



**Centro de Investigación y de Estudios Avanzados
del Instituto Politécnico Nacional
Unidad Zacatenco
Departamento de Matemáticas**

Control estocástico de tipo híbrido en tiempo discreto y aplicaciones

Tesis que presenta

Cristian Felipe Correa Rozo

Para obtener el grado de
Maestro en Ciencias

En la Especialidad de Matemáticas

Director de Tesis:
Dr. Héctor Jasso Fuentes

Ciudad de México

Marzo 2020

Dedicatoria

Esto va dedicado a ti, que me enseñaste a amar, a valorar cada momento de mi vida, irónicamente te convertiste en mi motor para continuar mi camino, hoy eres la fuerza que me ayuda a levantarme cada día. Gracias mi hermosa hermana, por cada momento, por cada instante de tenerte a mi lado, porque sin ti nada sería posible, espero pronto volverte a ver y estar a tu lado. Por esta razón esto va en tú honor Sandra Julieth Correa Rozo.

El Amor vive por siempre.

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología CONACYT por el apoyo económico brindado durante la realización de este proyecto.

Agradezco al Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional Cinvestav, esta institución de enorme calidad, que me brindó todo el apoyo durante mi estancia.

Presento mi gratitud al Departamento de Matemáticas del Cinvestav por el apoyo institucional, por estar siempre pendiente de mis avances, por el acompañamiento académico y sobre todo por la gran amabilidad, respeto, honestidad y compromiso de sus funcionarios, profesores, secretarias, etc. Fueron de gran ayuda para culminar satisfactoriamente esta etapa.

Quiero agradecerle a mi asesor de tesis, el Dr. Héctor Jasso Fuentes , por sus conocimientos invaluable, los cuales me ayudaron a llevar a cabo esta tesis, y sobre todo agradezco su gran paciencia para esperar a que este trabajo pudiera llegar a su fin.

Agradezco de antemano a los sinodales, por las valiosas contribuciones que harán al trabajo final y por el tiempo que dedicarán para revisarlo, aún a pesar de tantas actividades que los ocupan.

Presento gratitud a mi familia por estar siempre presente, a mis padres, hermanos y especialmente a mi esposa Katerin Semanate, por su gran disposición, compromiso, compañía, por darme ánimos en los momentos más difíciles y sobre todo por amarme tanto .

Doy Gracias a Jehova por darme fuerzas para seguir a pesar de todas las adversidades, siempre serás mi gran compañero de viaje, mi amigo y sobre todo el ser en el que más confié.

Agradezco a mis compañeros y amigos que siempre me animaron a continuar y a todos aquellos que se tomen la molestia de leer esta tesis.

Índice

Resumen	5
1. Introducción	7
2. Definición de modelo	11
2.1. Espacio de estado y acciones	11
2.2. Aleatoriedad y políticas de control	11
2.3. Evolución del estado y dinámica del proceso híbrido	12
2.4. Componente de tiempo y criterio de optimalidad	13
2.5. Planteamiento y solubilidad del problema	14
3. Programación dinámica: resultados de optimalidad	19
3.1. Conjunto de contacto y región de continuación	27
4. Aplicaciones	31
4.1. Problema de Inventario y Producción	31
4.2. Un problema de acumulación de contaminación	38
5. Conclusiones	49

Resumen

El propósito de este trabajo es analizar de manera formal y rigurosa un modelo estocástico de control híbrido a tiempo discreto y mostrar algunas aplicaciones que ilustren al mismo.

Los espacios de estados y de acciones considerados en el modelo son definidos de manera general como espacios de Borel y la función objetivo es tipo descontado en horizonte infinito. Para dar más intuición en los resultados, se considerará únicamente el caso en el que la dinámica se rige mediante una ecuación en diferencias.

Para dar resultados de optimalidad, se usará como herramienta el método de programación dinámica el cual caracteriza a la función de costo óptima como solución de cierta ecuación tipo funcional y los controles óptimos son también obtenidos mediante esta ecuación. También se mostrará una región clave contenida en el espacio de estados llamada región de continuidad (o continuación) la cual determina de manera óptima el tipo de control y dinámica a ejecutar a lo largo del tiempo.

Posterior a ello, se ilustrarán los resultados mediante dos aplicaciones: una que tiene que ver con un problema de inventario-producción y otra que trata un problema de acumulación de contaminación sobre una comunidad en la que existen diferentes tipos de contingencias ambientales. Para dar a entender un poco más estas aplicaciones, se dará un algoritmo elaborado en el software *Mathematica* con el fin de encontrar de manera aproximada los elementos óptimos asociados a éstas.

Finalizaremos el trabajo con algunas conclusiones y extensiones a realizarse en un futuro.

Abstract

The purpose of this work is to analyze in a formal and rigorous way a stochastic model of hybrid control at discrete time and show some applications that illustrate it.

The spaces of states and actions considered in the model are generally defined as Borel spaces and the objective function is a discounted type in an infinite horizon. To give more insight into the results, only the case in which the dynamics is governed by a difference equation will be considered.

To give results of optimality, the dynamic programming method will be used as a tool which characterizes the optimal cost function as a solution of a certain functional type equation and the optimal controls are also obtained through this equation. It will also show a key region contained in the state space called the continuity region (or continuation) which optimally determines the type of control and dynamics to be executed over time.

Subsequently, the results will be illustrated by two applications: one that has to do with an inventory-production problem and another that deals with a problem of accumulation of pollution over a community in which there are different types of environmental contingencies. To give a little more understanding of these applications, an algorithm developed in Mathematica software will be given in order to find approximately the optimal elements associated with them.

We will finish the work with some conclusions and extensions to be carried out in the future.

1. Introducción

Un sistema controlado se puede ver como un conjunto de componentes conectadas que interactúan entre sí mediante ciertas reglas y en donde el control determina el curso del estado del sistema. En particular, un sistema dinámico controlado es un sistema en el cual el tiempo juega un papel importante, en el sentido de que, cuando el control es seleccionado, el estado del sistema evoluciona con respecto al tiempo de acuerdo a una dinámica (o ley de evolución).

Los sistemas dinámicos controlados pueden ser clasificados de diversas formas, e.g., por el conjunto de parámetros que representan el tiempo (e.g. discreto o continuo) o por el tipo de espacios de acciones y de estados (e.g. finito o infinito numerable o no-numerable) o también por el tipo de función objetivo (e.g. descontada, promedio, sensible al riesgo). Respecto a la primera clasificación, hay modelos en los que se tienen que involucrar ambos tiempos para dar una adecuada interpretación del fenómeno a modelar. En el caso de modelos (o sistemas) controlados, el control juega un papel fundamental en cada tiempo, sea discreto y continuo. Es importante mencionar que dependiendo el fenómeno a modelar, en algunas ocasiones la interpretación que se le da a los tiempos continuo y discreto es reemplazada por tiempos rápido y lento, ambos discretos. En este trabajo estaremos considerando esta última interpretación.

