$, es de clase C^2. Así para $a \in (0,h(x))$

$$G_a^2(x,a) = U'(a) - \alpha E\left[V_1'(\xi(h(x)-a))\xi\right].$$

También se sabe que $f_2(x) \in (0,h(x))$ (por Lema 4.5.3) y maximiza $G^2(x,\cdot)$. Así $G_a^2(x,f_2(x)) = 0$, equivalentemente,

$$U'(f_2(x)) = \alpha E\left[V_1'(\xi(h(x) - f_2(x)))\xi\right]. \tag{4.28}$$

Además por el teorema de la envolvente $V_2'(x) = G_x^2(x,f_2(x))$, es decir,

$$V_2'(x) = \alpha E\left[V_1'(\xi(h(x) - f_2(x)))\xi\right]h'(x), \tag{4.29}$$

por (4.28) se tiene que (4.29) toma la forma siguiente

$$V_2(x) = U'(f_2(x))h'(x),$$

por la invertibilidad de U' se sigue que

$$f_2(x) = H\left(\frac{V_2(x)}{h'(x)}\right),$$

sustituyendo la ultima igualdad en (4.29) se tiene que (4.27) se satisface.

Sea $n > 2$, por el Lema 4.5.3 se sabe que V_{n-1} es de clase C^2 implicando que G^n es de clase C^2. Por lo tanto, si $a \in (0,h(x))$ se sigue que

$$G_a^n(x,a) = U'(a) + \alpha E\left[V_{n-1}'(\xi(h(x)-a))\xi\right].$$

Usando la condición de primer orden y el hecho de que $f_n(x) \in (0,h(x))$ es un maximizador de $G^n(x,\cdot)$, se tiene que

$$U'(f_n(x)) = \alpha E\left[V_{n-1}'(\xi(h(x)-a))\xi\right], \tag{4.30}$$

utilizando nuevamente el teorema de la Envolvente se tiene que

$$V_n(x) = \alpha E\left[V'_{n-1}(\xi(h(x)-a))\xi\right]h'(x), \tag{4.31}$$

por (4.30) se sigue que (4.31) toma la forma siguiente

$$V'_n(x) = U'(f_n(x))h'(x),$$

nuevamente por la invertibilidad de U' se sigue que

$$f_n(x) = H\left(\frac{V_n(x)}{h'(x)}\right).$$

Finalmente al sustituir esta ultima igualdad en (4.31) se tiene que (4.27) se satisface. ■

El siguiente lema nos permitirá deducir una Ecuación de Euler para las políticas, la cual determina cuando una política es óptima.

Lema 4.5.5 *Supóngase que la política óptima f es tal que $f(x) \in (0, h(x))$, para cada $x \in X$. Entonces para este ejemplo la función de valor óptimo $V \in C^2(int(X);\mathbb{R})$ y la política óptima $f \in C^1(int(X);A)$.*

Demostración. Sea $x > 0$ fijo. Defínase para $a \in A(x)$

$$G(x,a) := U(a) + \alpha E\left[V(\xi(h(x)-a))\right].$$

Por hipótesis $f(x) \in (0, h(x))$, entonces si se muestra que $G \in C^2(int(\mathbb{K});\mathbb{R})$ y que G_{aa} es definida negativa, por el Lema 3.3.5 el resultado se sigue.

Para probar que G es de clase C^2, se sabe que U es de clase C^2 entonces basta mostrar que la función $W(x,a) := E\left[V(\xi(h(x)-a))\right]$, con $a \in A(x)$, es de clase C^2. Sean $a \in A(x)$ y

$$W(x,a) := \int_0^\infty V(s(h(x)-a))\Delta(s)ds,$$

aplicando el siguiente cambio de variable, $u = s(h(x)-a)$, se tiene que

$$W(x,a) = \frac{1}{h(x)-a}\int_0^\infty V(u)\Delta\left(\frac{u}{h(x)-a}\right)du,$$

por el hecho de que la densidad Δ es de clase C^2 se sigue que W es de clase C^2.

Por otro lado, se sabe por el Lema 4.5.2 que V es estrictamente cóncava, por lo que W resulta ser cóncava. En resumen W es una función de clase C^2 y cóncava, implicando que W_{aa} es semidefinida negativa (véase [12]).

Finalmente dado que U_{aa} es definida negativa (por ser estrictamente cóncava) y que W_{aa} es semidefinida negativa entonces G_{aa} resulta ser definida negativa. ■

El siguiente resultado permite caracterizar a la política óptima para este ejemplo mediante su correspondiente ecuación de Euler.

Lema 4.5.6 *Para este ejemplo la política óptima f satisface la siguiente ecuación de Euler, para cada $x \in int(X)$*

$$U'(f(x)) = \alpha E\left[h'(\xi(h(x) - f(x)))U'\left(f\left(\xi\left(h(x) - f(x)\right)\right)\right)\xi\right]. \qquad (4.32)$$

Recíprocamente, si $f \in \mathbb{F}$ es una política tal que para cada $x \in int(X)$ satisface a (4.32) y la relación siguiente

$$\lim_{t\to\infty} E\left[\alpha^t h'(x_t)U'(f(x_t))x_t\right] = 0. \qquad (4.33)$$

Entonces f es óptima.

