Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional

Cinvestav

Unidad de Zacatenco

Departamento de Matemáticas

El Enfoque del Descuento Desvaneciente para el Control de Sistemas Deterministas con Costo Promedio

Tesis presentada por

Alexis Pantoja Pineda

Para obtener el grado de
Maestro en Ciencias
En la Especialidad de Matemáticas

Director de Tesis:

Dr. Onésimo Hernández Lerma.

Ciudad de México Septiembre 2021

Resumen.

En este trabajo estudiaremos problemas de control con costo promedio a largo plazo en sistemas deterministas a tiempo discreto. En el caso estocástico, el problema de costo promedio (o costo ergódico) está bien estudiado, por ejemplo si retrocedemos al artículo pionero escrito por Richar Bellman, *A Markovian decision process [2]. Sin embargo, en el caso determinista, no tenemos herramientas como "ergodicidad" o "irreducibilidad" que son usadas en los sistemas Markovianos. Por lo tanto, en este trabajo nos restringimos al enfoque del descuento desvaneciente, el cual usa algunos de los más conocidos teoremas abelianos. Nuestra presentación sobre el problema de costo promedio está principalmente basada en el artículo de Vega-Amaya [16], el cual de hecho presenta una versión más sencilla del trabajo hecho por Feinberg et al. [5]. Por lo tanto nuestro resultado principal puede ser remontado este último artículo.*

Este trabajo está organizado en cuatro capítulos. El capítulo 1 introduce el problema de control determinista a tiempo discreto que vamos a tratar. También presenta material preliminar de la teoría de control óptimo. En el capítulo 2 planteamos y probamos el resultado principal, el teorema 2.1 y el teorema 2.2. Para ilustrar estos resultados, en el capítulo 3 presentamos varios ejemplos y aplicaciones de los problemas de costo promedio. Concluimos en el capítulo 4 con un apéndice de material preliminar.

Abstract.

The vanishing discount approach to the long-run average cost of the deterministic control systems

We study long-run average cost problems for deterministic discrete-time control systems. In the stochastic case, the average cost problem (also known as ergodic control) is well-known, going back to the pioneering paper by Richard Bellman, A Markovian decision process [2]. In the deterministic case, however, we do not have the tools of "ergodicity" and "irreducibility" used for Markovian systems. Hence, in this work we restrict ourselves to the vanishing discount approach, which uses standard Abelian theorems. Our presentation on the average cost problem is mainly based on the paper by Vega-Amaya [16], which in fact presents a simpler version of the work by Feinberg et al. [5]. Hence, our main result can be traced back to the latter paper.

This work is organized in four chapters. Chapter 1 introduces the discrete-time deterministic control problem we are concerned with. It also presents background material on optimal control theory. In Chapter 2 we state and prove the main results, Theorem 2.1 and Theorem 2.2. To illustrate this results, in Chapter 3 we present several examples and applications of average cost problems. We conclude in Chapter 4 with an appendix on background material.

Índice general

Introdución

En matemáticas (y más generalmente, en la vida), los problemas no siempre tienen una sola respuesta correcta. De hecho, podemos tener varios métodos que lleguen a un mismo resultado o incluso tener varios métodos con diferentes resultados. Tal es el caso de los problemas de control óptimo. En este documento se estudiará el "Costo promedio a largo plazo", a veces también llamado costo ergódico, el cual es un famoso criterio para resolver problemas de control óptimo. Si tenemos una función de costo $c(\cdot,\cdot)$ *que modela cierto problema y depende del estado actual y de la acción a tomar en el estado actual, el costo promedio a largo plazo es el límite de los costos promedios acumulados, es decir, tomará la siguiente forma*

$$\overline{V}(\cdot,\cdot) := \limsup_{k\to\infty} \frac{1}{k}\sum_{n=0}^{k} c(\cdot,\cdot).$$

Este criterio de optimización fue originalmente introducido por Richard Bellman en 1957 [2] para un tipo especial de procesos de decisiones de Markov, sin embargo el concepto de costo promedio a largo plazo es también conocido en economía y la forma simbólica de definirlo es matemáticamente familiar, por ejemplo aparece en la ley de los grandes números, sumas de Cesàro, teoremas ergódicos o integrales expresadas por sumas de Riemann con particiones uniformes.

Más concretamente, la pequeña contribución de esta disertación será desarrollar la aproximación $\alpha-$*descontada y un resultado ergódico en el caso determinista en sistemas discretos, pues estos resultados son bien sabidos en el caso no determinista (ver por ejemplo [8]). La aproximación* $\alpha-$*descontada es un procedimiento "estandar" para el estudio del costo promedio a largo plazo, el cual tiene como objetivo principal encontrar una sucesión de acciones* π' *tales que alcancen el mínimo posible entre todas las posibles de acciones* π

$$\overline{V}(\cdot,\pi') = \inf_{\pi} \overline{V}(\cdot,\pi).$$

La aproximación $\alpha-$*descontada está inspirada en el descuento desvaneciente, la cual fue introducida primero por Blackwell en 1962 [3], que consiste en tomar problemas descontados por algún factor* α *en* $(0,1)$*, y despúes en la aproximación* $\alpha-$*descontada se toma el límite cuando* α *tiende a* 1 *por la izquierda. Por otro* *lado, el resultado ergódico retoma la forma clásica de teoremas ergódicos como la ley de los grandes números o el teorema de Birkhoff-Khinchin en la teoría de probabilidad. Parafraseando el resultado, nos dice bajo qué condiciones el costo promedio a largo plazo óptimo es equivalente al promedio en tiempo del costo acumulado óptimo, es decir*

$$\lim_{k\to\infty} \frac{1}{k} V_k(\cdot) = \overline{V}(\cdot),$$

donde $V_k(\cdot)$ *es el costo óptimo al tiempo* k.

La importancia de estudiar la ergodicidad en sistemas aleatorios ha ayudado en varios campos de la ciencia. Las aplicaciones van desde crear herramientas para facilitar cálculos en economía [13] con la llamada hipótesis ergódica hasta predecir el comportamiento humano [12] con la teoría ergódica de decisión/acción, donde se predice como los individuos deben tolerar el riesgo en diferentes entornos. En nuestro caso veremos como el teorema ergódico nos ayudará a calcular el costo promedio a largo plazo en ejemplos concretos.

Para cumplir con nuestros objetivos pasaremos por un par de herramientas. En el capítulo 1 empezaremos a ver propiedades de las funciones semicontinuas por abajo. Dichas funciones nos ayudarán a generalizar los resultados haciendo que las suposiciones sean menos restrictivas. También mencionáremos algunos resultados que se saben en la teoría de control óptimo, tales como la ecuación de optimalidad del costo promedio y el control óptimo descontado. En el capítulo 2 veremos los resultados importantes que son la aproximación $\alpha-$*descontada y el teorema ergódico. La aproximación* $\alpha-$*descontada está guiada e inspirada por [16] y parte de la contribución es simplificar las pruebas y las hipótesis de algunos resultados. La aproximación* α*-descontada desprende el resultado ergódico el cual sigue los pasos de [15]. En el último capítulo tenemos un conglomerado de ejemplos, que fueron encontrados en diversas fuentes como [8], [9], [10], [11], [1]. En dichas fuentes, los ejemplos se encuentran planteados en su caso no determinista y se trata de estudiar su costo promedio a largo plazo mediante los resultados del capítulo 2.*

Capítulo 1

Sistemas deterministas discretos.

1.1. El modelo determinista.

Pensemos en el siguiente modelo sistémico:
Considere que en cierto contexto estamos en un estado incial x_0. *Entre un repertorio de acciones que podemos tomar a partir del estado* x_0, *escogemos una acción* a_0 *y mediante una regla de transición* F *avanzaremos a un nuevo estado* x_1 *(que depende del par* (x_0, a_0)*). De nuevo podremos tomar otra acción* a_1 *entre muchas otras acciones para* x_1 *e inductivamente se escogerán* $(a_0, a_1, ...)$ *tal que*

$$x_{k+1} = F(x_k, a_k), \quad k = 0, 1, ...$$

Sin embargo tomar cada acción requiere un costo, ya sea dinero, energía, combustible, materia prima, etc. Dicho costo lo denotamos por c *y comunmente lo llamamos el costo por etapa. Como el costo, en general, depende del estado del sistema y de la acción que se tome, pensaremos que* c *es una función de dos variables "*$c(\cdot, \cdot)$*".*

Este modelo es considerado típico en la teoría de control óptimo. Pongamos en contexto matemático dicho modelo. Sea X *el espacio de estados y* A *el espacio de acciones admisibles (generalmente son solo espacios de Borel). Para cada estado* x *en* X *existe* $A(x) \subset A$, *llamado el conjunto de las acciones admisibles para el estado* x, *el cual es siempre medible en* A *y como lo dice su nombre, consta de todas las acciones* a *en* A *que admite el sistema dado que nos encontramos en el estado* x. *Definiendo* $\mathbb{K} := \{(x, a) : x \in X, a \in A(x)\} \in \mathcal{B}(X \times A)$, *entonces la regla de transición es una función medible* $F : \mathbb{K} \to X$. *Por último el costo por etapa es también una función medible* $c : \mathbb{K} \to \mathbb{R}$.

Dado un estado inicial x_0 *en* X, *por una política de lazo abierto (más adelante solo política) entenderemos una sucesión de acciones admisibles* $\pi = \{a_t\}_{t=0}^{\infty}$

en A tal que a_t pertenece a $A(x_t)$ para $t = 0, 1, \ldots$. Además definimos

$$\Pi := \{\{a_t\} \in A^{\mathbb{N}_0} : a_t \in A(x_t)\}.$$

Dada una política, nos podemos fijar en el costo que genera dicha política, es decir, el costo que genera la toma de acciones en cada etapa. En este escrito vamos a tener diferentes funciones de costo como objetivo, las cuales definimos a continuación.

Tomando α en $(0,1)$, x en X y π en Π definimos:

(I) El costo $\alpha-$descontado de horizonte finito,

$$V_k^\alpha(x, \pi) := \sum_{n=0}^{k-1} \alpha^n c(x_n, a_n).$$

(II) El costo $\alpha-$descontado de horizonte infinito,

$$V^\alpha(x, \pi) := \sum_{n=0}^{\infty} \alpha^n c(x_n, a_n).$$

(III) El costo promedio de horizonte finito,

$$\overline{V}_k(x, \pi) := \frac{1}{k} \sum_{n=0}^{k-1} c(x_n, a_n).$$

(IV) El costo promedio de horizonte infinito,

$$\overline{V}(x, \pi) := \limsup_{k\to\infty} \frac{1}{k} \sum_{n=0}^{k-1} c(x_n, a_n) = \limsup_{k\to\infty} \overline{V}_k(x, a).$$

Al problema de averiguar cuando estos costos alcanzan su mínimo (o máximo) entre todas las políticas disponibles se le llama problema de control óptimo. Es conveniente denotar el mínimo de cada costo con un simbolo que lo reconozca:

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

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

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

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

Cuando obtengamos una política π en Π tal que el ínfimo se alcance, esto es, para todo x en X:

$$V_k^\alpha(x) = V_k^\alpha(x,\pi)$$
$$V^\alpha(x) = V^\alpha(x,\pi)$$
$$\overline{V}_k(x) = \overline{V}_k(x,\pi)$$
$$\overline{V}(x) = \overline{V}(x,\pi)$$

diremos que π es α-óptima en el horizonte k, α-óptima, óptima en costo promedio en el horizonte k y óptima en costo promedio, respectivamente.

Para poder obtener resultados fundamentales, necesitamos suponer dos condiciones, que en realidad no son tan restrictivas, pero ayudarán con detalles técnicos.

Suposición 1. (a) El costo por etapa $c : \mathbb{K} \to \mathbb{R}$ es no negativo.

(b) Para todo $\alpha \in (0,1)$, existe una política de control $\pi \in \Pi$ tal que

$$V^\alpha(x,\pi) < \infty, \quad \forall x \in X.$$

$\Diamond$

Gracias a (a), el criterio $V^\alpha(\cdot,\pi)$ en (II) está bien definido, pues esta serie converge o será infinito, pues acabamos de convertir a $\{V_k^\alpha(x,\pi)\}_{k\in\mathbb{N}}$ en una sucesión creciente. Sin embargo, no sabemos si la sucesión $\{\overline{V}_k(x,\pi)\}_{k\in\mathbb{N}}$ converge, por lo que el criterio (IV) tiene que estar definido por el límite superior. Las condiciones para que exista el límite en (IV), sí pueden ser muy restrictivas o poco naturales. (b) es una condición para no caer en el caso trivial en que todo es infinito.