En años anteriores, el concepto de control híbrido se creó para hacer frente a fenómenos en los que bajo condiciones normales, dicho fenómeno evoluciona a través del tiempo (rápido) por medio de una dinámica estándar, sin embargo en algún momento cierto evento (endógeno o exógeno) puede presenciarse haciendo que el modelo original (con dinámica estándar) se vuelva inválido, lo cual fuerza al modelador a reconsiderar los datos del problema haciendo cambios estructurales en el modelo de forma, digamos, drástica. Ejemplo de ello sería el *problema de acumulación de contaminación*. En este, un país consume un bien o producto específico y como resultado de esto se genera contaminación. Bajo condiciones normales de contaminación nuestro modelo evolucionará de acuerdo a una dinámica estándar, sin embargo cuando la contaminación sea muy alta se generará un estado extremo, obteniendo así, un cambio a una dinámica especial, la cual pueda modelar de forma correcta lo que está ocurriendo. Otro caso, es el problema de *Inventario-Producción*, en el cual, una empresa tiene una máquina que produce un artículo. Para su abastecimiento tiene la opción de producir o comprar ese artículo a los competidores. Bajo condiciones normales nuestra empresa podrá producir su artículo y la cantidad de artículos en inventario será modelada por una dinámica estándar, pero en ciertos casos será más beneficioso comprarle el artículo a los competidores, ya que, genera un menor costo, es ahí donde se activa una dinámica especial. Esto se podrá ver con más detalle en los ejemplos del capítulo 4. De esta manera, un modelo híbrido consiste de una sub-dinámica estándar la cual es utilizada durante casi cualquier situación, pero bajo circunstancias extremas, una sub-dinámica especial se activa, anulando la sub-dinámica estándar. Estos cambios entre las sub-dinámicas pueden producir modificaciones estructurales al sistema y al mismo tiempo dan una oportunidad de redimensionar al estado del sistema.

La manera de cómo se aplica el control en este tipo de modelos consiste en lo siguiente: cuando la dinámica está en su fase de sub-dinámica estándar (respectivamente sub-dinámica especial), el control toma valores en un conjunto llamado acciones estándar (respectivamente acciones especiales). También existe otro control de tipo no convencional que determina el momento en los que hay que cambiar de sub-dinámica. Este último control no siempre es permitido y la libertad de su uso dependerá en donde se ubique el estado del sistema; en particular, existirán regiones en el espacio de estados en donde es obligatorio ejercer alguna de las dos sub-dinámicas (estándar y especial) con sus respectivos controles.

Respecto al estado del sistema, éste consiste en un par, cuya primer componente muestra la descripción o medición real del fenómeno (variable estándar o rápida), mientras que la segunda guarda los cambios

estructurales del sistema, mismos que se conocen como configuraciones del sistema (variable especial o lenta).

En resumen, un modelo de control híbrido consiste en lo siguiente: (1) un estado del sistema con dos componentes: variable rápida y variable lenta, (2) dos sub-dinámicas: estándar y especial las cuales evolucionan en dos escalas de tiempo: tiempo rápido para la estándar y tiempo lento para la especial. El cambio entre estas sub-dinámicas puede ser ejecutada a través del controlador por medio del control no convencional o de manera obligatoria de acuerdo a la ubicación del estado del sistema y (3) un control que toma diferentes valores en sus acciones dependiendo en que sub-dinámica estamos ubicados y un control no convencional que determina el momento en el que habrá que hacer cambios entre sub-dinámicas.

Literatura relacionada. El estudio de modelos de control de tipo híbrido es muy vasto y cubre tanto resultados teóricos como prácticos, en modelos determinísticos como estocásticos, a tiempo discreto y continuo. Por mencionar algunos de estos trabajos desde el punto de vista teórico, podemos mencionar a Azhmayakov, et al. [1], Barles et al. [2], Bensoussan [3], Bensoussan y Menaldi [4], Branicky et al. [7], Borkar et al/ [6], Dhamarmatti y Ramaswamy [8], Lygeros [14], Riedinger et al. [17], Shainkh y Caines [18], Jasso-Fuentes et. al [11], [12]. Desde el punto de vista de aplicaciones, específicamente en robótica, planeación de aviones y teoría de autómatas, podemos mencionar, por ejemplo los trabajos de Posa et al [15], Soler et al. [19], Tavernini [20], aunque existen muchas aplicaciones más en la literatura.

Los aspectos teóricos de esta tesis se basan en la referencia Jasso-Fuentes et al. [11] en el caso especial donde la ley de transición se realiza a través de una ecuación en diferencias y las políticas de control se toman de manera no-aleatorizada. En cuanto a las aplicaciones en el trabajo, se hicieron adaptaciones a este contexto de ejemplos tomados en Kawaguchi y Morimoto [13] y de Hernández-Lerma y Lasserre [9].

Contenido de la tesis. El contenido de esta tesis se clasifica de la siguiente manera: después de esta parte introductoria, el Capítulo 2 consiste en la definición del modelo, la cual contempla las definiciones de espacio de estados y espacio de acciones, la introducción de las políticas de control a tratar y la evolución el proceso mediante su dinámica. También se dará el concepto de la componente de tiempo y se mostrará el criterio de optimalidad (función objetivo) con el que trabajaremos a lo largo del documento. Para finalizar el capítulo plantaremos la solubilidad del problema en el sentido de mostrar que la función objetivo es finita para al menos una política de control. En el Capítulo 3 se introducirán las ecuaciones de programación dinámica (EPDs) las cuales son una variación de las ecuaciones clásicas de programación dinámica, además de esto se definirá el concepto de conjunto de contacto o región de continuación. Mediante estos últimos elementos se demostrará la existencia de políticas de control óptimas y la existencia y caracterización de la función objetivo optimizada, esta última vista como solución de dichas EPDs. En el Capítulo 4 se trabajarán dos aplicaciones referentes al modelo de control híbrido, las cuales consisten en (1) un problema de inventario-producción y (2) un problema de acumulación de contaminación. En estas aplicaciones se mostrará cómo el modelo de control híbrido puede describir de forma satisfactoria estos fenómenos, además de ello, se verificará que cada aplicación cumple las hipótesis expuestas en el modelo teórico y por último se dará una solución aproximada a las EPDs por medio de un algoritmo; éste será desarrollado en el software *Mathematica*. Para finalizar daremos las conclusiones relacionadas a este trabajo de tesis, mostrando resultados obtenidos y algunas extensiones del trabajo.