Demostración. Sea f la política óptima y $x > 0$. Por el Lema 4.5.5 se sabe que las funciones V y G son diferenciables, donde G está definida por

$$G(x,a) := U(a) + \alpha E\left[V(\xi(h(x) - a))\right],$$

$a \in A(x)$. Como $f(x)$ es el maximizador de G en $(0, h(x))$ entonces por la condición de primer orden se tiene que

$$U'(f(x)) = \alpha E\left[V'(\xi(h(x) - f(x)))\xi\right]. \qquad (4.34)$$

Además V' satisface la fórmula de la envolvente

$$V'(x) = \alpha E\left[V'(\xi(h(x) - f(x)))\xi\right] h'(x). \qquad (4.35)$$

Sustituyendo a (4.34) en (4.35) se concluye que

$$V'(x) = U'(f(x))h'(x),$$

para así obtener que

$$V'(\xi(h(x) - f(x))) = U'(f(\xi(h(x) - f(x))))h'(\xi(h(x) - f(x))), \qquad (4.36)$$

finalmente sustituyendo (4.36) en (4.34) se tiene que (4.32) es válida.

Por otra parte, sea $f \in \mathbb{F}$ satisfaciendo (4.32) y (4.33). Sea g otra política y denótese por $a_t = f(x_t)$ y $\overline{a}_t = g(x_t)$ los consumos generados por las respectivas políticas para $t = 0, 1, ...$

Dada la concavidad de U y por ser U de clase C^2, usando la Proposición 7.2.2 del Apéndice se tiene que

$$E\left[\sum_{t=0}^{T-1} \alpha^t \left(U(a_t) - U(\overline{a}_t)\right)\right] \geq E\left[\sum_{t=0}^{T-1} \alpha^t U'(a_t)(a_t - \overline{a}_t)\right].$$

Por definición, $x_{t+1} = \xi_t(h(x_t) - a_t)$ y $\overline{x}_{t+1} = \xi_t(h(\overline{x}_t) - \overline{a}_t)$, ignorando conjuntos de probabilidad cero correspondientes a los valores de $\xi = 0$, se sigue que

$$\frac{x_t - \overline{x}_t}{\xi_{t-1}} = h(x_{t-1}) - h(\overline{x}_{t-1}) - \left(a_{t-1} - \overline{a}_{t-1}\right),$$

equivalentemente,

$$a_{t-1} - \overline{a}_{t-1} = h(x_{t-1}) - h(\overline{x}_{t-1}) - \frac{x_t - \overline{x}_t}{\xi_{t-1}}.$$

Como h es una función cóncava y de clase C^2 se sigue por la Proposición 7.2.2 del Apéndice que

$$a_{t-1} - \overline{a}_{t-1} \geq h'(x_{t-1})(x_{t-1} - \overline{x}_{t-1}) - \frac{x_t - \overline{x}_t}{\xi_{t-1}}.$$

Así mediante algunos cálculos se llega a que

$$\begin{aligned}
\sum_{t=0}^{T-1} \alpha^t U'(a_t)(a_t - \overline{a}_t) \quad &\geq \quad \sum_{t=0}^{T-1} \alpha^t U'(a_t) \left(h'(x_t)(x_t - \overline{x}_t) - \frac{x_{t+1} - \overline{x}_{t+1}}{\xi_t}\right) \\
&= \quad \sum_{t=1}^{T-1} \left[\alpha^{t-1} \frac{(x_t - \overline{x}_t)}{\xi_{t-1}} \left[\alpha \xi_{t-1} h'(x_t) U'(a_t) - U'(a_{t-1})\right]\right] \\
&\quad\; - \frac{\alpha^{T-1}}{\xi_{T-1}} U'(a_{T-1})(x_T - \overline{x}_T) \\
&\geq \quad \sum_{t=1}^{T-1} \left[\alpha^{t-1} \frac{(x_t - \overline{x}_t)}{\xi_{t-1}} \left[\alpha \xi_{t-1} h'(x_t) U'(a_t) - U'(a_{t-1})\right]\right] \\
&\quad\; - \frac{\alpha^{T-1}}{\xi_{T-1}} U'(a_{T-1}) x_T.
\end{aligned}$$

La ultima desigualdad es verdad dado que U' es positiva y x_T, α, ξ también los son.