Debido a que la "ruta"que siguen los costos y los costos promedios dependen del estado inicial y la política seleccionada, es natural que pensemos en una función medible $f : X \to A$ tal que $f(x)$ está en $A(x)$ para todo x en X. A dicha función se le llama selector o selectora. Podemos pensar que un selector f forma "automaticamente" una política $\{f(x_t)\}_{t=0}^\infty \in \Pi$ y por esta razón, haremos un abuso de la notación para evaluar a f en los costos anteriores en vez de la política $\{f(x_t)\}_{t=0}^\infty$. Denotaremos al conjunto de los selectores como $\mathbb{F} := \{f : X \to A : f(x) \in A(x), \forall x \in X\}$.

Se dice que $\pi = \{a_t\}_{t=0}^\infty$ en Π es una política de lazo cerrado o markoviana o de retroalimentación si para todo $t = 0,1,\ldots$ tenemos $a_t = f_t(x_t)$ para alguna función $f_t : X \to A$ tal que $f_t(x_t)$ está en $A(x_t)$. Se dice que es estacionaria si para todo $t = 0,1,\ldots$ existe f en $\mathbb{F}$ tal que $a_t = f(x_t)$.

Nos propondremos dos objetivos:

(1) Encontrar condiciones para que exista $f \in \mathbb{F}$ tal que

$$\overline{V}(x, f) = \overline{V}(x), \quad \forall x \in X.$$

(2) Encontrar condiciones para que se cumpla que

$$\limsup_{k\to\infty} \overline{V}_k(x) = \overline{V}(x), \quad \forall x \in X.$$

El primer objetivo es clásico en la teoría de control óptimo, mientras que el segundo es muy parecido a los teoremas ergódicos en procesos estocásticos o a la ley de grandes números en teoría de probabilidad.

1.2. Teoría preliminar.

Para obtener cualquiera de nuestros objetivos, necesitamos de ciertos conceptos bastante técnicos y necesitamos conocer resultados que aportarán contexto e información. Esta es la razón por la que escribimos está sección.

Para el resto de esta disertación, supondremos que X y A son espacios métricos.
Empezamos con las siguientes definiciones:

Definición 1.1. Bajo los mismos objetos escritos en la sección anterior:

(a) Una función $u : X \to \mathbb{R} \cup \{\infty\}$ se dice semicontinua por abajo en un punto x en X, si
$$\liminf_{n\to\infty} u(x_n) \geqslant u(x),$$
para cualquier sucesión $\{x_n\}_{n\in\mathbb{N}}$ en X que converge a x. Se dice semicontinua por abajo en X, si es semicontinua por abajo en todo punto de X.

(b) Sea $u : X \to \mathbb{R} \cup \{\infty\}$ una función. La envoltura semicontinua por abajo de u, denotada por u^e, esta definida como
$$u^e(x) := \sup_{r>0} \inf_{y\in B_r(x)} u(y)$$
donde $B_r(x)$ es la bola abierta de radio r con centro en x. Note que u^e es semicontinua por abajo y es tal que si u' es semicontinua por abajo con $u' \leqslant u$, entonces $u' \leqslant u^e$.

(c) Una función $v : \mathbb{K} \to \mathbb{R}$ se dice:

(c1) inf-compacta, si para cada r en $\mathbb{R}$, el conjunto
$$\{(x, a) \in \mathbb{K} | v(x, a) \leqslant r\}$$
es compacto.

(c2) $\mathbb{K}$-inf-compacta, si para cada conjunto compacto K subconjunto de X y cada r en $\mathbb{R}$, el conjunto

$$\{(x,a) \in G(K) | v(x,a) \leqslant r\}$$

es compacto, donde $G(K) := \{(x,a) \in K \times A | a \in A(x)\}$.

(c3) inf-compacta en $\mathbb{K}$, si para todo x en X, la función $a \mapsto v(x,a)$ es inf-compacta en $A(x)$, esto es, para cada r en $\mathbb{R}$, el conjunto

$$\{a \in A(x) | v(x,a) \leqslant r\} \subset A$$

es compacto.

$\diamondsuit$

Se pueden demostrar varias propiedades acerca de las funciones semicontinuas y de las diferentes compacidades que acabamos de definir. A continuación mencionamos algunas, pero no las demostraremos.

Proposición 1.1. *Se cumplen las siguientes propiedades:*

(i) *Siempre tenemos que* $(c1) \implies (c2) \implies (c3)$.

(ii) *Sean* $u, u_1, \ldots, u_n$ *tal que* $u, u_k : X \to \mathbb{R} \cup \{\infty\}$ *son semicontinuas por abajo para cada* $k = 1, \ldots, n$ *y* $\alpha \geqslant 0$. *Entonces* αu, $u_1 + \cdots + u_n$ *y* $\min_i u_i$ *son funciones semicontinuas por abajo.*

(iii) *Sea* $u : X \to \mathbb{R} \cup \{\infty\}$ *semicontinua por abajo. Si* X *es compacto, entonces* u *alcanza el mínimo, es decir, existe un punto* x^* *en* X *tal que* $u(x^*) = \min_{x \in X} u(x)$.

(iv) *Para* $v : \mathbb{K} \to \mathbb{R}$ *semicontinua por abajo, acotada por abajo e inf-compacta en* $\mathbb{K}$, *tenemos que existe* f^* *en* $\mathbb{F}$ *tal que* $v(x, f^*(x)) = \min_{a \in A(x)} v(x,a)$.

(v) *Si* $u, u_n : X \to \mathbb{R} \cup \{\infty\}$ *con* n *en* $\mathbb{N}$ *son acotadas por abajo e inf-compactas en* $\mathbb{K}$ *y además* $u_n \uparrow u$, *entonces*

$$\lim_{n \to \infty} \inf_{a \in A(x)} u_n(x,a) = \inf_{a \in A(x)} u(x,a).$$

(vi) *Si* $u_n : X \to \mathbb{R} \cup \{\infty\}$ *con* n *en* $\mathbb{N}$ *son funciones semicontinuas por abajo, entonces* $\sup_{i \in \mathbb{N}} u_i(\cdot)$ *es semicontinua por abajo.*

Las propiedades $(i) - (iii)$ *y* (vi) *las puede ver en [8],* (iv) *en [14], mientras que* (v) *se pueden ver en un curso de control optimo por ejemplo [7].*

Es conveniente definir la siguiente notación. Para cualquier espacio de Borel B pongamos

$$L_b(B) := \{u : B \to \mathbb{R} | u \text{ es semicontinua por abajo y acotada por abajo}\}.$$

También definamos para cada α en $(0, 1]$ y $l : X \to \mathbb{R}$ el, a veces llamado, operador α de Bellman aplicado a l como

$$K^\alpha l(x) := \inf_{a \in A(x)} \{c(x, a) + \alpha l(F(x, a))\}, \quad x \in X,$$

y el operador

$$K_f^\alpha l(x) := c(x, f(x)) + \alpha l(F(x, f(x))), \quad x \in X, f \in \mathbb{F}.$$

Por convención pondremos $Kl(x) \equiv K^1 l(x)$ y $K_f l(x) \equiv K_f^1 l(x)$.

El operador de Bellman es crucial en la teoría de control óptimo, pues es muy frecuente preguntarse bajo que condiciones V^α es solución de la ecuación de Bellman o ecuación dinámica de programación o la ecuación de $\alpha-$óptimalidad, esto es, cuando se cumple que

$$V^\alpha(x) = \inf_{a \in A(x)} \{c(x, a) + \alpha V^\alpha(F(x, a))\}, \quad \forall x \in X$$

y equivalentemente

$$V^\alpha(x) = K^\alpha V^\alpha(x), \quad \forall x \in X.$$

Queremos que nuestros objetos de estudio tengan algunas propiedades de compacidad, por lo que necesitaremos suponer algunas cosas adicionales.

Suposición 2.

(a) La función de costo por etapa $c(\cdot, \cdot)$, es semicontinua por abajo en $\mathbb{K}$.

(b) La función de transición $F : \mathbb{K} \to X$ es continua.

(c) Si $\{(x_n, a_n)\}_{n\in\mathbb{N}}$ es una sucesión en $\mathbb{K}$, x_n converge a x en X y $\{c(x_n, a_n)\}_{n\in\mathbb{N}}$ es acotada por arriba, entonces existe a en $A(x)$ tal que a es un punto de acumulación de $\{a_n\}_{n\in\mathbb{N}}$.

$\Diamond$

Para entender mejor el inciso (c) de la suposición 2, veremos el siguiente ejemplo.

Ejemplo 1.1. Pongamos $X = A = [0, 1]$ con las métricas usuales. Definamos para cada x en $X \setminus \{0\}$, $A(x) := (0, x]$ y $A(0) := \{0\}$.
Supongamos que

$$c(x, a) := x + a \qquad \text{y} \qquad F(x, a) := 1 - a.$$

Tanto c como F son funciones continuas en $\mathbb{K}$.
Note que para $x_0 := 1$ y $a_n := \frac{1}{n+2}$ para $n = 0, 1, \ldots$ tenemos que $x_n = 1 - \frac{1}{n+1}$ para n en $\mathbb{N}$, por lo que

$$a_n = \frac{1}{n+2} \in \left(0, 1 - \frac{1}{n+1}\right] = A(x_n) \quad \text{y} \quad c(x_n, a_n) = 1 - \frac{1}{n+1} + \frac{1}{n+2} \leqslant 1,$$

para cada n en $\mathbb{N}$, así que en resumen $\{(x_n, a_n)\}_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$, $x_n \to 1 \in X$ y $\{c(x_n, a_n)\}_{n\in\mathbb{N}}$ es acotada por arriba. Sin embargo el único punto de acumulación de $\{a_n\}_{n\in\mathbb{N}}$ es 0 pero $A(1) = (0, 1]$, es decir que $a = 0$ no es una acción admisible en el estado $x = 1$, por lo que perderemos acciones admisibles cuando n tienda al límite.

$\Diamond$

La suposición 2 (c) nos brinda la seguridad de seguir encontrando acciones admisibles en el límite cuando todo se comporta "localmente bien". De hecho la suposición 2 (a) y (c) están íntimamente relacionados con la compacidad en $\mathbb{K}$ (ver nota después del teorema 1.1).

Teorema 1.1. *Considere las suposiciónes 1 y 2, para $\alpha \in (0, 1]$, $u \in L_b(X)$ arbitrarios se tiene:*

(i) $c(\cdot, \cdot)$ *es inf-compacta en* $\mathbb{K}$.

(ii) $c(\cdot, \cdot) + \alpha u(F(\cdot, \cdot))$ *está en* $L_b(\mathbb{K})$ *y es inf-compacta en* $\mathbb{K}$.

(iii) $K^\alpha u$ *está en* $L_b(X)$ *y existe* f *en* $\mathbb{F}$ *tal que* $K^\alpha u = K^\alpha_f u$.

Demostración
Considerese una función $u \in L_b(X)$ arbitraria y $\alpha \in (0, 1]$.

(i) *Gracias a la suposición 2 incisos (a) y (c) se puede demostrar que $c(\cdot, \cdot)$ es $\mathbb{K}-$inf-compacta, (vea lemma 2.5 de [6] o [4]), y por la proposición 1.1-(i) tenemos $c(\cdot, \cdot)$ es inf-compacta en $\mathbb{K}$.*

(ii) *Como F es continua, es fácil probar que $u(F(\cdot, \cdot))$ está en $L_b(\mathbb{K})$. Por la proposición 1.1-(ii), tenemos que $c(\cdot, \cdot) + \alpha u(F(\cdot, \cdot))$ está en $L_b(\mathbb{K})$. Por esto último y por que se puede verificar una propiedad similar a la de la suposición 2-(b), al igual en (i), podemos verificar que $c(\cdot, \cdot) + \alpha u(F(\cdot, \cdot))$ es inf-compacta en $\mathbb{K}$.*

(iii) *Por el inciso anterior y la proposición 1.1-(ii) tenemos que $K^\alpha u \in L_b(X)$. De nuevo, por el inciso anterior y por la proposición 1.1-(iv) tenemos que existe f en $\mathbb{F}$ tal que $K^\alpha u = K^\alpha_f u$.*

$\blacksquare$

NOTA. Los incisos (a) y (c) de la suposición 2 son equivalentes a que $c(\cdot, \cdot)$ sea $\mathbb{K}-$inf-compacta (ver lemma 2.5 de [6]). Pero esto sucede en el caso donde X y A son espacios métricos. Para ver generalizaciones sobre los espacios X y A

se recomienda ver [4]. La razón por la que se opta escribir (a) y (c) en vez de que $c(\cdot,\cdot)$ sea $\mathbb{K}-$inf-compacta es por que los incisos (a) y (c) son más fáciles de comprobar dada la función $c(\cdot,\cdot)$ $\diamond$

El siguiente teorema es un compilado de resultados conocidos en la teoría de control óptimo descontado $(i)-(iii)$ y costos promedios (iv), por lo que no presentaremos una prueba. Para ver las pruebas puede consultar [7].