Notación.

- $\mathbb{N}_0 = \{0, 1, \dots\} = \mathbb{N} \cup \{0\}$.
- A cada espacio métrico Z , lo dotaremos con su σ -álgebra de Borel denotada por $\mathcal{B}(Z)$ y el atributo *medible* se referirá tanto para conjuntos como para funciones.
- $\mathbb{M}(Z)$ denota a la familia de funciones medibles $g : Z \mapsto \mathbb{R} \cup \{\infty\}$. Análogamente $\mathbb{M}^+(Z)$ denota a la sub-familia de funciones no negativas de $\mathbb{M}(Z)$.
- Denotaremos a la familia de funciones medibles y acotadas $g : Z \mapsto \mathbb{R}$ por $\mathbb{B}(Z)$ y por $\mathbb{B}^+(Z)$ a la sub-familia de funciones no negativas de $\mathbb{B}(Z)$. Estos dos últimas, serán dotadas por la norma del supremo, denotada por $\|\cdot\|$.
- A la familia de funciones semi-continuas inferiormente (s.c.i) $g : Z \mapsto \mathbb{R} \cup \{\infty\}$ se denotará por $\mathbb{L}(X)$, mientras que al subconjunto de funciones s.c.i. no negativas por $\mathbb{L}^+(X)$.

2. Definición de modelo

En este capítulo presentaremos las bases para el desarrollo de nuestro problema de control híbrido. Definiremos los espacios de estados y el de acciones, el sistema controlado que tendrá una naturaleza aleatoria (o proceso estocástico controlado) y las políticas de control que llamaremos admisibles. También se definirá la función objetivo a tratar, la cual estará dada por medio de un costo acumulado en horizonte infinito y su solubilidad en el sentido de encontrar al menos una política de control que haga finito tal costo.

2.1. Espacio de estado y acciones

El *espacio de estados* X de un sistema de control híbrido se considerará de manera general como el producto $X = X^f \times X^s$ de dos conjuntos de Borel, donde las componentes $x^f \in X^f$ y $x^s \in X^s$ son llamados el estado rápido y estado lento, respectivamente. Por otro lado, el *espacio de acciones* denotado por A , va a ser considerado a través de la unión de dos conjuntos *disjuntos* de Borel, es decir, $A = A^f \cup A^s$ tal que $A^f \cap A^s = \emptyset$. A estos últimos se les llama el conjunto de acción rápida y el de acción lenta, respectivamente. El conjunto de *estado-acción admisible* está dado por el subconjunto medible $\mathbb{K} \subseteq X \times A$ con X -sección no vacía, es decir, $\emptyset \neq A(x^f, x^s) \subset A$, para cada $(x^f, x^s) \in X$.

Además asumiremos la existencia de dos conjuntos medibles

$$D^\wedge \subseteq D^\vee \subseteq X$$

tales que:

- (a) $A(x^f, x^s) \cap A^f = \emptyset$ cuando $(x^f, x^s) \in D^\wedge$, esto significa que cuando el estado del sistema se encuentra en D^\wedge , el controlador necesariamente debe escoger una acción en A^s (acción lenta).
- (b) $A(x^f, x^s) \cap A^s = \emptyset$ cuando $(x^f, x^s) \in X \setminus D^\vee$, significa que si el estado del sistema está por fuera de D^\vee , el controlador necesariamente debe escoger una acción en A^f (acción rápida).

De la definición anterior, los conjuntos D^\wedge y D^\vee resultan ser obligatorios para el controlador, ya que dentro de D^\wedge y fuera de D^\vee , el controlador está *forzado* a escoger una acción de una naturaleza específica (rápida o lenta). En contraste, cuando el estado del sistema (x^f, x^s) se encuentra en $D^\vee \setminus D^\wedge$, el controlador no tendrá una restricción a priori sobre la naturaleza de sus acciones y éste puede *elegir* cualquier acción, sea rápida o lenta.

2.2. Aleatoriedad y políticas de control

A lo largo de este documento asumiremos la existencia de un espacio de probabilidad *fijo*, definido como $\mathcal{E} := (\Omega, \mathcal{F}, P)$. Así pues consideraremos un espacio de Borel fijo denotado por Δ y a su vez la existencia de una sucesión de variables aleatorias independientes e idénticamente distribuidas (i.i.d) $\{\xi_t\}$ con valores en Δ y definidas sobre \mathcal{E} , mismas que tienen distribución de probabilidad común μ ; es decir, $P(\xi_t \in B) = \mu(B)$, para todo $t \in \mathbb{N}_0$ y $B \in \mathcal{B}(\Delta)$.

Ahora bien, para fines prácticos, consideraremos únicamente políticas no aleatorizadas (para más detalles sobre políticas aleatorizadas, ver Hernández-Lerma y Lasserre [9]). A saber, para $t \in \mathbb{N}_0$, definimos

el espacio H_t de las *historias admisibles* antes del tiempo t de la siguiente manera recursiva: $H_0 = X$, $H_t = (\mathbb{K} \times \Delta)^t \times X$, y $H_\infty = \mathbb{K}^\infty$, todos con el correspondiente producto de σ -álgebras. Un elemento genérico h_t de H_t , el cual es llamado una t -historia admisible, o simplemente t -historia, es un vector de la forma.

$$h_t = (x_0^f, x_0^s, a_0, \xi_0 \cdots, x_{t-1}^f, x_{t-1}^s, a_{t-1}, \xi_{t-1}, x_t^f, x_t^s) \in H_t.$$

Una política de control *admisibile* $\{\pi_t\}_{t \geq 0}$ es una sucesión de funciones medibles $\pi_t : H_t \mapsto A$ tal que

$$\pi_t(h_t) \in A(x_t^f, x_t^s) \quad \forall h_t \in H_t, t \in \mathbb{N}_0.$$

En particular, necesariamente se cumple que

$$\pi_t(h_t) \in A(x_t^f, x_t^s) \cap A^s \quad \text{si } (x_t^f, x_t^s) \in D^\wedge, \text{ y}$$

$$\pi_t(h_t) \in A(x_t^f, x_t^s) \cap A^f \quad \text{si } (x_t^f, x_t^s) \in X \setminus D^\vee.$$

Denotaremos por Π el conjunto de políticas de control admisibles.