Por otra parte como $a_t = f(x_t)$ y f satisface (4.32) se sigue que

$$\begin{aligned} U'(a_{T-1}) &= \alpha E[h'(\xi_{T-1}(h(x_{T-1}) - f(x_{T-1}))) \\ &\quad U'\left(f\left(\xi_{T-1}(h(x_{T-1}) - f(x_{T-1}))\right)\xi_{T-1}\right)] \\ &= \alpha E\left[h'(x_T)U'(a_T)\xi_{T-1}\middle|\, x_{T-1}\right]. \end{aligned}$$

De está última relación se concluye que

$$E\left[\alpha^{T-1}\frac{x_T}{\xi_{T-1}}U'(a_{T-1})\right] = E\left[\alpha^{T-1}\frac{x_T}{\xi_{T-1}}\alpha E\left[h'(x_T)U'(a_T)\xi_{T-1}\middle|\, x_{T-1}\right]\right],$$

en este caso obsérvese que $\frac{x_T}{\xi_{T-1}} = h(x_{T-1}) - f(x_{T-1})$, es decir, $\frac{x_T}{\xi_{T-1}}$ depende de x_{T-1}. Por propiedades de la esperanza condicional se tiene que

$$E\left[\alpha^{T-1}\frac{x_T}{\xi_{T-1}}U'(a_{T-1})\right] = E\left[\alpha^T h'(x_T)U'(f(x_T))x_T\right],$$

la cual resulta ser la condición de transversabilidad (4.33).

Similarmente, nótese que $\frac{x_t - \overline{x}_t}{\xi_{t-1}}$ depende de x_{t-1} y $\overline{x}_{t-1}$. De aquí,

$$E\left[\frac{(x_t - \overline{x}_t)}{\xi_{t-1}}\left[\alpha\xi_{t-1}h'(x_t)U'(a_t) - U'(a_{t-1})\right]\right] =$$

$$E\left[\frac{(x_t - \overline{x}_t)}{\xi_{t-1}}E\left[\alpha\xi_{t-1}h'(x_t)U'(a_t) - U'(a_{t-1})\middle|\, x_{t-1}\right]\right],$$

dado que f satisface (4.32) se tiene que

$$U'(a_{t-1}) = E\left[\alpha\xi_{t-1}h'(x_t)U'(a_t)\middle|\, x_{t-1}\right],$$

implicando que

$$E\left[\alpha\xi_{t-1}h'(x_t)U'(a_t) - U'(a_{t-1})\middle|\, x_{t-1}\right] = 0,$$

por consiguiente

$$E\left[\frac{(x_t - \overline{x}_t)}{\xi_{t-1}}\left[\alpha\xi_{t-1}h'(x_t)U'(a_t) - U'(a_{t-1})\right]\right] = 0.$$

Por lo tanto para T suficientemente grande

$$E\left[\sum_{t=0}^{T-1}\alpha^t\left(U(a_t) - U(\overline{a}_t)\right)\right] \geq 0.$$

Por lo cual se concluye que f es una política óptima. ■

Capítulo 5

ALGORITMOS NUMÉRICOS PARA LA ECUACIÓN DE EULER

En este capítulo se presentan ejemplos de PDMs tanto deterministas como estocásticos, para los cuales se elaboraron programas en MAPLE 11. En particular, se elaboró el método iterativo propuesto en el Capítulo 4 para resolver la ecuación de Euler. También se presenta un método para aproximar a la política óptima de un problema Lineal Cuadrático (LQ).

5.1. Métodos iterativos mediante la ecuación de Euler

El método iterativo que proporciona la ecuación de Euler (4.1) presentada en el Capítulo 4, resulta ser más fácil de implementar a diferencia de programación dinámica, debido a que solo depende de las derivadas de las funciones de iteración de valores (I.V) (2.5). El siguiente algoritmo describe dicho método.

Algoritmo Iterativo de la Ecuación de Euler

Input

$\alpha \to$ factor de descuento;

$r \to$ función de recompensa;

$F \to$ Dinámica;

$ds \to$ Distribución de la v.a ξ;

Let

ra y $rx \to$ Derivadas de r con respecto a a y x;

Fa y $Fx \to$ Derivadas de F con respecto a a y x;

$H := (rx * Fa - ra * Fx)\, Fa^{-1}$;

$dV[0] := 0 =: V[0] \to$Valor inicial de las I.V;

For i from 1 to n do

$dV[i] := rx + \alpha E\,[dV[i](F(x, a, s)] * Fx$;

$f[i] := solve(dV[i] = H, a)$;

$dV[i] := subs(f[i] = a, dV[i])$;

$u[i] := int(dV[i]))$;

$c[i]$:$\to$cálculo de constantes involucradas en la integración;

$V[i] := u[i] + c[i]$;

End

5.1.1. Ejemplo Determinista

Con relación al ejemplo determinista de la Sección 4.3, a continuación se presenta un caso particular resuelto en MAPLE 11 mediante la Ecuación de Euler obtenida. Para el caso sin empleo se consideran los siguientes datos: el horizonte del problema es $N = 5$ con $\gamma = 3$, y en el caso con empleo se consideran los mismos datos junto con la ganancia $w = 5$. Al implementar el programa se obtuvo la siguiente gráfica, para el caso sin empleo:donde la función de color azul corresponde a la función de valor óptimo, la cual proporciona el programa y está dada por

$$V(x) = 15 \log\left(\frac{x}{5}\right),$$

$x > 0$. Dicha función coincide con la dada en el Corolario 4.2.3.