Teorema 1.2. *Considere las suposicióńes 1 y 2. Sea $\alpha \in (0,1)$. Entonces*

(i) $V^\alpha(\cdot)$ *está en* $L_b(X)$ *y* $V^\alpha(\cdot) = K^\alpha V^\alpha(\cdot)$.

(ii) *Existe* f *en* $\mathbb{F}$ *tal que* $V^\alpha(\cdot) = K_f^\alpha V^\alpha(\cdot)$. *Además* f *es* α*-óptima descontada, esto es,* $V^\alpha(\cdot) = V^\alpha(\cdot, f)$.

(iii) $V_k^\alpha(\cdot) \uparrow V^\alpha(\cdot)$.

Sin considerar la suposición 2.

(iv) *Sean* v^* *en* $\mathbb{R}$ *y* $l : X \to \mathbb{R}$. *Suponga que para alguna* f *en* $\mathbb{F}$ *tenemos*

$$v^* + l(x) \geqslant c(x, f(x)) + l(F(x, f(x))), \quad \forall x \in X.$$

Si $\lim_{t\to\infty} l(x_t^f)/t$ *existe y es no negativo, entonces* $v^* \geqslant \overline{V}(x, f(x))$ *para todo* x *en* X.

Capítulo 2

Costo promedio determinista.

2.1. La aproximación $\alpha-$descontada.

La aproximación $\alpha-$descontada es utilizada para estudiar el costo promedio a largo plazo, más comunmente se encuentra para procesos Markovianos como en [8] o [16] y tiene la peculiaridad de desarrollarse en términos de objetos auxiliares, así que ahora definiremos los últimos objetos para alcanzar nuestras metas.

Definamos para cada α en $(0,1)$,

$$m_\alpha := \inf_{x\in X} V^\alpha(x), \quad y \quad h_\alpha : X \to \mathbb{R} \quad h_\alpha(\cdot) := V^\alpha(\cdot) - m_\alpha.$$

Aquí m_α se puede interpretar como el valor más pequeño que pudiera llegar a tomar $V^\alpha(\cdot)$ y pensando en la siguiente desigualdad $0 \leqslant m_\alpha \leqslant V^\alpha(\cdot)$ notamos que $h_\alpha(\cdot)$ es la distancia entre el costo del estado actual y el costo más pequeño que se puede llegar a tomar. En particular, considerando la suposición 1, tenemos $0 \leqslant m_\alpha < \infty$. Considerando la suposición 1 y 2, por el teorema 1.2-(i) tenemos que h_α esta en $L_b(X)$.

Para poder trabajar, necesitamos de las siguientes hipótesis.

Suposición 3.

(a) La constante $V^* := \inf_{x\in X} \overline{V}(x) < \infty$.

(b) La función $h(\cdot) := \liminf_{\alpha\uparrow 1} h_\alpha(\cdot)$ es finita.

$\Diamond$

Con (a), evitamos el caso trivial en donde $\overline{V}(x) = \infty$, para toda x en X. Más adelante servirá para verificar la finitud de una constante.
Con (b) aseguramos que los h_α's seran finitos en el límite inferior cuando α tiende crecientemente a 1, lo que nos ayuda en al menos en una vecindad de 1, pero es menos restrictivo que pedir que el límite exista y que sea finito.

Ahora pongamos $H_\alpha(\cdot) := \inf_{\beta\in[\alpha,1)} h_\beta(\cdot)$, para cada α en $(0,1)$. Note que dicha colección de funciones es creciente respecto a su índice; por lo que

$$0 \leqslant h(\cdot) = \lim_{\alpha\uparrow 1} H_\alpha(\cdot) = \sup_{\alpha\in(0,1)} H_\alpha(\cdot).$$

Por el momento solo reescribimos a la función $h(\cdot)$. Al tomar $H_\alpha^e(\cdot)$, la envoltura semicontinua por abajo de $H_\alpha(\cdot)$ para cada α en $(0,1)$, note que tambien tenemos una colección de funciones crecientes, pues para $\alpha < \beta$, ambos en $(0,1)$, tenemos

$$H_\alpha^e(\cdot) \leqslant H_\alpha(\cdot) \leqslant H_\beta(\cdot),$$

con $H_\alpha^e(\cdot)$ semicontinua por abajo; así que por definición de envoltura semicontinua por abajo

$$H_\alpha^e(\cdot) \leqslant H_\beta^e(\cdot).$$

Por lo tanto, $\{H_\alpha^e(\cdot)\}_{\alpha\in(0,1)}$ es creciente y además no negativa (0 es una función semicontinua por abajo menor o igual que $H_\alpha(\cdot)$) Así que el límite cuando $\alpha \uparrow 1$ estará bien definido. Tomemos

$$\rho^* := \limsup_{\alpha\uparrow 1} (1-\alpha) m_\alpha, \qquad h^*(\cdot) := \lim_{\alpha\uparrow 1} H_\alpha^e(\cdot) = \sup_{\alpha\in(0,1)} H_\alpha^e(\cdot).$$

NOTA. Considerando la suposición 3, como $h(\cdot)$ es finita, entonces $h^*(\cdot)$ es finita. Además ρ^* es finita, pues por algunos teoremas abelianos (ver colorario 4.1 en el apéndice) tenemos la siguiente desigualdad

$$\limsup_{\alpha\uparrow 1} (1-\alpha) V^\alpha(x) \leqslant \overline{V}(x), \qquad \forall x \in X,$$

y por definición de m_α tenemos

$$\rho^* = \limsup_{\alpha\uparrow 1} (1-\alpha) m_\alpha \leqslant \limsup_{\alpha\uparrow 1} (1-\alpha) V^\alpha(x), \qquad \forall x \in X.$$

Uniendo ambos hechos

$$\rho^* \leqslant \overline{V}(x) \qquad \forall x \in X.$$

Luego

$$\rho^* \leqslant V^* < \infty.$$

Ahora considere la suposición 1. Entonces existe $\{\alpha_n\}_{n\in\mathbb{N}}$ contenida en $(0,1)$ tal que $\alpha_n \uparrow 1$ y además $\rho^* = \lim_{n\to\infty}(1-\alpha_n) m_{\alpha_n}$. En efecto, por la suposición 1 (b), para todo α en $(0,1)$ existe π en Π tal que

$$V^\alpha(x,\pi) < \infty, \quad \forall x \in X.$$

Luego para α en $(0,1)$ y cualquier x en X

$$(1-\alpha)m_\alpha \leqslant m_\alpha \leqslant V^\alpha(x) < \infty.$$

Por lo tanto

$$\{(1-\alpha)m_\alpha\}_{\alpha\in(0,1)} \subset \mathbb{R}.$$

Como cada sucesión en $\mathbb{R}$ tiene una subsucesión monotona, obtenemos lo deseado. $\diamond$

Es momento de probar uno de nuestros resultados principales. La demostración sigue las ideas de [16].

Teorema 2.1. *Considere la suposición 1, 2 y 3. Entonces existe f^* en $\mathbb{F}$ tal que*

$$\rho^* + h^*(\cdot) \geqslant K_{f^*}h^*(\cdot).$$

Además f^* es costo promedio-óptima, esto es, $\overline{V}(\cdot) = \overline{V}(\cdot, f^*)$.

Demostración
Sea α en $(0,1)$. Por el teorema 1.2-(i) tenemos $V^\alpha(\cdot) = K^\alpha V^\alpha(\cdot)$ y como $h_\alpha(\cdot)$ está en $L_b(X)$, entonces

$$\begin{aligned} K^\alpha h_\alpha(\cdot) &= K^\alpha(V^\alpha(\cdot) - m_\alpha) = K^\alpha V^\alpha(\cdot) - \alpha m_\alpha \\ &= V^\alpha(\cdot) - \alpha m_\alpha \\ &= h_\alpha(\cdot) + (1-\alpha)m_\alpha. \end{aligned}$$

Dado que $H^e_\alpha(\cdot) \leqslant H_\alpha(\cdot) \leqslant h_\beta(\cdot)$ para α en $(0,1)$ y β en $[\alpha,1)$, entonces por la definición de K^α tenemos $K^\alpha H^e_\alpha(\cdot) \leqslant K^\alpha h_\beta(\cdot)$, luego

$$K^\alpha H^e_\alpha(\cdot) \leqslant h_\beta(\cdot) + (1-\beta)m_\beta \qquad \forall\alpha\in(0,1), \beta\in[\alpha,1). \qquad (1)$$

Por otro lado, por definición de ρ^*, para $\epsilon > 0$ arbitrario, existe α_0 en $(0,1)$ tal que

$$\epsilon + \rho^* \geqslant (1-\beta)m_\beta, \qquad \forall\beta\in(\alpha_0,1). \qquad (2)$$

Combinando (1) y (2) obtenemos

$$K^\alpha H^e_\alpha(\cdot) \leqslant h_\beta(\cdot) + \epsilon + \rho^*, \qquad \forall\alpha\in(\alpha_0,1), \beta\in[\alpha,1),$$

luego

$$K^\alpha H^e_\alpha(\cdot) \leqslant H_\alpha(\cdot) + \epsilon + \rho^*, \qquad \forall\alpha\in(\alpha_0,1).$$

Debido a que $H^e_\alpha(\cdot)$ es semicontinua por abajo, por el teorema 1.1-(iii), tenemos que $K^\alpha H^e_\alpha(\cdot)$ es semicontinua por abajo y por definición de envoltura semicontinua por abajo

$$K^\alpha H^e_\alpha(\cdot) \leqslant (H_\alpha(\cdot) + \epsilon + \rho^*)^e \quad \forall\alpha\in(\alpha_0,1).$$

Como ϵ y ρ^* son constantes, tendremos que $(H_\alpha(\cdot)+\epsilon+\rho^*)^e = H_\alpha^e(\cdot)+\epsilon+\rho^*$ *y por lo tanto*

$$K_\alpha H_\alpha^e(\cdot) \leqslant H_\alpha^e(\cdot)+\epsilon+\rho^* \quad \forall \alpha \in (\alpha_0, 1) \qquad (3).$$

Ahora tomemos $\alpha_n \uparrow 1$ sucesión arbitraria y definamos las siguientes funciones de $\mathbb{K}$ en $\mathbb{R}$,

$$v_n(\cdot,\cdot) := c(\cdot,\cdot)+\alpha_n H_{\alpha_n}^e(F(\cdot,\cdot)), \quad \forall n \in \mathbb{N}, \qquad v(\cdot,\cdot) := c(\cdot,\cdot)+h^*(F(\cdot,\cdot)).$$

Note que:

1. *Por los cálculos hechos antes del teorema $H_{\alpha_n}^e \uparrow h^*$, entonces $v_n \uparrow v$.*
2. *Por el teorema 1.1-(ii), $v_n(\cdot,\cdot)$ y $v(\cdot,\cdot)$ son semicontinuas por abajo e inf-compactas en $\mathbb{K}$.*

Así que por la proposición 1.1-(v), tenemos

$$\begin{aligned} &\lim_{n\to\infty} \inf_{a\in A(x)} v_n(x,a) = \inf_{a\in A(x)} v(x,a) \\ \Rightarrow &\lim_{n\to\infty} K^{\alpha_n}(H_{\alpha_n}^e(\cdot)) = Kh^*(\cdot). \end{aligned}$$

Por esto último, tomando el límite cuando $\alpha \uparrow 1$ en (3), tenemos

$$Kh^*(\cdot) \leqslant \lim_{\alpha\uparrow 1} H_\alpha^e(\cdot)+\epsilon+\rho^* = h^*(\cdot)+\epsilon+\rho^*.$$

Como $\epsilon > 0$ fue arbitrario,

$$K(h^*(\cdot)) \leqslant h^*(\cdot)+\rho^*.$$

Por la proposición 1.1-(vi), sabemos que h^* es semicontinua por abajo, luego por el teorema 1.1-(iii) sabemos que existe f^* en $\mathbb{F}$ tal que $Kh^*(\cdot) = K_{f^*}h^*(\cdot)$. Finalmente

$$\rho^*+h^*(\cdot) \geqslant K_{f^*}h^*(\cdot). \qquad (4)$$

Por algunos cálculos hechos antes sabemos que

$$\rho^* \leqslant \overline{V}(x) \quad \forall x \in X.$$

Al tener (4) y $h^*(\cdot) \geqslant 0$ (que es límite de funciones no negativas), por el teorema 1.2-(iv) se tiene

$$\rho^* \geqslant \overline{V}(x, f^*(x)), \quad \forall x \in X.$$

Con ambas desigualdades llegamos a $\overline{V}(x,f^*(x)) \leqslant \overline{V}(x)$ para todo x en X, y siempre tenemos $\overline{V}(x,f(x)) \geqslant \overline{V}(x)$ para todo x en X y toda f en $\mathbb{F}$. Por lo tanto

$$\overline{V}(\cdot) = \overline{V}(\cdot, f^*).$$

■

Corolario 2.1. Bajo las hipótesis del teorema anterior

$$\rho^* = \overline{V}(\cdot).$$

Demostración
Del teorema anterior $\overline{V}(x, f^*(x)) \leqslant \rho^* \leqslant \overline{V}(x)$, *para todo* x *en* X. *Pero por la conclusión se tiene que* $\overline{V}(\cdot, f^*) = \overline{V}(\cdot)$. ■