Definimos por \mathbb{F} al espacio de funciones medibles $f : X \rightarrow A$ tal que $f(x^f, x^s) \in A(x^f, x^s)$, para toda $(x^f, x^s) \in X$. Si para algún $f \in \mathbb{F}$ tenemos que la política $\pi \in \Pi$ satisface que $\pi_t(h_t) = f(x_t^f, x_t^s)$ para todo $h_t \in H_t$ y $t \in \mathbb{N}_0$, entonces decimos que π es una política *determinista estacionaria Markoviana* o simplemente *estacionaria*. Además de esto, si identificamos las políticas admisibles con la propiedad anterior, tendremos que $\mathbb{F} \subseteq \Pi$.

2.3. Evolución del estado y dinámica del proceso híbrido

Asumamos la existencia de dos funciones medibles $F : X \times A^f \times \Delta \rightarrow X^f$ y $G : X \times A^s \times \Delta \rightarrow X$, tales que cumplan las siguientes propiedades: Dado un estado inicial $(x^f, x^s) \in X$ y una política de control $\pi = \{a_t\}_t \in \Pi$, se tiene que:

$$(x_0^f, x_0^s) = (x^f, x^s), \quad \text{y} \quad a_0 = \pi_0(x^f, x^s).$$

para $t \in \mathbb{N}_0$,

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = (F(x_t^f, x_t^s, a_t, \xi_t), x_t^s)}_{\text{sub-dinámica estándar}} \quad \text{si } a_t \in A^f, \quad (2.1)$$

o

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = G(x_t^f, x_t^s, a_t, \xi_t)}_{\text{sub-dinámica especial}} \quad \text{si } a_t \in A^s, \quad (2.2)$$

y

$$a_t = \pi_t(h_t), \quad \forall t \in \mathbb{N}_0,$$

con $\{\xi_t\}$ la sucesión de variables aleatorias i.i.d. con valores en Δ previamente definidas. También podemos definir la evolución de nuestro sistema dinámico de la siguiente manera compacta.

$$(x_{t+1}^f, x_{t+1}^s) = \mathbf{F}(x_t^f, x_t^s, a_t, \xi_t) = \begin{cases} (F(x_t^f, x_t^s, a_t, \xi_t), x_t^s) & \text{si } a_t \in A^f \\ G(x_t^f, x_t^s, a_t, \xi_t) & \text{si } a_t \in A^s, \end{cases} \quad (2.3)$$

$$a_t = \pi_t(h_t). \quad (2.4)$$

Intuitivamente, nuestro modelo de control híbrido puede resumirse en dos sub-modelos, en el cual el modelo de sub-dinámicas *rápidas* evoluciona de acuerdo a (2.1), mientras el de sub-dinámica *lenta* evoluciona de acuerdo a (2.2).

Entonces, por el teorema de Ionescu-Tulcea (ver e.g. Bertsekas y Shreve [5] pp. 140-141), para cada $x = (x^f, x^s) \in X$ y $\pi \in \Pi$, existe una única medida de probabilidad P_x^π en $(X \times A \times \Delta)^\infty$, con soporte en H_∞ tal que, para todo $B \in \mathcal{B}(X)$, $C \in \mathcal{B}(A)$, $D \in \mathcal{B}(\Delta)$ y $t \in \mathbb{N}_0$,

(a) $P_x^\pi(x \in B) = \delta_x(B)$.

(b) $P_x^\pi(a_t \in C | h_t) = \delta_{\pi_t(h_t)}(C)$.

(c) $P_x^\pi(\xi_t \in D | h_t) = P(\xi_t \in D) = \mu(D)$.

(d) Se tiene que

$$P_x^\pi((x_{t+1}^f, x_{t+1}^s) \in B | h_t, a_t) = P(\mathbf{F}(x_t^f, x_t^s, a_t, \xi_t) \in B | x_t^f, x_t^s, a_t) = \mu\left(\{\varrho \in \Delta | (F(x_t^f, x_t^s, a_t, \varrho), x_t^s) \in B\}\right) \mathbf{1}_{\{a_t \in A^f\}} + \mu\left(\{\varrho \in \Delta | G(x_t^f, x_t^s, a_t, \varrho) \in B\}\right) \mathbf{1}_{\{a_t \in A^s\}}.$$

Es claro que

$$\mu\left(\{\varrho \in \Delta | (F(x_t^f, x_t^s, a_t, \varrho), x_t^s) \in B\}\right) \mathbf{1}_{\{a_t \in A^f\}} = \mu\left(\{\varrho \in \Delta | F(x_t^f, x_t^s, a_t, \varrho) \in B^f\}\right) \delta_{\{x_t^s\}}(B^s) \mathbf{1}_{\{a_t \in A^f\}},$$

donde B^f y B^s son las proyecciones de B en X^f y X^s , respectivamente.

Definimos las leyes de transición como:

$$Q^f(B^f | x^f, x^s, a) := \mu\left(\{\varrho \in \Delta | F(x^f, x^s, a, \varrho) \in B^f\}\right) = \int_{\Delta} I_{B^f}[F(x^f, x^s, a, \varrho)] \mu(d\varrho) \quad (2.5)$$

y

$$Q^s(B | x^f, x^s, a) := \mu\left(\{\varrho \in \Delta | G(x^f, x^s, a, \varrho) \in B\}\right) = \int_{\Delta} I_B[G(x^f, x^s, a, \varrho)] \mu(d\varrho) \quad (2.6)$$

para todo $B \in \mathcal{B}(X)$ y $B^f := \{x^f \in X^f | (x^f, x^s) \in B\}$. Esta notación será utilizada más tarde.

2.4. Componente de tiempo y criterio de optimalidad

En este trabajo asumiremos que las transiciones al aplicar la sub-dinámica lenta (o especial) se realizan de manera *instantánea*. De esta manera el tiempo real o componente de tiempo en la k -ésima transición, se define por medio de la sucesión de variables aleatorias $\{\tau_k : k \geq 0\}$ con valores en \mathbb{N} cuyos elementos

τ_k representan el numero de veces que, previo a k , la acción en A^f ha sido tomada. La sucesión se define así: $\tau_0(\omega) = 0$ y en general, para cada $k \geq 1$ y $\omega \in \Omega$

$$\tau_k(\omega) = \sum_{j=0}^{k-1} \mathbf{1}_{A^f}(a_j(\omega)). \quad (2.7)$$

Por otro lado, definiremos el costo por unidad de tiempo como la función medible $l : \mathbb{K} \cap (X \times A^f) \rightarrow [0, \infty)$, así como el costo impulsivo (o costo de salto) $\ell : \mathbb{K} \cap (X \times A^s) \rightarrow [0, \infty)$, éste último también función medible.