Para el caso con empleo MAPLE 11 presenta la siguiente gráfica, suponiendo que $x > 5$ En este caso la función en morado es la de valor óptimo y está dada por

$$V(x) = 15 \log\left(\frac{x + 20}{5}\right),$$

$x \geq 5$. Nuevamente la función que proporciona el programa coincide con la dada en el Corolario 4.2.6.

5.1.2. Ejemplo Lineal Cuadrático

Con relación al ejemplo Lineal Cuadrático presentado en la Sección 4.4, el método iterativo que proporciona la ecuación de Euler también fue implementado. Considere los siguientes datos: $\alpha = 0{,}5$, $q = r = \gamma = \beta = 1$ y la v.a ξ tiene una distribución normal estándar. Al implementar el programa se obtuvo la siguiente gráfica:la cual permite ver la convergencia de las funciones de iteración de valores a la función de valor óptimo que de acuerdo al Corolario 4.4.3 es de la forma

$$V(x) = \sqrt{2}(x^2 + 1).$$

También el programa proporciona los maximizadores f_n de las funciones de iteración de valores, cuyas gráficas se muestran a continuación: las cuales convergen a la política óptima:

$$f(x) = -\left(\frac{1}{1+\sqrt{2}}\right)x.$$

5.2. Método de la Proyección

El objetivo de esta sección es proporcionar un método numérico denominado el método de la proyección (o método de residual ponderado, véase [14]). Dicho método sirve para aproximar la solución de una ecuación funcional, esta aproximación se representa como una combinación lineal de alguna familia de funciones, por ejemplo una familia de polinomios. Los coeficientes involucrados en dicha aproximación son los objetivos a ser calculados para obtener una solución aproximada. Estos coeficientes se determinan al considerar el residual de la ecuación funcional igual a cero en un sentido promedio (de aquí el nombre de residual ponderado).

Se desea aproximar a una función $f : X \to Y$, donde X y Y son subconjuntos de $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente. Esta función se encuentra definida implícitamente por la ecuación

$$W(f) = 0,$$

donde $W : E_1 \to E_2$ es un funcional, E_1 y E_2 son espacio de funciones dados.

Dada una familia de funciones $\{\varphi_t\} \subseteq E_1$, se aproxima a f como una combinación lineal de los primeros $p+1$ miembros de esta familia, es decir,

$$\widehat{f}(x) := \sum_{i=0}^{p} \gamma_i \varphi_i(x),$$

con $x \in X$, dicha familia puede ser por ejemplo polinomios ortogonales (véase sección 7.3 del Apéndice). La función residual es obtenida al evaluar a $\widehat{f}$ en el funcional W, es decir,

$$R(\overline{\gamma}, x) := W(\widehat{f}(\overline{\gamma}, x)),$$

donde $\overline{\gamma} := (\gamma_0, ..., \gamma_p)$. Supóngase que existen un conjunto de funciones $\{g_i\}_{i=0}^p \subset E_1$ conocidas como funciones proyección y una función de peso w definida sobre X, tales que inducen un producto interior (y por tanto una distancia) sobre E_1 dado por

$$\int R(\overline{\gamma}, x) g_i(x) w(x) dx.$$

Se desea determinar al vector de parámetros $\overline{\gamma}$ tal que tenga la siguiente propiedad: para cada $i = 0, 1, ..., p$

$$\int R(\overline{\gamma}, x) g_i(x) w(x) dx = 0.$$

La forma específica de las funciones proyección $\{g_i\}_{i=0}^p$ y de la función de peso w, permiten determinar al vector de parámetros $\overline{\gamma}$, por ejemplo los métodos de Galerking o mínimos cuadrados (véase [14]), este último servirá como ejemplo.

De acuerdo al método de mínimos cuadrados, se desea minimizar a

$$\int R(\overline{\gamma}, x)^2 dx,$$

con respecto al vector $\overline{\gamma} := (\gamma_0, ..., \gamma_p)$. Las condiciones de primer orden para este problema están dadas por

$$\int R(\overline{\gamma}, x) \frac{dR(\overline{\gamma}, x)}{d\gamma_i} dx = 0, \tag{5.1}$$

para $i = 0, 1, ..., p$. Resulta fácil ver que las funciones proyección son $g_i = \frac{dR}{d\gamma_i}$, para $i = 0, 1, ..., p$ y la función de peso es $w = 1$.

Algoritmo del Método de la Proyección

Propósito: Aproximar a $f : X \to Y$ de una ecuación funcional $W(f) = 0$.
Paso 1. Elegir una familia de funciones $\{\varphi_i\}_{i=0}^\infty$ con $\varphi_i : X \to Y$.

Paso 2. Elegir el grado p de la aproximación y defina $\widehat{f}(x) := \sum_{i=0}^{p} \gamma_i \varphi_i(x)$.