Corolario 2.2. Bajo las hipótesis del teorema anterior

$$\rho^* + h^*(\cdot) = Kh^*(\cdot).$$

Demostración
Sea $\alpha_n \uparrow 1$ *una susesión. Sea* n *en* $\mathbb{N}$. *Note que por el teorema* 1.2-(i)

$$(1-\beta)m_\beta + h_\beta(\cdot) = K^\beta V^\beta(\cdot) - \beta m_\beta = K^\beta h_\beta(\cdot), \quad \forall \beta \in [\alpha_n, 1)$$

lo que implica que

$$\begin{aligned}
\inf_{\beta\in[\alpha_n,1)} (1-\beta)m_\beta + H^e_{\alpha_n}(\cdot) &\leqslant (1-\beta)m_\beta + h_\beta(\cdot) \\
&\leqslant c(\cdot, f) + \beta h_\beta(F(\cdot, f)) \\
&\leqslant c(\cdot, f) + h_\beta(F(\cdot, f)), \quad \forall f \in \mathbb{F}, \beta \in [\alpha_n, 1).
\end{aligned}$$

Como $\inf_{\beta\in[\alpha_n,1)}(1-\beta)m_\beta + H^e_{\alpha_n}(\cdot) - c(\cdot, f)$ *está por debajo de* $h_\beta(F(\cdot, f))$ *para toda* β *en* $[\alpha_n, 1)$, *entonces*

$$\inf_{\beta\in[\alpha_n,1)} (1-\beta)m_\beta + H^e_{\alpha_n}(\cdot) - c(\cdot, f) \leqslant H_{\alpha_n}(F(\cdot, f)),$$

pero el lado izquierdo es una función semicontinua por abajo (proposición 1.1*), por lo que*

$$\inf_{\beta\in[\alpha_n,1)} (1-\beta)m_\beta + H^e_{\alpha_n}(\cdot) - c(\cdot, f) \leqslant H^e_{\alpha_n}(F(\cdot, f)).$$

Por la nota antes del teorema 2.1, *sabemos que existe* $\{\alpha_n\}_{n\in\mathbb{N}}$ *contenida en* $(0, 1)$ *tal que* $\alpha_n \uparrow 1$ *y* $\rho^* = \lim_{n\to\infty}(1-\alpha_n)m_{\alpha_n}$. *Para esta sucesión, tomando el límite cuando* $n \to \infty$ *en la última desigualdad obtenemos*

$$\liminf_{\alpha_n\uparrow 1} (1-\alpha_n)m_{\alpha_n} + h^*(\cdot) \leqslant c(\cdot, f) + h^*(F(\cdot, f)),$$

y luego

$$\rho^* + h^*(\cdot) \leqslant c(\cdot, f) + h^*(F(\cdot, f)) \qquad \forall f \in \mathbb{F}.$$

Así que

$$\rho^* + h^*(\cdot) \leqslant Kh^*(\cdot).$$

Por el teorema anterior tenemos la otra desigualdad.

■

2.2. Un teorema ergódico.

Ahora toca probar el segundo resultado principal. Afortunadamente ya está todo preparado gracias a la sección anterior.

La idea de la demostración es tomada del teorema principal en [15].

Para cada $t = 1, 2, \ldots$ *considere* $V_t(\cdot, \cdot)$ *el costo de horizonte finito dado por* $V_t(x, \pi) := t\, \overline{V}_t(x, \pi) = \sum_{n=0}^{t-1} c(x_n, a_n)$ *y definimos* $V_t(x) := \inf_{\pi \in \Pi} V_t(x, \pi)$.

Teorema 2.2. *Suponga las hipótesis 1, 2 y 3. Además suponga que* $\lim_{t\to\infty} \frac{1}{t} h^*(x_t^{f_t})$ *existe y es* 0 *para cada* f_t *en* $\mathbb{F}$ *tal que* $V_t(\cdot, f_t) = V_t(\cdot)$*. Entonces*

$$\lim_{k\to\infty} \overline{V}_k(x) = \overline{V}(x), \quad \forall x \in X.$$

Demostración
Sea $x = x_0$ *en* X *arbitrario. Para cada* f *en* $\mathbb{F}$ *podemos formar* $\{x_k^f\}_{k=0}^{\infty}$ *una sucesión en* X *tal que* $x_k^f = F(x_{k-1}, f(x_{k-1}))$ *para todo* $k = 1, 2, \ldots$*, y por el corolario anterior tenemos que*

$$\rho^* + h^*(x_k) \leqslant c(x_{k+1}, f(x_k)) + h^*(x_{k+1}).$$

Sumando las primeras n *ecuaciones que se obtienen para* $k = 0, \ldots, n-1$ *obtenemos*

$$n\rho^* + h^*(x_0) \leqslant \sum_{k=0}^{n-1} c(x_{k+1}, f(x_k)) + h^*(x_n) = V_n(x, f) + h^*(x_n).$$

Tomando f_n *en* $\mathbb{F}$ *tal que* $V_n(x, f_n) = V_n(x)$ *(similar al teorema 1.2-(ii)) obtenemos*

$$n\rho^* + h^*(x_0) \leqslant V_n(x) + h^*(x_n^{f_n}).$$

Por otro lado, como $\lim_{t\to\infty} \frac{1}{t} h^*(x_t^{f_t}) = 0$*, entonces para* $\epsilon > 0$ *existe* n_0 *natural, tal que*

$$\frac{1}{n} h^*(x_n^{f_n}) < \epsilon \qquad \forall n \geqslant n_0,$$

luego tenemos

$$\rho^* \leqslant \rho^* + \frac{1}{n} h^*(x_0) < \frac{1}{n} V_n(x) + \epsilon, \qquad \forall n \geqslant n_0. \qquad (1)$$

Por otro lado, podemos obtener f^* *como en el teorema 2.1 y proceder de manera análoga para obtener*

$$\begin{aligned} n\rho^* + h^*(x_0) &\geqslant \sum_{k=0}^{n-1} c(x_{k+1}, f^*(x_k)) + h^*(x_n^{f^*}) \\ &\geqslant V_n(x) + h^*(x_n^{f^*}) \\ &\geqslant V_n(x) \\ \Rightarrow \rho^* + \frac{1}{n} h^*(x_0) &\geqslant \frac{1}{n} V_n(x). \end{aligned}$$

La penúltima desigualdad se debe a que $h^*(\cdot)$ *es siempre positiva. Como* $\frac{1}{n}h^*(x_0) \to 0$ *cuando* $n \to \infty$, *pues* $h^*(\cdot) \leqslant h(\cdot)$ *y* $h(\cdot)$ *es finita por la suposición 3, entonces existe* n_1 *natural tal que si* $n \geqslant n_1$ *se tiene* $\frac{1}{n}h^*(x_0) < \epsilon$, *así*

$$\rho^* + \epsilon > \frac{1}{n}V_n(x), \qquad \forall n \geqslant n_1. \qquad (2)$$

Finalmente, de (1) *y* (2), *para* $n \geqslant \max\{n_0, n_1\} =: k$, *obtenemos*

$$-\epsilon < \rho^* - \frac{1}{n}V_n(x) < \epsilon,$$

es decir que para $\epsilon > 0$ *existe* k *tal que si* $n \geqslant k$, *entonces* $|\rho^* - \frac{1}{n}V_n(x)| < \epsilon$. *Por lo tanto*

$$\lim_{k\to\infty} \overline{V}_k(x) = \overline{V}(x), \quad \forall x \in X.$$

$\blacksquare$

Capítulo 3

Ejemplos

Este capítulo está dedicado a ejemplos que ilustren la aproximación $\alpha-$descontada y el teorema ergódico. Pasaremos primero por ejemplos sencillos con un contexto muy pobre (y que se deja a la imaginación del lector) y después por otros ejemplos que tienen aplicación y un contexto muy específico (aunque fácilmente adaptable a otros contextos).

En cada ejemplo se dará la referencia de donde fue "tomado". Sin embargo en cada referencia se presenta una versión no determinista, es decir, dependerán también de algunas variables aleatorias.

3.1. Ejemplos sencillos.

El siguiente ejemplo se puede encontrar en [9] capítulo 10.

Ejemplo 3.1. Sea $X = \mathbb{Z}^+$. Pongamos $A = A(x) := \{1, 2\}$, para toda x en X. Definamos $F : \mathbb{K} \to X$ como

$$F(0, a) := 0, \quad F(x, 1) := 0, \quad F(x, 2) := x + 1, \quad \forall a \in A, \forall x \in X \backslash \{0\}.$$

Definamos el costo por etapa $c : \mathbb{K} \to \mathbb{R}$ como

$$c(0, a) := 0, \quad c(x, 1) := \frac{1}{x} + 1, \quad c(x, 2) := 0, \quad \forall a \in A, \forall x \in X \backslash \{0\}.$$

En este sistema, podemos verificar las suposciones 1, 2 y 3. Brevemente las mencionamos. La función c es positiva y continua, la función F es continua, para $\pi := \{1, 0, \dots\} \in \Pi$ tenemos que $V^\alpha(x, \pi) < \infty$, para todo x en X. Toda sucesión en $A := \{1, 2\}$ debe tener algún punto de acumulación en A, así que se cumple el inciso (c) de la suposición 3. Como $\overline{V}(x) \leqslant 2$, entonces $V^* = \inf_{x \in X} \overline{V}(x) < \infty$. Note que $V^\alpha(x)$ es acotada por arriba y por abajo, luego $h_\alpha(x) = V^\alpha(x) - m_\alpha < \infty$, para toda x en X. Así que se cumplen las suposciones 1, 2 y 3. Luego podemos aplicar los teoremas 2.1 y 2.2.

Por un lado, si ponemos $f(x) = 1$ para todo x en X, tenemos

$$V_n(0, f) = 0, \quad V_n(x, f) = \frac{1}{x} + 1, \quad \forall x \geqslant 1, n \geqslant 1.$$

Por otro, no es muy difícil ver que

$$V_n(0) = 0, \quad V_n(x) = \frac{1}{x + n - 1} + 1, \quad \forall x \geqslant 1, n \geqslant 1.$$

Luego

$$\lim_{n \to \infty} \frac{1}{n} [V_n(f, x) - V_n(x)] = \lim_{n \to \infty} \frac{1}{n} \left[\frac{1}{x} + 1 - \frac{1}{x + n - 1} - 1 \right] = 0. \quad (*)$$

Note que

$$\overline{V}(x) \leqslant \overline{V}(x, f) = \limsup \frac{1}{n} V_n(x, f)$$

y que para π en Π

$$\begin{aligned} \frac{1}{n} V_n(x) \leqslant \frac{1}{n} V_n(x, \pi) &\implies \limsup_{n \to \infty} \frac{1}{n} V_n(x) \leqslant \limsup_{n \to \infty} \frac{1}{n} V_n(x, \pi) = \overline{V}(x, \pi) \\ &\implies -\overline{V}(x) \leqslant -\limsup_{n \to \infty} \frac{1}{n} V_n(x). \end{aligned}$$

Entonces

$$0 \leqslant \overline{V}(x, f) - \overline{V}(x) \leqslant \lim_{n \to \infty} \frac{1}{n} [V_n(f, x) - V_n(x)] = 0.$$

Por lo tanto $\overline{V}(x, f) = \overline{V}(x)$ para toda x en X. Es decir, f es óptima en costo promedio.
Por el teorema 2.2, tenemos

$$\overline{V}(x) = \lim_{n \to \infty} \frac{1}{n} V_n(x) = \lim_{n \to \infty} \frac{1}{n} \left[\frac{1}{x + n - 1} + 1 \right] = 0.$$

$\diamondsuit$

NOTA. Cuando una política cumple $(*)$, se dice que es óptima en costo promedio F-fuerte (o óptima en costo promedio fuerte en el sentido de Flynn). En general, F-fuerte implica óptimo en costo promedio y la demostración es muy similar a lo que presentamos en el ejemplo anterior. $\diamondsuit$

El siguiente ejemplo se puede encontrar en [8] capítulo 5.