Con estos elementos podemos definir el costo por cada transición como sigue

$$c(x^f, x^s, a) := l(x^f, x^s, a) \mathbf{1}_{A^f}(a) + \ell(x^f, x^s, a) \mathbf{1}_{A^s}(a). \quad (2.8)$$

Supondremos además que el costo (2.8) tiene un factor de descuento que no es constante. Esto permite considerar modelos más generales que algunos estudiados en la literatura. Así pues, consideraremos como factor de descuento a la función medible (fija) $\alpha : \mathbb{K} \rightarrow [0, 1]$, la cual, como se puede observar, puede tomar valores inclusive de *cero* o *uno*. Más supuestos sobre esta función se discutirán más adelante en la Hipótesis 2.1.

Dado un estado inicial $x := (x^f, x^s) \in X$ y una política de control $\pi \in \Pi$ definimos

$$J(x, \pi) = E_x^\pi \sum_{t=0}^{\infty} \left[c(x_t^f, x_t^s, a_t) \prod_{i=0}^{t-1} \alpha(x_i^f, x_i^s, a_i) \right]. \quad (2.9)$$

Ahora impongamos las siguientes condiciones a los elementos de (2.8) y (2.9).

Hipótesis 2.1. Para cada $(x, a) \in \mathbb{K}$:

- (a) La función costo ℓ es acotada por arriba, mientras que por abajo existe una constante $\ell_0 > 0$ tal que $\ell(x, a) \geq \ell_0$.
- (b) La función l es acotada por arriba y no negativa.
- (c) El factor de descuento satisface:

$$\begin{cases} \alpha(x, a) = 1, & \text{cuando } a \in A^s, \\ \alpha(x, a) \leq \alpha_0 < 1, & \text{cuando } a \in A^f, \end{cases} \quad (2.10)$$

para alguna constante dada $0 < \alpha_0 < 1$.

2.5. Planteamiento y solubilidad del problema

El problema de control óptimo a analizar se define de la siguiente manera: Dado J el costo total definido en (2.9), queremos encontrar condiciones para la existencia de una política $\pi^* \in \Pi$ tal que, para cada $x \in X$

$$J^*(x) := \inf_{\pi \in \Pi} J(x, \pi) = J(x, \pi^*) \quad \forall x \in X. \quad (2.11)$$

En tal caso decimos que π^* es *óptima* y a la función J^* se le llama *función de valor*.

En este apartado veremos que el problema de control óptimo planteado arriba, es soluble en el sentido de que existirá al menos una política de control $\pi \in \Pi$ bajo la cual el costo total definido en (2.9) es finito, luego entonces la función de valor J^* será finita también.

Para comenzar, haremos uso de la siguiente notación.

$$\mathbf{Q}(dz \times dm|y, n, a) = \begin{cases} Q^f(dz|y, n, a)\delta_n(dm) & \text{para } a \in A^f, \\ Q^s(dz \times dm|y, n, a) & \text{para } a \in A^s. \end{cases} \quad (2.12)$$

Además denotaremos por $\mathbf{Q}(\cdot|x, \mathbf{f}) := \mathbf{Q}(\cdot|x, \mathbf{f}(x))$, para todo $x := (x^f, x^s) \in X$ y $\mathbf{f} \in \mathbb{F}$.

Definición 2.2. La probabilidad de transición en t -pasos $\mathbf{Q}^t(B|x, \mathbf{f}) = P_x^{\mathbf{f}}(x_t \in B)$ se define de manera recursiva de la siguiente manera:

$$\mathbf{Q}^0(B|x, \mathbf{f}) := \delta_x(B) \quad \text{y} \quad \mathbf{Q}^t(B|x, \mathbf{f}) := \int_X \mathbf{Q}(B|\eta, \mathbf{f})\mathbf{Q}^{t-1}(d\eta|x, \mathbf{f}), \quad t \geq 1,$$

con $d\eta := dy^f \times dy^s$, $y \eta = y^f \times y^s$.

La siguiente hipótesis es crucial para la solubilidad del problema de control. Esta condición reemplaza el supuesto en otros trabajos de considerar el factor de descuento α acotado uniformemente por un valor menor a 1.

Hipótesis 2.3. Existe una política de control estacionaria $\mathbf{f} \in \mathbb{F}$, y una constante $m \geq 1$, tal que para $\epsilon > 0$ y para todo, $x := (x^f, x^s) \in X$,

$$\mathbf{Q}^m(D|x, \mathbf{f}) \leq 1 - \epsilon,$$

donde $D = \{x \in X : \mathbf{f}(x) \in A^s\}$.

La meta será el demostrar que, bajo nuestros supuestos ya dados, el costo (2.9) es finito cuando utilizamos la política $\mathbf{f} \in \mathbb{F}$ de arriba. Esto se verá reflejado en de las siguientes dos proposiciones.

Proposición 2.4. Supongamos que el estado inicial $x = (x^f, x^s)$ se encuentra en el conjunto D definido en la Hipótesis 2.3. Definimos $T = \min\{k \geq 0 : x_k \notin D\}$ como el tiempo de salida de D . Entonces, bajo las Hipótesis 2.1 y 2.3, se tiene que

$$P_x^{\mathbf{f}}\{T > km\} \leq (1 - \epsilon)^k \quad \text{para todo} \quad k \geq 0 \quad (2.13)$$

y

$$E_x^{\mathbf{f}}[T] \leq m/\epsilon \quad (2.14)$$

Demostración. Demostraremos esta proposición por inducción matemática: Observe que para el caso $k = 0$, la relación (2.13) se cumple trivialmente. Ahora tomemos $k = 1$, tenemos que

$$P_x^{\mathbf{f}}\{T > m\} = P_x^{\mathbf{f}}\{x_1, x_2, \dots, x_m \in D\} \leq P_x^{\mathbf{f}}\{x_m \in D\} = \mathbf{Q}^m(D|x, \mathbf{f}) \leq (1 - \epsilon),$$

donde la última desigualdad se debe a la Hipótesis 2.3.