Paso 3. Definir el residual

$$R(\overline{\gamma}, x) := W(\widehat{f}(\overline{\gamma}, x)).$$

Paso 4. Elegir las funciones proyección $\{g_i\}_{i=0}^{p}$, la función de peso w y encontrar el valor de $\overline{\gamma}$ tal que para $i = 0, 1, ..., p$

$$\int R(\overline{\gamma}, x) g_i(x) w(x) dx = 0.$$

Paso 5. Verificar la calidad de la solución $\overline{\gamma}$. Sí es necesario, regresar al Paso 2 e incrementar el grado del polinomio o regresar al Paso 1 y elegir otra familia de funciones.

A continuación se presenta un ejemplo, el cual se ha trabajado en capítulos anteriores usando la ecuación de Euler. A pesar de que este ejemplo se puede resolver de forma directa el objetivo de la siguiente sección es comparar la solución aproximada con la obtenida en el Capítulo 4.

5.2.1. Ejemplo Lineal Cuadrático

Con relación al ejemplo Lineal Cuadrático presentado en la Sección 4.4, considere que el espacio de estados y de acciones es $X = A = [-10, 10]$. Considere también los siguientes datos: $\alpha = 1/2$, $q = r = \gamma = \beta = 1$, y la dinámica tiene la forma siguiente

$$x_{t+1} = x_t + a_t.$$

Para este ejemplo, el Corolario 4.4.3 asegura que la política óptima es de la forma

$$f(x) = -\left(\frac{1}{1+\sqrt{2}}\right) x, \tag{5.2}$$

$x \in X$. Por el Lema 4.4.4 se sabe que f satisface la siguiente Ecuación de Euler

$$[W(f)](x) := \alpha x + (\alpha + 1) f(x) - \alpha f(x + f(x)) = 0.$$

Se aproxima la política óptima f mediante una combinación lineal de miembros de una familia de polinomios ortogonales $\{T_i\}$, conocida como polinomios de Chebyshev (véase Definición 7.3.3 del Apéndice), dicha aproximación se considera de grado 5, es decir,

$$\widehat{f}(x) := \sum_{i=0}^{5} \gamma_i T_i(Z(x)),$$

donde $Z(x) = 2x/20$ (véase (7.2) en la sección 7.3 del Apéndice).

Las funciones proyección que se usan son las otorgadas por mínimos cuadrados, debido al Teorema de la mejor aproximación (Véase Teorema 7.3.2 del Apéndice). Se define el residual como

$$R(\overline{\gamma}, x) = \alpha x + (\alpha + 1)\,\widehat{f}(x) - \alpha \widehat{f}\left(x + \widehat{f}(x)\right),$$

$x \in X$. Para resolver el sistema no lineal (5.1) se aplica el algoritmo de Newton-Raphson modificado (véase Sección 7.4 del Apéndice). El programa elaborado en MAPLE 11 presenta la siguiente gráfica.

Dicha gráfica ilustra la solución (5.2) y la aproximación por el método de la proyección, en la cual se observa un ajuste muy adecuado. La aproximación dada por el programa presenta los siguientes datos para el vector $\overline{\gamma}$

$$\overline{\gamma} = \begin{bmatrix} 3{,}59554 * 10^{-15} \\ -{,}414213 \\ 2{,}43905 * 10^{-16} \\ -1{,}65622 * 10^{-16} \\ 2{,}63568 * 10^{-18} \\ 1{,}62068 * 10^{-18} \end{bmatrix},$$

de lo cual se concluye que

$$f(x) \approx -0{,}414213x.$$

Al comparar con la solución dada en (5.2), se obtuvo el error de $-0{,}0000035624$.

Capítulo 6

CONCLUSIONES

En el trabajo se estudió la Teoría de Procesos de Decisión de Markov (PDMs) y sus conceptos generales planteando al problema de control óptimo (PCO). Se trabajó con la herramienta esencial para resolver dicho problema, conocida como programación dinámica (PD), cuya característica es el Algoritmo de Iteración de Valores, el cual consiste en determinar a la función de valor óptimo mediante una sucesión de funciones conocidas como funciones de iteración de valores.

Se estudian propiedades cualitativas sobre el modelo de control de Markov como la unicidad de la política óptima, concavidad, continuidad y diferenciabilidad de la función de valor óptimo y de las funciones de iteración de valores. Las cuales permiten obtener una de las contribuciones principales de la tesis, una versión de la ecuación de Euler (EE) en el contexto de PDMs, para el caso estocástico. Dicha ecuación resulta ser un método complementario a PD y se utiliza de manera iterativa para resolver el PCO.

También se presentan ejemplos los cuales se resuelven vía la EE, destacando un problema de consumo y producción. Para este se encontraron condiciones suficientes para caracterizar la política óptima mediante la ecuación de Euler.

Por último se han derivado problemas que se pretenden trabajar a futuro, tales como:

- Obtener condiciones suficientes para caracterizar la política óptima vía la EE.
- Garantizar la interioridad de políticas óptimas (véase Capítulos 3 y 4).
- Dar condiciones para mostrar la diferenciabilidad de orden superior de la función de valor óptimo y de la política óptima para el caso determinista (véase Capítulo 3).
- Estudiar aproximaciones asintóticas de la solución óptima (función de

valor y política óptima).