Ejemplo 3.2. Sea $X := \{1, 2, 3\}$ y $A := \{1\}$. Definamos el costo por etapa como $c(x, 1) := c_x$ con $0 < c_1 < c_2, c_3$ y la función de transición por $F(x, 1) := x$ si $x = 1, 2$ y $F(3, 1) := 1$.
Entonces

$$V^\alpha(1) = V^\alpha(1, 1) = \sum_{k=0}^{\infty} \alpha^k c(1, 1) = \frac{c_1}{1 - \alpha},$$

y análogamente $V^{\alpha}(2) = c_2/(1-\alpha)$ y $V^{\alpha}(3) = c_3 + \alpha c_1/(1-\alpha)$.

Ahora $m_{\alpha} = \operatorname{ínf}_{x\in X} V^{\alpha}(x) = c_1/(1-\alpha)$, por lo que

$$h_{\alpha}(x) = V^{\alpha}(x) - m_{\alpha} = \begin{cases} 0 & \text{si } x = 1, \\ \frac{c_2 - c_1}{1-\alpha} & \text{si } x = 2, \\ c_3 - c_1 & \text{si } x = 3. \end{cases}$$

Luego

$$h(x) = \limsup_{\alpha\uparrow 1} h_{\alpha}(x) = \begin{cases} 0 & \text{si } x = 1, \\ +\infty & \text{si } x = 2, \\ c_3 - c_1 & \text{si } x = 3. \end{cases}$$

Así que no se cumple la suposición 3 (b).
Por otro lado, al poner $\rho^* := \limsup_{\alpha\uparrow 1}(1-\alpha)m_{\alpha} = c_1$ y f en $\mathbb{F}$ como $f(x) := 1$, entonces

$$\rho^* + h(x) \geqslant K_f(h)(x), \quad \forall x \in X;$$

esto es, se cumple una conclusión del teorema 2.1 (De hecho se cumple la igualdad, aunque para $x = 2$ la igualdad es $\infty = \infty$). Además note que

$$V_n(x) = V_n(x, f) = \sum_{k=0}^{n-1} c(x, f(x)) = nc_x,$$

y luego $\overline{V}(x) = \overline{V}(x, f) = \limsup_{n\to\infty} \frac{1}{n} V_n(x, f) = c_x$. Así que se cumple el resultado del teorema 2.2

$$\overline{V}(x) = \lim_{n\to\infty} \frac{1}{n} V_n(x).$$

Sin embargo, NO se cumple que $\rho^* = \overline{V}(x)$ para todo x en X.

$\Diamond$

3.2. Sistema LQ.

En esta parte mostraremos el sistema LQ en su forma más sencilla. Las siglas LQ provienen del ingles "linear quadratic" y se deben a que la función de transición del sistema adquiere una forma lineal respecto a sus argumentos mientras que el costo por etapa tiene una forma cuadrática. Las aplicaciones para este sistema son extensas pero en particular los problemas de seguimiento que son típicos en economía e ingeniería son descritos através de sistemas LQ. El problema de seguimiento lo trataremos más adelante, después de introducir el sistema LQ.

Ejemplo 3.3. Sistema LQ.
Sea $X = A = \mathbb{R}$. Definamos

$$F(x, a) := \gamma x + \beta a, \quad t = 0, 1, 2, \ldots$$

con γ, β en $\mathbb{R}$ distintas de cero. La función objetivo o la función a optimizar es

$$V_T(x,\pi) = \sum_{t=0}^{T-1} (qx_t^2 + ra_t^2),$$

con $q, r \geqslant 0$.
Note que:

(a) $c : X \times A \to \mathbb{R}$ definida como $c(x,a) := qx^2 + ra^2$, es positiva y continua, por lo tanto, tambien es semicontinua por abajo.

(b) Sea $x \in X$ y $\pi := \{0, 0, \ldots\}$. Entonces

$$V(x,\pi) := \sum_{t=0}^{\infty} qx_t^2 = \sum_{t=0}^{\infty} q(\gamma^t x)^2 = qx^2 \sum_{t=0}^{\infty} (\gamma^2)^t,$$

suponiendo que $\gamma^2 < 1$, tenemos que $V(x,\pi) < \infty$ para toda x en X y luego para toda α en $(0,1)$ tenemos

$$V^\alpha(x,\pi) < \infty, \quad \forall x \in X.$$

(c) $F : \mathbb{K} \to X$ es continua.

(d) Suponga que $x \mapsto A(x)$ esta definida. Sea $\{(x_n, a_n)\}_{n\in\mathbb{N}}$ una sucesión en $\mathbb{K}^{\mathbb{N}}$ tal que x_n converge a x en X y $\{c(x_n, a_n)\}_{n\in\mathbb{N}}$ es acotada por arriba. Por la continuidad de c, existe M en $\mathbb{R}$ tal que $0 \leqslant qx_t^2 + ra_t^2 \leqslant M$ para toda t en $\mathbb{N}$. Luego $-\sqrt{M/r} \leqslant a_t \leqslant \sqrt{M/r}$. Así que $\{a_t\}_{t=0}^{\infty}$ es acotada en $\mathbb{R}$ y por el teorema de Bolzano-Weierstrass existe una subsucesión $\{a_{n_k}\}_{k\in\mathbb{N}}$ convergente a algún a en A. Luego $a \in \overline{\{a_{n_k}\}}_{k\in\mathbb{N}} \subset \overline{\{a_n\}}_{n\in\mathbb{N}}$. Suponga que siempre tendremos a en $A(x)$ (por ejemplo $A(x) = \mathbb{R}$).

(e) $V^* = \operatorname{ínf}_{x\in X} \overline{V}(x) \leqslant \overline{V}(0) \leqslant \overline{V}(0,\pi)$, donde
$\overline{V}(0,\pi) = \operatorname{lím\,sup}_{k\to\infty} \frac{1}{k} \sum_{n=0}^{k} c(0,0) = 0$, por lo tanto $V^* \leqslant 0 < \infty$.

(f) Es fácil notar que $V^\alpha(x) \geqslant 0$ para todo x en X. Por otro lado $V^\alpha(0,\pi) = 0$ lo que implica que $m_\alpha = \operatorname{ínf}_{x\in X} V^\alpha(x) = 0$. Por lo tanto $h_\alpha(\cdot) = V^\alpha(\cdot) - m_\alpha = V^\alpha(\cdot)$. Así que

$$h(\cdot) = \liminf_{\alpha\uparrow 1} h_\alpha(\cdot) = \liminf_{\alpha\uparrow 1} V^\alpha(\cdot) \leqslant \liminf_{\alpha\uparrow 1} V^\alpha(\cdot,\pi),$$

luego para toda x en X

$$h(x) \leqslant \liminf_{\alpha\uparrow 1} V^\alpha(x,\pi) = \liminf_{\alpha\uparrow 1} \sum_{t=0}^{\infty} \alpha^t q(x_t)^2 = \liminf_{\alpha\uparrow 1} qx^2 \sum_{t=0}^{\infty} (\alpha\gamma^2)^t,$$

como suponemos que $\gamma^2 < 1$ y $\alpha < 1$, entonces

$$h(x) \leqslant \liminf_{\alpha\uparrow 1} \frac{qx^2}{1-\alpha\gamma^2} = \frac{qx^2}{1-\gamma^2} < \infty. \qquad (*)$$

Por tanto $h(\cdot)$ es finita.

Con estos incisos verificamos que se cumple la suposición 1, 2 y 3. Para conseguir que este sistema verifique el teorema 2.2, necesitamos que $\frac{1}{t}h^*(x_t^{f_t}) \to 0$, para cada f_t en $\mathbb{F}$ tal que $V_t(\cdot, f_t) = V_t(\cdot)$, es decir, f_t genera una política óptima para el horizonte t. Para probarlo necesitaremos apoyarnos en un resultado que calcula la política óptima que se conoce como el teorema de programación dinámica (ver Apendice 4.2).

Para nuestro modelo LQ de horizonte finito T, definamos

$$F(x,a) := \alpha x + \beta a \qquad y \qquad V_T(x,\pi) := \sum_{k=0}^{T-1}(qx_k^2 + ra_k^2) + q_T x_T,$$

donde q_T es un coeficiente positivo que se interpreta como un costo final que en nuestro caso $q_T = 0$; las ecuaciones de programación dínamica, al resolverlas quedan como:

$$J_T(x) := 0, \quad J_s(x) := K_s x^2, \quad s = t-1, \ldots, 0,$$

con

$$K_T := q_T, \quad K_s := (r + K_{s+1}\beta^2)^{-1} K_{s+1} r\alpha^2 + q, \quad s = t-1, \ldots, 0,$$

y por el teorema de programación dinámica se cumple que $J_0(x) = V_T(x) = \text{ínf}_{\pi\in\Pi} V_T(x,\pi)$. Además

$$a_s^*(x) := G_s x, \text{ donde } G_s := -(r + K_{s+1}\beta^2)^{-1} K_{s+1}\alpha\beta, \quad s = T-1, \ldots, 0$$

es tal que $f_T(x) = \pi^*(x) := (a_0^*(x), a_1^*(x), \ldots, a_{T-1}^*(x))$ es una política óptima.

Suponga que $K_{s+1} \geqslant 0$ para algún s en $\{T-2, \ldots, 0\}$. Entonces, por definición $K_s > 0$ pues $r, q > 0$ y además

$$\begin{aligned} x_{s+1}^{f_T} = x_{s+1} &= F(x_s, f_T(x_s)) = \alpha x_s + \beta a_s^*(x_s) \\ &= \alpha x_s - (r + K_{s+1}\beta^2)^{-1} K_{s+1}\alpha\beta^2 x_s \\ &= \alpha x_s[1 - (r + K_{s+1}\beta^2)^{-1} K_{s+1}\beta^2]. \end{aligned}$$

Como $r > 0$ y $K_{s+1} \geqslant 0$ entonces $r + k_{s+1}\beta^2 > k_{s+1}\beta^2 \geqslant 0$ lo que implica $1 > (r + K_{s+1}\beta^2)^{-1} K_{s+1}\beta^2 \geqslant 0$ y finalmente

$$0 < 1 - (r + K_{s+1}\beta^2)^{-1} K_{s+1}\beta^2 \leqslant 1,$$

así que en resumen:

$$\text{Sí } K_{s+1} \geqslant 0 \text{ entonces } K_s \geqslant 0 \text{ y } |x_{s+1}^{f_T}| \leqslant |\alpha x_s| = |\alpha x_s^{f_T}|.$$

Como $K_T = 0$ entonces $K_{T-1} = q > 0$ y aplicando recursivamente el argumento anterior obtenemos

$$|x_T^{f_T}| \leqslant |\alpha^T x_0^{f_t}| = |\alpha^T x|, \quad \forall T \in \mathbb{N}.$$

Usando $(*)$ y lo anterior

$$\lim_{t\to\infty}\frac{1}{t}h^*(x_t^{f_t}) \leqslant \lim_{t\to\infty}\frac{1}{t}\frac{q(x_t^{f_t})^2}{1-\gamma^2} \leqslant \lim_{t\to\infty}\frac{1}{t}\frac{q(\gamma^t x)^2}{1-\gamma^2} = \frac{qx^2}{1-\gamma^2}\lim_{t\to\infty}\frac{(\gamma^2)^t}{t},$$

y como $\alpha^2 < 1$, finalmente obtenemos

$$\lim_{t\to\infty}\frac{1}{t}h^*(x_t^{f_t}) = 0.$$

Usando el teorema 2.2, obtenemos que

$$\lim_{k\to\infty}\overline{V}_k(x) = \overline{V}(x), \quad \forall x \in X,$$

lo que resulta muy útil, pues $\overline{V}_k(x)$ se puede obtener de manera explícita gracias al teorema de programación dinámica.

$\Diamond$

Suponga que tenemos un controlador que actua sobre una máquina o un proceso, tales como un inyector de gasolina en un satélite o un regulador de temperatura en un reactor químico. Además dicha máquina tiene una "ruta"de funcionamiento pre-establecida tanto de estados como de control, que en nuestros ejemplos podría ser una órbita junto con el mínimo combustible de funcionamiento o temperaturas específicas de acuerdo al tiempo. Al problema de control óptimo que mejor se aproxime a las rutas pre-establecidas se le llama problema de seguimiento. Formalmente se describen de la siguiente forma:

Dada una ruta de control nominal $\{a_t^*\}_{t=0}^{\infty}$ y una trayectoria de estados nominal $\{x_t^*\}_{t=0}^{\infty}$, deseamos optimizar el "costo de seguimiento"

$$V(x,\pi) := \sum_{t=0}^{\infty}[(x_t - x_t^*)^2 + (a_t - a_t^*)^2].$$

Básicamente deseamos minimizar la distacia de ℓ^2 entre la trayectoria de estados y control $\{(x_t, a_t)\}_{t=0}^{\infty}$ y la trayectoria nominal $\{(x_t^*, a_t^*)\}_{t=0}^{\infty}$. Si este sistema está sujeto a una función de transición dada como en el ejemplo 3.3, este problema se reduce a un sistema LQ.