En general, supongamos que el resultado se cumple para $k \geq 1$ y lo demostremos para el caso $k + 1$. En efecto,

$$\begin{aligned}
P_x^f\{T > (k+1)m\} &= P_x^f\{x_1, x_2, \dots, x_{(k+1)m} \in D\} \\
&= E_x^f[E_x^f[\mathbf{1}_{\{x_1, x_2, \dots, x_{(k+1)m} \in D\}} | h_{km}]] \\
&= E_x^f[\mathbf{1}_{\{x_1, x_2, \dots, x_{km} \in D\}} E_x^f[\mathbf{1}_{\{x_{km+1}, \dots, x_{(k+1)m} \in D\}} | h_{km}]] \\
&\leq E_x^f[\mathbf{1}_{\{x_1, x_2, \dots, x_{km} \in D\}} (1 - \epsilon)] \leq (1 - \epsilon)^{k+1},
\end{aligned}$$

en donde en la última expresión hemos utilizado la estacionariedad del proceso, el caso $k = 1$ y la hipótesis de inducción. Por lo tanto la expresión (2.13) se cumple para todo $k \geq 0$.

Usando (2.13) y la caracterización de la esperanza para variables aleatorias no negativas, tenemos que para $x \in D$,

$$E_x^f[T] = \sum_{k=0}^{\infty} P_x^f\{T > k\} \leq m \sum_{k=0}^{\infty} (1 - \epsilon)^k = m/\epsilon. \quad (2.15)$$

□

Proposición 2.5. *Bajo las Hipótesis 2.1, y 2.3, la función de valor J^* definida en (2.11) es acotada y por lo tanto finita.*

Demostración. Definamos dos sucesiones de variables aleatorias $\{T_r\}_{r \geq 0}$ y $\{S_r\}_{r \geq 0}$, como los tiempos de salida de D y D^c respectivamente. De forma específica definamos $T_1 = T$ y $S_1 = \min\{k > T_1 : x_k \in D\}$ y, para $r \geq 2$,

$$T_r = \min\{k > S_{r-1} : x_k \notin D\} \quad \text{y} \quad S_r = \min\{k > T_r : x_k \in D\}. \quad (2.16)$$

Podemos notar que si el estado inicial x esta en D , entonces $T_1 > 0$, mientras que si x no esta en D , entonces $T_1 = 0$. Si para algún $r \geq 1$, $S_r = \infty$, entonces el proceso no retorna a D , en caso contrario si para algún r tenemos que $S_r < \infty$, entonces por argumentos anteriores tendríamos que T_{r+1} es finito casi seguramente.

Así pues, supondremos primero que el estado inicial x esta en D y haremos la convención de $S_0 = 0$. Ahora bien, para algún $k \geq 1$ y para cualquiera $\omega \in \Omega$, definamos a $n_k(\omega)$ como la cantidad de veces que el proceso esta por fuera de D durante los primeros $k - 1$ periodos, es decir

$$n_k(\omega) = \sum_{j=0}^{k-1} \mathbf{1}_{D^c}(x_j).$$

Ya que $\alpha(y, f(y)) = 1$ cuando $y \in D$ y $\alpha(y, f(y)) \leq \alpha_0$ cuando $y \notin D$, tenemos que

$$\begin{aligned}
J(x, f) &\leq \|c\| E_x^f \left[\sum_{k=0}^{\infty} \alpha_0^{n_k} \right] \\
&= \|c\| + \|c\| \sum_{r=1}^{\infty} \left(E_x^f \left[\sum_{S_{r-1} < k \leq T_r} \alpha_0^{n_k} \right] + E_x^f \left[\sum_{T_r < k \leq S_r} \alpha_0^{n_k} \right] \right).
\end{aligned}$$

De esta manera, cuando $T_r < k \leq S_r$ los términos n_k se incrementan en uno. Agrupando todos los términos, obtenemos la serie $\|c\| \sum \alpha_0^k$. Por el contrario, cuando $S_{r-1} < k \leq T_r$, el término n_k permanece constante y es igual a

$$n_k = \sum_{j=1}^{r-1} (S_j - T_j) \geq r - 1$$

con esto $\alpha_0^{n_k} \leq \alpha_0^{r-1}$. Luego,

$$J(x, \mathbf{f}) \leq \frac{\|c\|}{1 - \alpha_0} + \|c\| \sum_{r=1}^{\infty} \alpha_0^{r-1} E_x^{\mathbf{f}} [T_r - S_{r-1}].$$

Por la propiedad de Markov, al tiempo S_{r-1} el proceso se encuentra en D y luego por la estacionariedad del proceso y usando la Proposición 2.4, tenemos ccc Usando un razonamiento análogo se llega a la última desigualdad cuando el proceso inicia en $x \notin D$. □

3. Programación dinámica: resultados de optimalidad

En esta sección vamos a estudiar la solución del problema de control híbrido (2.11). Para tal fin utilizaremos el conocido método de programación dinámica, el cual consiste en estudiar la solución de cierta ecuación funcional. Resulta ser que la función de valor J^* definida en (2.11) es solución de la mencionada ecuación; además, bajo ciertas condiciones, la solución es única. Finalmente se verá que el control óptimo se puede obtener a través de este tipo de ecuaciones y mostraremos una región del espacio de estados, llamada región de continuidad (o conjunto de contacto), la cual muestra la manera óptima de cambiar entre dinámicas y aplicar los controles óptimos correspondientes a cada una de ellas.

Ecuación de programación dinámica. Dadas una función medible $u \in \mathbb{M}(X)$ y \mathbf{F} como en (2.3), definimos el operador de programación dinámica $\mathcal{T}u$ en X de la siguiente manera:

$$\mathcal{T}u(x) = \inf_{a \in A(x)} \left\{ c(x, a) + \alpha(x, a) \int_{\Delta} u(\mathbf{F}(x, a, \varrho)) \mu(d\varrho) \right\}, \quad x := (x^f, x^s) \in X.$$

Tomando en cuenta la naturaleza del modelo de control híbrido, podemos definir los siguientes operadores $\mathcal{M}u$ y $\mathcal{H}u$ en X como sigue:

$$\mathcal{M}u(x) = \inf_{a \in A(x) \cap A^s} \left\{ \ell(x, a) + \int_{\Delta} u(G(x, a, \varrho)) \mu(d\varrho) \right\}, \quad x := (x^f, x^s) \in X \quad (3.1)$$

y

$$\mathcal{H}u(x) = \inf_{a \in A(x) \cap A^f} \left\{ l(x, a) + \alpha(x, a) \int_{\Delta} u((F(x, a, \varrho), x^s)) \mu(d\varrho) \right\}, \quad x := (x^f, x^s) \in X. \quad (3.2)$$

Por lo tanto, el operador de programación dinámica \mathcal{T} puede ser escrito en términos de éstos operadores de la siguiente forma:

$$\mathcal{T}u(x) = \begin{cases} \mathcal{M}u(x), & \text{si } x \in D^\wedge, \\ \min \{ \mathcal{M}u(x), \mathcal{H}u(x) \}, & \text{si } x \in D^\vee \setminus D^\wedge, \\ \mathcal{H}u(x), & \text{si } x \in X \setminus D^\vee. \end{cases}$$

De lo anterior, definimos la *ecuación de programación dinámica* (EDP) como:

$$u(x) = \mathcal{T}u(x) \quad \text{para todo } x \in X, \quad (3.3)$$

la cual por simplificación se escribirá como $u = \mathcal{T}u$.