• Realizar un análisis cualitativo de soluciones óptimas de problemas de control mediante la EE.

• Desarrollar métodos numéricos basados en la EE, como el método de la proyección implementado en el Capítulo 5.

Capítulo 7

APÉNDICE

7.1. Funciones Semicontinuas

Definición 7.1.1 *Sean (X, d) un espacio métrico y $u : X \to \mathbb{R} \cup \{+\infty\}$ una función tal que $u(x) < \infty$ para al menos una $x \in X$, la función v es semicontinua superiormente (u.s.c) en $x \in X$, si*

$$\limsup_{n\to\infty} u(x_n) \leq u(x),$$

para cualquier sucesión $\{x_n\}$ en X convergente a $x \in X$.

Si u es u.s.c para toda $x \in X$, entonces u se dice que es semicontinua superiormente (u.s.c) en X.

La función u es semicontinua inferiormente (l.s.c) si $-u$ es u.s.c.

Proposición 7.1.2 *Una función u es u.s.c (l.s.c) y acotada superiormente (inferiormente) en X, si y sólo si, existe una sucesión $\{u_n\}$ de funciones continuas y acotadas sobre X, tales que u_n convergen de manera decreciente (creciente) a u ($u_n \downarrow u$, $u_n \uparrow u$, respectivamente).*

Demostración. Véase [3] o [6]. ■

7.2. Funciones Cóncavas

Sean X y Y subconjuntos convexos no vacíos de $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente y suponga que Y es dotado de un orden parcial, por ejemplo el orden lexicográfico. Una función $u : X \to Y$ es cóncava si para cualesquiera $x, y \in X$ y $\lambda \in [0, 1]$ se tiene que

$$u(\lambda x + (1 - \lambda)y) \geq \lambda u(x) + (1 - \lambda)u(y).$$

Si la desigualdad anterior es estricta entonces u es una función estrictamente cóncava.

Proposición 7.2.1 *Supóngase que X es abierto y sea u es función cóncava en X. Entonces u es continua en X.*

Demostración. Véase [12]. ■

Proposición 7.2.2 *Supóngase que X es abierto y u es una función de clase C^1 en X. Entonces u es cóncava, si y sólo si, para cualesquiera $x, y \in X$ se tiene que*

$$u(x) - u(y) \leq u'(y)(x - y).$$

Además, u es estrictamente cóncava, si sólo si, para cualesquiera $x, y \in X$, la desigualdad anterior es estricta.

Demostración. Véase [12]. ■

7.3. Polinomios Ortogonales

Sea w una función de peso definida sobre $[a, b]$ tal que $\int_a^b w(x)dx < \infty$. Se induce el siguiente producto interior sobre $\mathbb{C}([a, b])$. Para $f, g \in \mathbb{C}([a, b])$

$$\langle f, g \rangle = \int_a^b f(x)g(x)w(x)dx. \tag{7.1}$$

Definición 7.3.1 *Una familia $\{\varphi_i\}$ en $\mathbb{C}([a, b])$ es un sistema ortogonal si*

$$\left\langle \varphi_i, \varphi_j \right\rangle = 0,$$

siempre que $i \neq j$.

Supóngase que se desea aproximar a $f \in \mathbb{C}([a, b])$ por

$$\widehat{f}(x) := \sum_{i=0}^{p} \gamma_i \varphi_i(x),$$

donde $\{\varphi_i\}$ es una familia de polinomios ortogonales en $\mathbb{C}([a, b])$ y $\gamma_0, \gamma_1, ...\gamma_p$, son números arbitrarios. El objetivo es determinar los parámetros γ_i, $i =$

$0, 1, ..., p$ tales que $\widehat{f}$ sea la mejor aproximación a f. Para medir el error de dicha aproximación se usa la norma inducida por el producto interior (7.1). El Lema siguiente dice cómo encontrar a los parámetros que dan la mejor aproximación (método de mínimos cuadrados).

Lema 7.3.2 *Sean $f \in \mathbb{C}([a,b])$ y $\{\varphi_i\}$ una familia ortogonal en $\mathbb{C}([a,b])$. Defínase a*

$$\widehat{g}(x) := \sum_{i=0}^{p} \lambda_i \varphi_i(x) \text{ y } \widehat{f}(x) := \sum_{i=0}^{p} \gamma_i \varphi_i(x),$$

donde para cada $i = 0, ..., p$, $\lambda_i \in \mathbb{R}$ y $\gamma_i = \langle f, \varphi_i \rangle$. Entonces

$$\left\| f - \widehat{f} \right\| \leq \| f - \widehat{g} \| ,$$

donde $\|\cdot\|$ es la norma inducida por el producto interior definido en (7.1). Además la igualdad ocurre, sí y sólo si, $\lambda_i = \gamma_i$, para $i = 0, ..., p$.