3.3. Sistema de administración forestal de Mitra-Wan.

Con anterioridad mencionamos que en ocasiones se desea maximizar en vez de minimizar las funciones de costo. Sin embargo, este caso se puede analizar con la teoría que hemos desarrollado y se podría pensar que "se pone un signo menos" a todas las hipótesis y resultados anteriores.

Suponga que $r : \mathbb{K} \to \mathbb{R}$ es una función medible que en cierto modelo sistémico representa una ganancia o recompensa. Entonces será deseable encontrar cuando

$$U_k^\alpha(x,\pi) := \sum_{n=0}^{k-1} \alpha^n r(x_n, a_n), \quad \alpha \in [0,1], x \in X, \pi \in \Pi,$$

alcance un máximo. Así que si definimos (de manera analoga) $U^\alpha(x,\pi), \overline{U}(x,\pi)$, deseamos encontrar $\pi^*, \tilde{\pi}$ en Π tal que para toda x en X

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

$$\overline{U}(x,\tilde{\pi}) = \sup_{\pi\in\Pi} \overline{U}(x,\pi) =: \overline{U}(x).$$

Tomando $c : \mathbb{K} \to \mathbb{R}$ como $c(\cdot,\cdot) := -r(\cdot,\cdot)$ y definiendo $V_k^\alpha, V^\alpha, \overline{V}$ de la manera usual, tenemos que

$$\textit{minimizamos} \begin{cases} V_k^\alpha(\cdot,\cdot), \\ V^\alpha(\cdot,\cdot), \\ \overline{V}(\cdot,\cdot), \end{cases} \quad \textit{sujeto a } F(\cdot,\cdot)$$

$$\Longleftrightarrow$$

$$\textit{maximizamos} \begin{cases} U_k^\alpha(\cdot,\cdot), \\ U^\alpha(\cdot,\cdot), \\ \overline{U}(\cdot,\cdot), \end{cases} \quad \textit{sujeto a } F(\cdot,\cdot).$$

Gracias a esta analogía, podremos obtener los mismos resultados del teorema 2.1 y teorema 2.2 con sus respectivas suposiciones análogas. Se deja al lector profundizar en los detalles.

Para la suposición 1, tenemos la siguiente analogía:

(a) La recompensa por etapa $r : \mathbb{K} \to \mathbb{R}$ es acotada por arriba.

(b) Para toda α en $(0,1)$ existe una política de control π en Π tal que

$$U^\alpha(x,\pi) > -\infty, \quad \forall x \in X.$$

Para la suposición 2, obtenemos:

(a) La función de recompensa por paso $r(\cdot,\cdot)$ es semicontinua por arriba en $\mathbb{K}$ (f es semicontinua por arriba si y solo si $-f$ es semicontinua por abajo).

(b) La función de transición $F : \mathbb{K} \to X$ es continua.

(c) Si $\{(x_n, a_n)\}_{n\in\mathbb{N}}$ es una sucesión en $\mathbb{K}$, x_n converge a x en X y $\{r(x_n, a_n)\}_{n\in\mathbb{N}}$ es acotada por abajo, entonces existe a en $A(x)$ tal que a es punto de acumulación de $\{a_n\}_{n\in\mathbb{N}}$.

Para la suposición 3, obtenemos:

(a) La constante $U^* := \sup_{x\in X} \overline{U}(x) > -\infty$.

(b) La función $g(\cdot) := \limsup_{\alpha\uparrow 1} g_\alpha(\cdot)$ *es finita. Donde* $g_\alpha(\cdot) := U^\alpha(\cdot) - \eta_\alpha$ *y* $\eta_\alpha := \sup_{x\in X} U^\alpha(x)$.

Con estas nuevas suposiciones obtenemos, por el teorema 2.1 y corolario 2.2 que existe f^* *en* $\mathbb{F}$ *tal que*

$$\sigma^* + g^*(\cdot) = r(\cdot, f^*) + g^*(F(\cdot, f^*)),$$

para $\sigma^* := -\rho^*, g^* := -h^*$; *además,* $\overline{U}(\cdot) = \overline{U}(\cdot, f^*)$.
Si suponemos que $\lim_{t\to\infty} \frac{1}{t} g^*(x_t^{f_t}) = 0$, *para cada* f_t *en* $\mathbb{F}$ *tal que* $U_t(\cdot, f_t) = U_t(\cdot)$. *Entonces por el teorema 2.2*

$$\lim_{k\to\infty} \overline{U}_k(x) = \overline{U}(x), \quad \forall x \in X.$$

Con un ejemplo mostraremos la utilidad de los cálculos.

Ejemplo 3.4. Sistema de manejo forestal de Mitra-Wan.
Supongamos que tenemos un terreno fértil del cual deseamos extraer madera. Podemos considerar a todo el terreno como una unidad y que tiene árboles de una especie en particular de distintas edades. Supongamos que las edades oscilan de 1 año a n años. Dicho de esta forma podemos considerar que el terreno (la unidad) puede ser dividido en porciones que correspondan a cada edad. Esto es, podemos describir al terreno vía $x = (x_1, \ldots, x_n)$ donde x_i es la fracción del terreno que está ocupada por árboles de edad i y $x_1 + \cdots + x_n = 1$. Además suponga que después de n-años, el árbol muere o que pierde su valor económico o que después de esta edad el árbol no crecerá más y por tanto, no producirá más madera en los siguientes años. Por este motivo, pasado el año n, los árboles de n años son cortados y la superficie es replantada.

Definamos Δ como el conjunto de todas las posibles configuraciones/estados del terreno, es decir

$$\Delta := \{x \in [0,1]^n : x_1 + x_2 + \cdots + x_n = 1\}.$$

Pensemos que en el período del tiempo $t \in \{0, 1, \ldots, T-1\}$, observamos nuestro terreno y nos encontramos la configuración $x(t) = (x_1(t), \ldots, x_n(t))$ en Δ. Entonces deseamos encontrar una acción para recolectar madera, la cual podemos escribir como $u(t) = (u_1(t), \ldots, u_n(t))$, donde $u_i(t)$ es el área a talar de árboles de i-años, por lo que $0 \leqslant u_i(t) \leqslant x_i(t)$. Como despues de n-años, los árboles pierden su valor económico, supondremos que $u_n(t) = x_n(t)$ para todo $t = 0, 1, \ldots, T-1$. Además, como el terreno lo mantendremos en constante reforestación, tendremos que plantar un área $\sum_{i=1}^n u_i(t)$ de árboles de edad 1. Con esta información ya sabremos que la siguiente configuración o estado será

$$x(t+1) = (\sum_{i=1}^n u_i(t), x_1(t) - u_1(t), \ldots, x_{n-1}(t) - u_{n-1}(t)).$$

También tenemos que dada una configuración x en Δ, el conjunto de las acciones admisibles para x es

$$A(x) := \{u \in [0,1]^n : u_i \leqslant x_i, i = 1, \ldots, n-1, u_n = x_n\},$$

y el conjunto de estados admisibles es

$$\mathbb{K} := \{(x,u) : x \in \Delta, u \in A(x)\}.$$

Definiendo

$$A := \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \qquad \text{y} \qquad B := \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ -1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 0 \end{pmatrix},$$

podemos ver que la función de transición $F : \mathbb{K} \to \Delta$ está dada como

$$F(x(t), u(t)) := x(t+1) = Ax(t) + Bu(t), \quad u(t) \in A(t), t = 0, 1, \ldots, T-1.$$

La producción de madera por unidad de área puede ser diferente y depende de las edades de los árboles. Por este motivo, relacionamos la producción de madera con un vector $\xi = (\xi_1, \ldots, \xi_n)$ en $\mathbb{R}^n$ donde $\xi_i \geqslant 0, i = 0, \ldots, n$ donde ξ_i representa la madera producida por unidad de área con arboles de i-años. Por lo que la madera producida en toda la recolección al tiempo t es

$$\langle \xi, u(t) \rangle = \sum_{i=1}^{n} \xi_i u_i(t).$$

Por las mismas razones, podemos considerar un vector $\gamma := (\gamma_1, \ldots, \gamma_n)$ con $\gamma_i \geqslant 0, i = 1, \ldots, n$ donde γ_i representa los recursos de mantenimiento (agua, abono, etc) de árboles de i-años por unidad de área. Así que los recursos administrados totales al tiempo t son

$$\langle \gamma, x(t) \rangle = \sum_{i=1}^{n} \gamma_i x_i(t).$$

Ahora considere una función de beneficio (o función de ingreso) por la madera producida $p : [0,\infty) \to [0,\infty)$. Entonces dado el plan de recolección u, el beneficio obtenido por $\langle \xi, u(t) \rangle$ madera recolectada es $p(\langle \xi, u(t) \rangle)$. De manera análoga, considere una función de costo por mantenimiento $c : [0,\infty) \to [0,\infty)$ tal que dado un estado x del terreno, el costo por $\langle \gamma, x(t) \rangle$ recursos de mantenimiento es $c(\langle \gamma, x(t) \rangle)$. Finalmente la función de recompensa dado el estado x y la acción u, es decir $r : \mathbb{K} \to \mathbb{R}$ es

$$r(x,u) := p(\langle \xi, u(t) \rangle) - c(\langle \gamma, x(t) \rangle).$$

En general, las funciones p y c pueden ser cualesquiera, pero para fijar ideas (y seguir los resultados de [10] o [11]), supondremos que p es cóncava, c es convexa y ambas continuas. De esta forma, tenemos r una función continua cóncava.

Está claro que dado un estado inicial x_0 en Δ, queremos encontrar una política π en Π para maximizar la recompensa. Por el momento pensemos en maximizar la recompensa promedio $\overline{U}(x,\pi)$.

Definamos para cada $i = 1, \ldots, n$

$$x^i := (\frac{1}{i}, \ldots, \frac{1}{i}, \frac{1}{i}, 0, \ldots, 0) \in \Delta, \quad u^i := (0, \ldots, 0, \frac{1}{i}, 0, \ldots, 0) \in A(x^i).$$

Se puede demostrar (para más detalles vea [10] o [11]) que si pedimos que exista un único k en $\{1, \ldots, n\}$ tal que

$$r(x^k, u^k) = \max\{r(x^i, u^i) : i = 1, \ldots, n\},$$

y además que r sea estrictamente cóncava en (x^k, u^k), entonces

$$\overline{U}(x) = p(\frac{\xi_k}{k}) - c(\frac{\gamma_1 + \cdots + \gamma_k}{k}) = r(x^k, u^k), \forall x \in \Delta. \quad (1)$$

En [10] o [11] se nos presentan un grupo de políticas en las que el supremo se alcanza. Suponga que queremos encontrar una de estas políticas vía el teorema 2.1. Entonces deseamos verificar que se cumplen las hipótesis análogas a la suposiciónes 1, 2 y 3.

Para la suposición 1 no hay mucho problema, pues para x en Δ y u en $A(x)$

$$0 \leqslant \langle \xi, u \rangle = \sum_{i=1}^{n} \xi_i u_i \leqslant \max_i \xi_i \sum_{i=1}^{n} u_i \leqslant \max_i \xi_i \sum_{i=1}^{n} x_i = \max_i \xi_i,$$

así que podemos restringir el dominio de p a $[0, \max_i \xi_i]$ y como p es continua, entonces p alcanza un máximo m. Así $r(x,u) \leqslant p(\langle \xi, u \rangle) \leqslant m$. Por lo tanto r es acotada por arriba. Es relativamente fácil imponer condiciones sobre p y c para que dado α en $(0,1)$ podamos encontrar π en Π tal que $U^\alpha(x,\pi) > -\infty$ para toda x en Δ.

Como p y c son continuas entonces r es semicontinua por arriba. Por la definición de la función de transición $F : \mathbb{K} \to X$, tenemos que F es continua. Si hacemos los calculos pertinentes, podemos verificar el inciso (c) del análogo a la suposición 2 (Se hace similar al ejemplo LQ usando el teorema de Bolzano-Weierstrass).

Por lo anterior tenemos que la suposición 1 y 2 valen para este ejemplo. Sin embargo, en general no se cumplirá la suposción 3. Para esto, nos concentraremos en un caso particular del modelo que describimos a continuación.