En el resto de la sección, mostraremos que la función de valor J^* es una solución de la ecuación de programación dinámica (3.3), es decir $J^* = \mathcal{T}J^*$. También se mostrará cómo se puede obtener una política óptima de esta ecuación. En esta tesis trabajaremos el caso cuando las funciones medibles $F(\cdot, \delta)$ y $G(\cdot, \delta)$, definidas en (2.1) y (2.2) respectivamente son funciones continuas en $X \times A$ para toda $\delta \in \Delta$; esto nos llevará a demostrar que la transición \mathbf{Q} es débilmente continua.

Hipótesis 3.1. Para todo $x = (x^f, x^s) \in X$ tenemos:

(a) La función $x \rightarrow A(x)$ es compacta valuada y semi-continua superior.

(b) Las funciones medibles $F(\cdot, \delta)$ y $G(\cdot, \delta)$ son continuas en $X \times A$ para cada $\delta \in \Delta$

(c) Las funciones c y α en (2.8) y (2.9) son semi-continuas inferiores en \mathbb{K}

Proposición 3.2. *Bajo la Hipótesis 3.1, la función de transición \mathbf{Q} definida en (2.12) es débilmente continua; es decir, para toda función continua y acotada $u : X \mapsto \mathbb{R}$, la función*

$$(x, a) \mapsto \int_X u(y) \mathbf{Q}(dy|x, a), \quad (3.4)$$

es continua y acotada en \mathbb{K} .

Demostración. El hecho de que (3.4) es acotada, se sigue directamente de la propiedad de que u es acotada y $\mathbf{Q}(\cdot|x^f, x^s, a)$ es una medida de probabilidad para todo $(x^f, x^s, a) \in \mathbb{K}$.

La continuidad de (3.4) se basa del siguiente hecho: Dada una función continua y acotada $u : X \rightarrow \mathbb{R}$ y un elemento $x := (x^f, x^s) \in X$, se tiene

$$\int_X u(y^f, y^s) \mathbf{Q}(dy^f \times dy^s|x^f, x^s, a) = \begin{cases} \int_X u(y^f, y^s) Q^f(dy^f|x^f, x^s, a) \delta_{x^s}(dy^s) & \text{si } a \in A^f, \\ \int_X u(y^f, y^s) Q^s(dy^f \times dy^s|x^f, x^s, a) & \text{si } a \in A^s. \end{cases} \quad (3.5)$$

Aplicando la fórmula de cambio de variable para integrales y de (2.5), tenemos

$$\int_X u(y^f, y^s) Q^f(dy^f|x, a) \delta_{x^s}(dy^s) = \int_{\Delta} u(F(x, a, \varrho), x^s) \mu(d\varrho) \quad (3.6)$$

Ya que que $x \mapsto u(x)$ es continua y acotada sobre X y $(x^f, x^s, a) \mapsto F(x^f, x^s, a, \cdot)$ es continua en \mathbb{K} , entonces se tiene que $(x^f, x^s, a) \mapsto u(F(x^f, x^s, a, \cdot), x^s)$ es continua y acotada. Ahora bien tomemos una sucesión $\{(x_n, a_n)\}_{n \geq 0} \in \mathbb{K}$, tal que $(x_n, a_n) \rightarrow (x, a)$ en \mathbb{K} , definamos:

$$g_n(\varrho) := u(F(x_n, a_n, \varrho), x_n^s).$$

Notemos que la sucesión de funciones $\{g_n\}_{n \geq 0}$ es medible y acotada, además $g_n(\varrho) \rightarrow u(F(x, a, \varrho), x^s)$, cuando $n \rightarrow \infty$. Entonces, haciendo tender $n \rightarrow \infty$, $(x_n, a_n) \rightarrow (x, a)$, y por el teorema de convergencia dominada,

$$\lim_{n \rightarrow \infty} \int_{\Delta} g_n(\varrho) \mu(d\varrho) = \int_{\Delta} u(F(x, a, \varrho), x^s) \mu(d\varrho).$$

Esto implica que la función

$$(x, a) \rightarrow \int_{\Delta} u(F(x, a, \varrho), x^s) \mu(d\varrho),$$

es continua en \mathbb{K} , esto es justamente lo que se deseaba. Análogamente, si $a \in A^s$ y $u : X \rightarrow \mathbb{R}$ es una función continua y acotada, de (2.6) tenemos:

$$\int_X u(y^f, y^s) Q^s(dy^f \times dy^s|x, a) = \int_{\Delta} u(G(x, a, \varrho)) \mu(d\varrho). \quad (3.7)$$

Aplicando el hecho de que $(x^f, x^s, a) \mapsto G(x^f, x^s, a, \cdot)$ es continua en \mathbb{K} y siguiendo argumentos análogos que el anterior, se puede ver fácilmente que (3.7) es continua en \mathbb{K} . Así, \mathbf{Q} es débilmente continua usando su definición en (2.12). \square

En lo que sigue presentaremos la definición de semi-continuidad inferior de funciones y mostraremos algunos resultados de estas funciones, mismos que nos serán de gran utilidad más adelante.

Definición 3.3 (Funciones semi-continuas inferiormente). *Dado X un espacio métrico y $u : X \mapsto \mathbb{R} \cup \{\infty\}$ tal que $u(x) < \infty$ para al menos un punto $x \in X$. Decimos que u es una función semi-continua inferiormente (s.c.i.) en $x \in X$ si*

$$\liminf_{n \rightarrow \infty} u(x_n) \geq u(x) \quad (3.8)$$

para alguna sucesión $\{x_n\}$ en X que converja a x .