Demostración. Véase [2]. ■

Polinomios de Chebyshev

Una clase particular de polinomios ortogonales son lo de Chebyshev, estos polinomios tiene la propiedad de ser calculados de manera iterativa, facilitando los cálculos computacionales y la implementación.

Definición 7.3.3 *El dominio de los polinomios de Chebsyshev es el intervalo $[-1, 1]$ y el n-ésimo miembro de esta familia se define por*

$$T_n(z) = \cos(n \arccos x) .$$

La función de peso para la cual esta familia es ortogonal es

$$w(z) := \frac{1}{\sqrt{1 - z^2}}.$$

Es muy común que la función que se desea aproximar esté definida en un intervalo $[a, b]$ diferente al $[-1, 1]$. Sin embargo no hay problema de aproximar a f mediante una combinación lineal de polinomios de Chebyshev. Para ello se utiliza la siguiente transformación, para $x \in [a, b]$

$$Z(x) := \frac{2x - (a + b)}{b - a}, \tag{7.2}$$

así la aproximación es

$$\widehat{f}(x) := \sum_{i=0}^{p} \gamma_i T_i(Z(x)).$$

El polinomio de Chebyshev de grado $n+1$ se puede calcular conociendo a los polinomios de grado n y $n-1$, es decir, los polinomios de Chebyshev satisfacen la relación

$$T_{n+1}(z) = 2xT_n(z) - T_{n-1}(z),$$

para $n \geq 1$, donde $T_0(z) = 1$ y $T_1(z) = z$ (véase [14]). Esta fórmula recursiva es eficiente para evaluar el polinomio en un punto dado z y economiza los cálculos computacionales.

7.4. Método de Newton-Raphson

El método de Newton-Raphson es un algoritmo numérico que permite aproximar las raices de un sistema de ecuaciones no lineales (por ejemplo el sistema (5.1) en la Sección 5.2). Supóngase que se desea resolver el sistema de n ecuaciones

$$\begin{aligned} f^1(x_1, x_2, ...x_n) &= 0, \\ f^2(x_1, x_2, ...x_n) &= 0, \\ \vdots \quad &= \vdots \\ f^n(x_1, x_2, ...x_n) &= 0. \end{aligned} \tag{7.3}$$

donde el vector $x := [x_1, x_2, ...x_n]$ es desconocido. El sistema anterior es equivalente a

$$\mathbf{f}(\overline{x}) = 0,$$

donde $\mathbf{f} = \left[f^1, f^2, ..., f^n\right]^{tr}$. La matriz Jacobiana de $\mathbf{f}$ es denotada por $J(x)$, con

$$J(x) := \begin{bmatrix} f_1^1 & f_2^1 & \cdots & f_n^1 \\ f_1^2 & f_2^2 & \cdots & f_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ f_1^n & f_2^n & \cdots & f_n^n \end{bmatrix},$$

donde

$$f_j^i := \frac{\partial f^i(x)}{dx_j},$$

para $i = 1, 2, ..., n$ y $j = 1, 2, ..n$.

La bola abierta con centro en x y radio $r > 0$, se denota por $B_r(x)$ y el Jacobiano J se dice ser de Lipschitz en $B_r(x)$ con constante γ si para cualesquiera $y, z \in B_r(x)$ se tiene que

$$\|J(y) - J(z)\| \leq \gamma \|y - z\|.$$

La aproximación lineal de $\mathbf{f}$ en un punto inicial x_0 es

$$g(x) := \mathbf{f}(x_0) + J(x_0)dx,$$

con $dx = (x - x_0)$. Resolviendo para $g(x_1) = 0$ y suponiendo que $J(x_0)$ es invertible, se tiene que

$$x_1 = x_0 - J(x_0)^{-1}\mathbf{f}(x_0),$$

así x_1 es una primera aproximación a la raíz x^* del sistema (7.3). En caso contrario se busca de manera local un x_1' tal que sea raíz de

$$g(x) := \mathbf{f}(x_1) + J(x_1)(x - x_1),$$

generando así el método iterativo

$$x_{k+1} = x_k - J(x_k)^{-1}\mathbf{f}(x_k). \tag{7.4}$$

El siguiente teorema muestra que bajo ciertas condiciones la sucesión $\{x_k\}$ convergen a la raíz x^*, si x_0 es cercano a x^*.

Teorema 7.4.1 *Sea $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ una función continuamente diferenciable en $D \subset \mathbb{R}^n$, donde D es abierto y convexo. Supóngase que existen $x^* \in \mathbb{R}^n$, y $r > 0$, tales que $B_r(x^*) \subset D$, $\mathbf{f}(x^*) = 0$, $J(x^*)^{-1}$ existe y J es Lipschit en $B_r(x^*)$ con constante γ. Entonces existe $\varepsilon > 0$ tal que para cualquier $x_0 \in B_\varepsilon(x^*)$ la sucesión gereada por (7.4) está bien definida y converge a x^*.*

Demostración. Véase [11]. ■

Algoritmo de Newton-Raphson

Propósito. Resolver $\mathbf{f}(x) = 0$, donde $\mathbf{f} = \left[f^1, f^2, ..., f^n\right]^{tr}$ y $x \in \mathbb{R}^n$.