Pensemos que los árboles a plantar en todo el terreno son árboles endémicos, esto quiere decir que se pueden reproducir y crecer con los recursos que el ecosistema entrega. Por este motivo podemos pensar que $c(\cdot) = 0$ y luego $r(x,u) = p(\langle \xi, u\rangle)$. Además tomemos $p(t) := \sqrt{t}$. Tanto c como p, cumplen las condiciones previamente escritas.

Sea T en $\mathbb{N}$ y α en $(0,1)$. Por el teorema de la programación dinámica (ver Apéndice 4.2), si definimos

$$J_s^\alpha(x) := \max_{a\in A(x)} \{r(x,a) + \alpha J_{s+1}(F(x,a))\}, \quad \forall x \in \Delta, s = T-1,\dots,0,$$

con $J_T^\alpha(x) := 0$ para toda x en Δ, entonces $U_T^\alpha(x) = J_0^\alpha(x)$. Resolvamos estas ecuaciones “hacia atras”para el punto $x_i := (\delta_{i,m})_{m=1}^n$ con i en $\{1,\dots,n\}$. Por definición tenemos que $A(x_i) := \{a \in [0,1]^n : a_i \leqslant 1, a_j = 0, j \neq i\}$. Luego

$$J_{T-1}^\alpha(x_i) = \max_{a\in A(x_i)} \{r(x,a)\} = \max_{a\in A(x_i)} \{\sqrt{\xi_i a_i}\} = \sqrt{\xi_i},$$

lo que implica que

$$J_{T-2}^\alpha(x_i) = \max_{a\in A(x_i)} \{r(x,a) + \alpha\sqrt{\xi_i}\} = \max_{a\in A(x_i)} \{\sqrt{\xi_i a_i} + \alpha\sqrt{\xi_i}\} = \sqrt{\xi_i} + \alpha\sqrt{\xi_i},$$

e inductivamente

$$J_s^\alpha(x_i) = \sqrt{\xi_i}(1 + \alpha + \cdots + \alpha^{T-s-1}) = \sqrt{\xi_i}\frac{1-\alpha^{T-s}}{1-\alpha}.$$

Así que

$$U_T^\alpha(x_i) = J_0^\alpha(x_i) = \sqrt{\xi_i}\frac{1-\alpha^T}{1-\alpha}.$$

Al cumplirse la suposición 1 y 2, por el teorema 1.2 (iii) tenemos que

$$U^\alpha(x_i) = \lim_{T\to\infty} U_T^\alpha(x_i) = \frac{\sqrt{\xi_i}}{1-\alpha}.$$

Lo siguiente está inspirado en la proposición 4.2 de [1]. Ahora mostremos que para toda z en Δ, $U^\alpha(z) \geqslant \frac{\alpha^k}{1-\alpha}\sqrt{\frac{\xi_k}{k}}$ con k definido como en la hipótesis de la ecuación (1).

Sea z en Δ y $T > k$. Considere la siguiente política $\pi = \{a_t\}_{t=0}^T$ definida como

$$
\begin{aligned}
&z_0 = z && a_0 = z \\
&z_1 = (1, 0, \ldots, 0) && a_1 = (\frac{k-1}{k}, 0, \ldots, 0) \\
&z_2 = (\frac{k-1}{k}, \frac{1}{k}, 0, \ldots, 0) && a_2 = (\frac{k-2}{k}, 0, \ldots, 0) \\
&\vdots && \vdots \\
&z_j = (\frac{k-j+1}{k}, \frac{1}{k}, \ldots, \frac{1}{k}, 0, \ldots, 0) && a_j = (\frac{k-j}{k}, 0, \ldots, 0) \\
&\vdots && \vdots \\
&z_t = x^k && a_t = u^k \quad \forall T \geqslant t \geqslant k.
\end{aligned}
$$

Entonces

$$
\begin{aligned}
U_T^\alpha(z) &\geqslant \sum_{t=0}^{T-1} \alpha^t r(z_t, a_t) \geqslant \sum_{t=k}^{T-1} \alpha^t r(x^k, u^k) = \sum_{t=k}^{T-1} \alpha^t p(\langle \xi, u^k \rangle) \\
&= p(\frac{\xi_k}{k}) \frac{\alpha^k - \alpha^T}{1-\alpha},
\end{aligned}
$$

y de nuevo por el teorema 1.2 (iii)

$$
U^\alpha(z) = \lim_{T\to\infty} U_T^\alpha(z) = 0 \geqslant \lim_{T\to\infty} p(\frac{\xi_k}{k}) \frac{\alpha^k - \alpha^T}{1-\alpha} = \frac{\alpha^k}{1-\alpha}\sqrt{\frac{\xi_k}{k}}.
$$

Finalmente, $\eta_\alpha \geqslant \frac{\alpha^k}{1-\alpha}\sqrt{\xi_k/k}$ y luego

$$
\begin{aligned}
g^\alpha(x_i) = U^\alpha(x_i) - \eta_\alpha &\leqslant \frac{\sqrt{\xi_i}}{1-\alpha} - \frac{\alpha^k}{1-\alpha}\sqrt{\frac{\xi_k}{k}} \\
&= \frac{\sqrt{\xi_i} - \alpha^k\sqrt{\xi_k/k}}{1-\alpha},
\end{aligned}
$$

así que si tomamos i en $\{1, \ldots, n\}$ tal que $\sqrt{\xi_i} < \sqrt{\xi_k/k}$, entonces

$$
g(x_i) = \limsup_{\alpha\uparrow 1} g^\alpha(x_i) = -\infty.
$$

Por lo tanto $g(\cdot)$ no es finita y luego no se cumple la condición (b) del análogo a la suposición 3.

Para ver que en general, puede que exista dicho i, pongamos $n = 4$ y $\xi = (0, 2, 9, 11)$. En este caso $k = 3$ y $\sqrt{\xi_k/k} = \sqrt{3}$, por lo que $i = 1, 2, 4$ funcionará.

$\Diamond$

3.4. Sistema de Consumo e inversión.

Las ideas del siguiente ejemplo se pueden encontrar en [8] capítulo 3.6.

Ejemplo 3.5. Consumo e inversión.
Suponga que queremos asignar presupuesto al consumo e inversión en nuestra cartera o en nuestro balance. Dada nuestra riqueza al tiempo t, digamos x_t, podemos invertir cierta cantidad a_t en $A(x_t) := [0, x_t]$ y por tanto, dejar una cantidad para el consumo $x_t - a_t$ en cada periodo $t = 0, 1, \ldots$. Hasta este punto definimos $X := [0, M]$ con $M >> x_0$, y $A := [0, M]$.

Suponga que al invertir tenemos cierta tasa de retorno fija (o porcentaje de ganancia fijo), de tal manera que

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

con $m > 1$. Finalmente , tenemos una función de retorno en un paso

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

para alguna función $u : \mathbb{R}^+ \to \mathbb{R}^+$ llamada en este contexto utilidad de consumo. Para este ejemplo, pondremos $u = (b/\gamma)x^\gamma$, con $b > 0$ y $0 < \gamma < 1$; la cual es llamada como utilidad isoelástica marginal.

Demostremos que las suposiciones 1 y 2 se cumplen. Note que r es acotada por arriba por que $r(x, a) \leqslant r(x, 0) \leqslant (b/\gamma)M^\gamma$. Además, como $0 \leqslant r(x, a)$ entonces para todo α en $(0, 1)$, x en X y π en Π tenemos $U^\alpha(x, \pi) > -\infty$. Así que la suposición 1 se cumple. Por definición, r y F son continuas (luego r es semicontinua por arriba). La suposición 2-(c) es fácil comprobar usando el teorema de Bolzano-Weierstrass similar al sistema LQ. Todas estas ideas verifican las suposiciones 1 y 2.

Ahora veamos que pasa con la suposición 3. Una parte es sencilla, por que $U^* = \sup_{x \in X} \overline{U}(x) \geqslant 0 > -\infty$. La otra parte es calcular $g(x) = \limsup_{\alpha \uparrow 1} U^\alpha(x) - \eta_\alpha$ con $\eta_\alpha = \sup_{x \in X} U^\alpha(x)$ y verificar si es finita. Para esto, primero tendremos que maximizar

$$U_T^\alpha(x, \pi) = \sum_{n=0}^{T-1} \alpha^n u(x_n - a_n),$$

para T en $\mathbb{N}$ y α en $[0, 1]$. La manera de calcular la recompensa descontada de horizonte T la haremos mediante la programación dinámica, la misma que usamos en el ejemplo 3.3, solo que en su versión máximo. Esto es, encontrar para todo x en X

$$J_t(x) := \max_{a \in A(x)} [u(x - a) + \alpha J_{t+1}(F(x, a))], \quad t = T - 1, \ldots, 0,$$

con $J_k(x) := 0$. De esta forma, $U_T^\alpha(x) = J_0(x)$.

Note que

$$J_{T-1}(x) = \max_{a \in A(x)} (b/\gamma)(x-a)^\gamma = (b/\gamma)x^\gamma, \quad \forall x \in X.$$

Vamos a demostrar inductivamente que para todo x en X,

$$J_t(x) = (b/\gamma)D_t x^\gamma, \quad t = N-1, \dots, 0,$$

y la política que alcanza el máximo esta dada por

$$a_{T-1} = 0, \quad a_t = \frac{x}{1 + \delta D_{t+1}^{\frac{1}{\gamma-1}}}, \quad t = T-2, \dots, 0,$$

donde $\delta := (\alpha m^\gamma)^{\frac{1}{\gamma-1}}$ y D_t como $D_{T-1} = 1$ y $D_t = \delta^{\gamma-1} D_{t+1}/[1 + \delta D_{t+1}^{\frac{1}{\gamma-1}}]^{\gamma-1}$ para $t = T-2, \dots, 0$.

Para $T-1$ se cumple. Así que suponga que se cumple $J_{t+1}(x) = (b/\gamma)D_{t+1}x^\gamma$ para algún $0 \leqslant t \leqslant T-2$. Luego

$$J_t(x) = \max_{a \in A(x)} (b/\gamma)[(x-a)^\gamma + \alpha m^\gamma D_{t+1} a^\gamma].$$

Pongamos $f(a) := (x-a)^\gamma + \alpha m^\gamma D_{t+1} a^\gamma$. Note que J_{t+1} es una función estrictamente cóncava y como u tambien es cóncava, entonces f es estrictamente cóncava y además diferenciable, por lo que podemos calcular f' y luego resolver $f'(a) = 0$ para a en $A(x)$, y así a será tal que $J_t(x) = (b/\gamma)f(a)$. Note que

$$f'(a) = -\gamma(x-a)^{\gamma-1} + \alpha m^\gamma D_{t+1} \gamma a^{\gamma-1},$$

de aquí que $f'(a) = 0$ implica que

$$a_t = \frac{x}{1 + \delta D_{t+1}^{\frac{1}{\gamma-1}}}$$

y como $\delta, D_{t+1} \geqslant 0$ entonces $0 \leqslant a_t < x$, por lo que a_t está en $A(x)$. Luego

$$J_t(x) = (b/\gamma)D_t x^\gamma,$$

como lo deseabamos. Por otro lado, se puede demostrar recursivamente que $D_t = \left[\frac{\delta^{T-t-1}}{1+\delta+\cdots+\delta^{T-t-1}}\right]^{\gamma-1}$, lo que implica que

$$D_0 = \left[\frac{\delta^{T-1}(1-\delta)}{1-\delta^T}\right]^{\gamma-1}.$$

Así que

$$U_T^\alpha(x) = J_0(x) = \frac{b}{\gamma}\left[\frac{\delta^{T-1}(1-\delta)}{1-\delta^T}\right]^{\gamma-1} x^\gamma, \quad \delta = (\alpha m^\gamma)^{\frac{1}{\gamma-1}}.$$

Luego, por el teorema 1.2-(*iii*) y para α lo suficientemente cercano a 1 tal que $\delta \geqslant 1$, tenemos que

$$\begin{aligned} U^{\alpha}(x) &= \lim_{T\to\infty} U_T^{\alpha}(x) = \lim_{T\to\infty} \frac{b}{\gamma}\left[\frac{\delta^{T-1}(1-\delta)}{1-\delta^T}\right]^{\gamma-1} x^{\gamma} \\ &= \lim_{T\to\infty} \frac{b}{\gamma}\left[\frac{(1-\delta)}{\frac{1}{\delta^{T-1}}-\delta}\right]^{\gamma-1} x^{\gamma} = \frac{b}{\gamma}\left[\frac{(1-\delta)}{-\delta}\right]^{\gamma-1} x^{\gamma} \\ &= \frac{b}{\gamma}\left[\frac{\delta-1}{\delta}\right]^{\gamma-1} x^{\gamma}. \end{aligned}$$

Ahora, como $x \in X = [0, M]$, entonces

$$\eta_{\alpha} = \sup_{x\in X} U^{\alpha}(x) = \frac{b}{\gamma}\left[\frac{\delta-1}{\delta}\right]^{\gamma-1} M^{\gamma}.$$

Finalmente

$$\begin{aligned} g(x) = \limsup_{\alpha\uparrow 1} U^{\alpha}(x) - \eta_{\alpha} &= \limsup_{\alpha\uparrow 1} \frac{b}{\gamma}\left[\frac{(\alpha m^{\gamma})^{\frac{1}{\gamma-1}}-1}{(\alpha m^{\gamma})^{\frac{1}{\gamma-1}}}\right]^{\gamma-1}(x^{\gamma}-M^{\gamma}) \\ = \frac{b}{\gamma}\left[\frac{m^{\frac{\gamma}{\gamma-1}}-1}{m^{\frac{\gamma}{\gamma-1}}}\right]^{\gamma-1}(x^{\gamma}-M^{\gamma}) &= \frac{b}{\gamma}\frac{[m^{\frac{\gamma}{\gamma-1}}-1]^{\gamma-1}}{m^{\gamma}}(x^{\gamma}-M^{\gamma}). \end{aligned}$$

Así que g es finita. En resumen, se cumple la suposición 3. Así que estamos en condiciones de aplicar el teorema 2.1. Verifiquemos la última hipótesis del teorema 2.2. Como $m > 1$ y $0 < \gamma < 1$, entonces $0 < m^{\frac{\gamma}{\gamma-1}} < 1$. Además $x^{\gamma} \leqslant M^{\gamma}$, por lo que $g(\cdot) \geqslant 0$. Por la definición de g^* (via envolturas semicontinuas) tenemos que $0 \leqslant g^*(\cdot)$. Ahora no es dificil observar que

$$\begin{aligned} 0 \leqslant \frac{1}{t} g(x_t^{f_t}) &\leqslant \frac{1}{t}\frac{b}{\gamma}\frac{[m^{\frac{\gamma}{\gamma-1}}-1]^{\gamma-1}}{m^{\gamma}}((x_t^{f_t})^{\gamma}-M^{\gamma}) \\ &\leqslant \frac{1}{t}\frac{b}{\gamma}\frac{[m^{\frac{\gamma}{\gamma-1}}-1]^{\gamma-1}}{m^{\gamma}}(-M^{\gamma}) \to_{t\to\infty} 0, \end{aligned}$$

por lo que $0 \leqslant \frac{1}{t} g^*(x_t^{f_t}) \leqslant \frac{1}{t} g(x_t^{f_t}) \to_{t\to\infty} 0$. Finalmente $\lim_{t\to\infty} \frac{1}{t} g^*(x_t^{f_t}) = 0$.

Aplicando el teorema 2.2, obtenemos

$$\begin{aligned} \overline{U}(x) = \lim_{T\to\infty} \frac{1}{T} U_T(x) &= \lim_{T\to\infty} \frac{1}{T}\frac{b}{\gamma}\left[\frac{(m^{\frac{\gamma}{\gamma-1}})^{T-1}(1-m^{\frac{\gamma}{\gamma-1}})}{1-(m^{\frac{\gamma}{\gamma-1}})^{T}}\right]^{\gamma-1} x^{\gamma} \\ &= 0 \cdot \frac{b}{\gamma}\left[\frac{m^{\frac{\gamma}{\gamma-1}}-1}{m^{\frac{\gamma}{\gamma-1}}}\right]^{\gamma-1} x^{\gamma} = 0. \end{aligned}$$

$\Diamond$

3.5. Nota del capítulo 3.

Como se puede ver a lo largo de esta sección, hemos intuido un camino "estandar" para verificar la suposición 3–(b) y aplicar el teorema 2.2, pues en principio los objetos $h(\cdot)$, $\overline{V}_k(\cdot)$ y $\overline{V}(\cdot)$ no se conocen a priori. A continuación presentamos un breve resumen.

Suponga que las suposiciones 1 y 2 se cumplen.

Para verificar la finitud de $h(\cdot)$ es necesario obtener $h_\alpha(\cdot) = V^\alpha(\cdot) - m_\alpha$ para α en $(0,1)$, así que basta calcular $V^\alpha(\cdot)$, pues m_α está en terminos de V^α. Para calcular $V^\alpha(\cdot)$ primero calculamos $V_k^\alpha(\cdot)$ resolviendo inductivamente hacia atras las ecuaciones del teorema de programación dinámica (ver apéndice 4.2). Usando el teorema 1.2–(iii) obtenemos que

$$V^\alpha(\cdot) = \lim_{k\to\infty} V_k^\alpha(\cdot),$$

luego calculamos $m_\alpha = \inf_{x\in X} V^\alpha(x)$ y finalmente calculamos

$$h(\cdot) = \liminf_{\alpha\uparrow 1} h_\alpha(\cdot) = \liminf_{\alpha\uparrow 1} V^\alpha(\cdot) - m_\alpha.$$

El teorema de programación dinámica nos ayuda a encontrar explicitamente $x_t^{f_t}$, donde f_t es tal que $V_t(\cdot, f_t) = V_t(\cdot)$ para cada $t = 0, 1, \ldots$ y así poder verificar que

$$\lim_{t\to\infty} \frac{1}{t} h^*(x_t^{f_t}) = 0.$$

De nuevo mediante el teorema de programación dinámica aplicado a $V_t(\cdot,\cdot)$ podemos encontrar $\overline{V}_t(\cdot)$ pues

$$\overline{V}_t(\cdot) = \inf_\pi \frac{1}{t} V_t(\cdot, \pi) = \frac{1}{t} \inf_\pi V_t(\cdot, \pi) = \frac{1}{t} V_t(\cdot).$$

Finalmente podemos calcular el límite correspondiente al teorema 2.2

$$\overline{V}(\cdot) = \lim_{t\to\infty} \overline{V}_t(\cdot) = \lim_{t\to\infty} \frac{1}{t} V_t(\cdot).$$

Capítulo 4

Apéndice.

4.1. Teorema abeliano

Aquí repasamos algunos teoremas abelianos que ocupamos en la nota anterior al teorema 2.1. Hay diferentes teoremas abelianos pero cada uno dependen del método para sumar, es decir, de una función $L : \ell_c \to \mathbb{R}$, donde $\ell_c := \{(s_n)_{n\in\mathbb{N}_0} \in \mathbb{R}^{\mathbb{N}_0} | \quad (s_n)_n$ converge en $\mathbb{R}\}$ y $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. En general, podemos pensar que un teorema abeliano tiene la siguiente estructura.

Definición 4.1. Sea L un método para sumar. Decimos que L tiene un teorema abeliano si para cualquier sucesión $S := (s_n)_{n\in\mathbb{N}_0}$ en ℓ_c con $S \to s$, entonces $L(S) = s$.

$\Diamond$

Nosotros sólo utilizamos dos teoremas abelianos, que popularmente se conocen como el teorema Abeliano para la suma de Cesàro y el teorema de Abel para series de potencias.

Teorema 4.1. *Suma de Cesàro.*

Sea $L : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ definida como

$$L((s_n)_{n\in\mathbb{N}_0}) := \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} s_k.$$

Entonces L tiene teorema Abeliano.

Teorema 4.2. *Teorema de Abel.*

Sea $L : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ definida como

$$L((s_n)_{n\in\mathbb{N}_0}) := \limsup_{\alpha\uparrow 1} \sum_{k=0}^{\infty} \alpha^k s_k.$$

Entonces L tiene teorema Abeliano.

Usando ambos teoremas tenemos el siguiente corolario.

Corolario 4.1. Sea $(s_n)_{n\in\mathbb{N}_0}$ una sucesión convergente. Entonces

$$\limsup_{\alpha\uparrow 1}\sum_{k=0}^{\infty}\alpha^k s_k = \limsup_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}s_k.$$

4.2. Teorema de programación dinámica.

El teorema de programación dinámica es una herramienta poderosa en la teoría de control óptimo, ya que nos permite encontrar de manera inductiva el costo mínimo $\alpha-$descontado de horizonte finito, además de proveer una política $\alpha-$óptima de manera explícita. En general hay muchas variantes para este teorema como la variante $\alpha-$ descontada (que presentaremos a continuación), la variante para la máximización, la variante no descontada, etc, pero cada una de ellas tienen estructura muy similar.

Sean X y A los espacios de estados y de acciones admisibles respectivamente. Sean T en $\mathbb{N}$, α en $(0,1]$ y $c, C_T : \mathbb{K}\to\mathbb{R}$. Considere x_0 en X y π en Π, definimos

$$V_T^{\alpha}(x,\pi) := \sum_{t=0}^{T-1}\alpha^t c(x_t,a_t) + \alpha^T C_T(x_T).$$

Definamos las siguientes funciones recursivamente hacia atras

$$\begin{aligned} J_T(x) &:= C_T(x), \\ J_s(x) &:= \min_{a\in A(x)}\left[c(x,a) + \alpha J_{s+1}(F(x,a))\right] \quad \forall s = 0,\ldots,T-1, \end{aligned}$$

Las funciones anteriores son llamadas las ecuaciones de programación dinámica y son las protagonistas en el teorema de interés.

Teorema 4.3. *Teorema de programación dinámica.*
Suponga que para cada $s = 0,1,\ldots,T-1$ existe una función $a_s^* : X \to A$ tal que alcanza el mínimo del lado derecho de las ecuaciones de programación dinámica para todo x en X. Entonces la política de control $\pi^* := \{a_0^*,\ldots,a_{T-1}^*\}$ es $\alpha-$óptima en el horizonte T y además coincide con el valor en J_0, es decir

$$\inf_{\pi\in\Pi} V_T^{\alpha}(x,\pi) = V_T^{\alpha}(x,\pi^*) = J_0(x), \quad \forall x\in X.$$

Bibliografía

[1] M. Ali Khan, A. Piazza, *Classical turnpike theory and the economics of forestry, Journal of Economic Behavior & Organization (2011) 194-210.*

[2] R. Bellman, *A Markovian decision process. Journal of Mathematics and Mechanics, vol. 6, no. 5, 1957, pp. 679–684.*

[3] D. Blackwell, *Discrete dynamic programming, Ann. Math. Statist. 33 (1962), 719-726.*

[4] E.A. Feinberg, P.O. Kasyanov, M. Voorneveld, *Berge's maximum theorem for noncompact image sets, J. Math. Anal. Appl. 413 (2014), 1040-1046.*

[5] E.A. Feinberg, P.O. Kasyanov, N. V. Zadoianchuk, *Average Cost Markov Decision Processes with Weakly Continuous Transition Probabilities, Math. Oper. Res. 37 (2012), 591-607.*

[6] E.A. Feinberg, P.O. Kasyanov, N. V. Zadoianchuk, *Berge's theorem for noncompact image sets, J. Math. Anal. Appl. 397 (2013), 255-259.*

[7] D. González-Sánchez, O. Hernández-Lerma, L. R. Laura-Guarachi, S. Mendoza-Palacios, *An introduction to optimal control theory: Deterministic and stochastic models, Springer (Libro en preparación).*

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

[9] O. Hernández-Lerma, J.B. Lasserre, *Further Topics on Discrete-Time Markov Control Processes, Springer (1999), New York.*

[10] L. R. Laura-Guarachi, *An optimal control problem in forest management. In: Mendoza-Palacios S, and Mercado A. (eds) Games and Evolutionary Dynamics: Selected Theorical and Applied Developments, Ciudad de México, El Colegio de México, ISBN 978-607-564-285-7. (2021-for coming), Chapter 10, pp.189-209.*

[11] L. R. Laura-Guarachi, O. Hernández-Lerma, *The Mitra-Wan forestry model: A discrete-time optimal control problem, Natural Resource Modeling (2015), p. 152-168.*

[12] D. Meder, F. Rabe, T. Morville, H. H. Madsen, M. T. Koudahl, R. J. Dolan, H: R. Siebner, O. J. Hulme, Ergodicity-breaking reveals time optimal decision making in humans, Cornell University (2019).

[13] O. Peters, The ergodicity problem in economics, Nat. Phys. 15 (2019), 1216-1221.

[14] U. Rieder, Measurable selection theorems for optimization problems, Manuscripta Math. 24 (1978), 507-518.

[15] A.N. Vargas, E.F. Costa, J.B.R do Val, Advances in the control of Markov jump linear, Springer briefs in electrical and computer engineering-control, Automation and Robotics (2016), p.35-45.

[16] O. Vega-Amaya, On the vanishing discount factor approach for Markov decision processes with weakly continuous transition probabilities, J. Math. Anal. Appl. 426 (2015) 978-985.