Paso 1. Elegir el valor inicial x_0.

Paso 2. Calcular el jacobiano de $\mathbf{f}$ en x_0, $J(x_0)$ y resolver $J(x_0)dx = -\mathbf{f}(x_0)$. Si $x_1 = x_0 - dx$ no es buena aproximación elegir $\lambda \in (0, 1)$ tal que $x_2 = x_0 - \lambda dx$ si lo sea y definir $x_1 = x_2$.

Paso 3. Verificar la convergencia: sí $\|\mathbf{f}(x_1)\|_\infty < \varepsilon$ y/o $\left|x_i^1 - x_i^0\right| / \left(1 + \left|x_i^0\right|\right) < \varepsilon$, para toda $i = 1, 2, ..., n$ y una tolerancia dada ε, parar, de lo contrario $x_0 = x_1$ y regresar al paso 2.

Bibliografía

[1] Aliprantis C. D. and Burkinshaw O., Principles of Real Analysis, Academic Press, New York, 1981.

[2] Apostol T., Mathematical Analysis, Addison-Wesley, 1981.

[3] Ash R. B. and Doléans-Dade C. A., Probability and Measure Theory. Academic Press Elsevier, San Diego, 2005.

[4] Bellman R., Dynamic Programming. Dover, 2003.

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

[6] Bertsekas D. P. and S. E.Shreve, Stochastic Optimal Control: The Discrete-Time Case. Athena Scientific, 1996.

[7] Bhattarai K., Consumption, investment and financial intermediation a Ramsey model, Applied Financial Economics Letters, Vol. 1, pp. 329–333, 2005.

[8] Cruz-Suárez D., Montes-de-Oca R y Salem-Silva F., Conditions for uniqueness of optimal policies of discounted Markov decision processes, Mathematical Methods of Operations Research, Springer Verlag,V. 60, N. 3, 2004.

[9] Cruz-Suárez H. y Montes-de-Oca R., Discounted Markov control processes induced by deterministic systems. Kybernetika (Prague), V. 42 N. 6, pp. 647-664, 2006.

[10] Cruz-Suárez H. y Montes-de-Oca R., An envelope theorem and some applications to discounted Markov decision processes, Mathematical Methods of Operations Research, Springer Verlag,V. 67, N. 2, 2008.

[11] Dennis J. and Schnabel R., Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Society for Industrial Mathematics, SIAM, 1987.

[12] De la Fuente A., Mathematical Methods and Models for Economists, Cambridge University Press, 2000.

[13] Giraldo E., Aplicación de un regulador lineal cuadrático usando un observador de orden mímino sobre un puente grua, Scienta Et Technica, Vol. 13 no. 37, pp. 91-95, 2007.

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

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

[16] Hernández-Lerma O. and Lasserre J. B., Further Topics on Discrete-Time Markov Control Processes, Springer-Verlag, New York, 1999.

[17] Himmelberg C.J., Parthasarathy T. and Van Vleck F.S., Optimal Plans for Dynamic Programming Problems, Mathematics of Operations Research, V. 1 N. 4, pp 390-394, 1976.

[18] Janz N., Interpreting estimation result of Euler equation investment models when factor markets are imperfectly competitive, Center for European Economic Research in its Series, pp. 97-29, 1997.

[19] Kenneth D., A Variance Bounds Test of the Linear Quadratic Inventory Model, The Journal of Political Economy, Vol. 94, No. 2, pp. 374-401, 1986.

[20] Levhari D and Srinivasan T.N., Optimal savings under uncertainty, Review of Economic Studies, pp 153-163, 1968.

[21] Mirman L. and Zilcha I., On optimal Growth under unvertainty, Journal of Economic Theory, Vol. 2, pp 329-339, 1975.

[22] Ize J., Cálculo de Variaciones, V Coloquio del Departamento e Matemáticas, Centro de Investigación y de Estudios Avanzados del IPN, 1987.

[23] Puterman M. L., Markov Decision Processes: Discrete Stochastic Dynamic Programming, Wiley, New York, 1994.

[24] Royden H. L., Real Analysis, Macmillan, 1988.

[25] Schäl M., Conditions for optimality and for the limit of n-stage optimal policies to be optimal, Z. Wahrs.Verw, Gerb. 32, V. 32, pp 179-196, 1975.

[26] Semmler W., Consumption Based Asset Princing Models,

[27] Sendra Portero F., Radiobiología en braquiterapia de baja tasa de dosis, Radiobiología 2, pp. 26-33, 2002.

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

[29] Zacarías-Espinoza G. and Cruz-Suárez H., Un Método iterativo para resolver un modelo Lineal Cuadrático, Memorias de la Octava Conferencia Iberoamericana en Sistemas, Cibernética e Informática CISCI 2009, Orlando Fl., V. 1, 2009.

[30] Zietz J., Dynamic Programming: an introduction by example, Department of economics and finance, working paper series, 2004.