Recuerde en la parte introductoria de esta tesis que el conjunto de funciones s.c.i. $u : X \mapsto \mathbb{R} \cup \{\infty\}$ se denotará por $\mathbb{L}(X)$, mientras que al subconjunto de funciones s.c.i. no negativas por $\mathbb{L}^+(X)$. A continuación se dan algunas propiedades de estas funciones. Para más detalles sobre las demostraciones, ver [16], [9].

Proposición 3.4. (a) *Si $u \in \mathbb{L}(X)$, entonces existe una sucesión de funciones continuas y acotadas u_n en X tal que $u_n \uparrow u$.*

(b) *Dado X un espacio métrico completo y separable, $u, v \in \mathbb{L}^+(X)$, donde $u > 0$. Entonces $u + v \in \mathbb{L}^+(X)$ y $uv \in \mathbb{L}^+(X)$.*

Notación: Para simplificar la notación, consideraremos de ahora en adelante la siguiente notación compacta. $x := (x^f, x^s) \in X$ y $dy := dy^f \times dy^s$.

Proposición 3.5. *Bajo las Hipótesis 3.1 (b) y (c), para toda $u \in \mathbb{L}^+(X)$ la función $(x, a) \rightarrow c(x, a) + \alpha(x, a) \int_{\Delta} u(\mathbf{F}(x, a, \varrho)) \mu(d\varrho)$ es s.c.i. en \mathbb{K} .*

Demostración. Tomemos un punto $(x, a) \in \mathbb{K}$, de tal forma que podamos elegir una sucesión $(x_n, a_n) \in \mathbb{K}$ que converge a (x, a) . Dado que u es el límite puntual de una sucesión de funciones monótonas no decrecientes, no negativas, acotadas y continuas $u_n \uparrow u$, tenemos que para cada $k \geq 1$

$$\liminf_{n \rightarrow \infty} \int_{\Delta} u(\mathbf{F}(x_n, a_n, \varrho)) \mu(d\varrho) \geq \liminf_{n \rightarrow \infty} \int_{\Delta} u_k(\mathbf{F}(x_n, a_n, \varrho)) \mu(d\varrho) = \int_{\Delta} u_k(\mathbf{F}(x, a, \varrho)) \mu(d\varrho)$$

donde la última desigualdad, se sigue de la continuidad de \mathbf{F} en \mathbb{K} , ya que al ser u_k una familia de funciones acotadas y continuas podemos utilizar el teorema de convergencia dominada. Ahora bien por teorema de convergencia monótona, se obtiene

$$\liminf_{n \rightarrow \infty} \int_{\Delta} u(\mathbf{F}(x_n, a_n, \varrho)) \mu(d\varrho) \geq \lim_{k \rightarrow \infty} \int_{\Delta} u_k(\mathbf{F}(x, a, \varrho)) \mu(d\varrho) = \int_{\Delta} u(\mathbf{F}(x, a, \varrho)) \mu(d\varrho) \quad (3.9)$$

esto prueba que la función $(x, a) \rightarrow \int_{\Delta} u(\mathbf{F}(x, a, \varrho)) \mu(d\varrho)$ es s.c.i. en \mathbb{K} .

Así, el resultado se sigue de la Hipótesis 3.1 (c) y la Proposición 3.4(b). \square

Para $u \in \mathbb{L}^+(X)$, tenemos que $\mathcal{T}u : X \rightarrow [0, \infty]$ está definido como:

$$\mathcal{T}u(x) = \inf_{a \in A(x)} \left\{ c(x, a) + \alpha(x, a) \int_{\Delta} u(\mathbf{F}(x, a, \varrho)) \mu(d\varrho) \right\} \quad \text{para todo } x \in X \quad (3.10)$$

A continuación daremos algunos comentarios sobre este operador.

Observación 3.6. (a) *Note que existe la posibilidad que $\mathcal{T}u(x)$ sea infinita para algún punto $x \in X$.*

(b) *Debido a la no-negatividad de los elementos de \mathcal{T} , éste último resulta ser un operador monótono, esto significa que para $u, v \in \mathbb{L}^+(X)$ con $u \leq v$, tenemos $\mathcal{T}u \leq \mathcal{T}v$.*

Finalmente llegamos a un resultado auxiliar pero crucial para nuestro análisis.

Proposición 3.7. *Bajo las Hipótesis 2.1 y 3.1, para toda $u \in \mathbb{L}^+(X)$ el operador $\mathcal{T}u$ es s.c.i. y existe $\mathbf{f} \in \mathbb{F}$, tal que:*

$$\begin{aligned} \mathcal{T}u(x) &= \inf_{a \in A(x)} \left\{ c(x, a) + \alpha(x, a) \int_{\Delta} u(\mathbf{F}(x, a, \varrho)) \mu(d\varrho) \right\} \\ &= c(x, \mathbf{f}(x)) + \alpha(x, \mathbf{f}(x)) \int_{\Delta} u(\mathbf{F}(x, \mathbf{f}(x), \varrho)) \mu(d\varrho) \end{aligned} \quad (3.11)$$

para cada $x \in X$.

Demostración. Por la Proposición 3.5 sabemos que $(x, a) \mapsto c(x, a) + \alpha(x, a) \int_{\Delta} u(\mathbf{F}(x, a, \varrho)) \mu(d\varrho)$ es s.c.i., luego la primera parte del resultado se obtiene usando la Proposición D.5(b) en Hernández-Lerma y Lasserre [9]. La segunda parte de la demostración concerniente a la existencia de un mínimo, se sigue del hecho de que estamos minimizando una función s.c.i. en un conjunto compacto $A(x)$ —ver Hipótesis 3.1(a), luego tenemos que el mínimo en (3.10) se alcanza para cada $x \in X$ o equivalentemente (3.11) se sigue —ver más detalles en Himmelberg et. al. [10]. \square

Ahora definamos recursivamente a u_k para $k \geq 0$. Dado $u_0 := \mathbf{0}$ es la función nula en X y para $k \geq 0$, tenemos $u_{k+1} := \mathcal{T}u_k$. Vamos a probar algunas propiedades de la sucesión $\{u_k\}_{k \geq 0}$.

Teorema 3.8. *Supongamos que las Hipótesis 2.1, 2.3 y 3.1 se cumplen, entonces*

(a) *Para cada $k \geq 0$, la función u_k se encuentra en $\mathbb{L}^+(X)$*

(b) *La sucesión $\{u_k\}_{k \geq 0}$ converge monótonamente a una función $u^* \in \mathbb{L}^+(X)$ con $u^* \leq J^*$.*

(c) *La función u^* satisface $u^* = \mathcal{T}u^*$*

Demostración. (a) Vamos a verificar que, para cada $k \geq 0$, cada elemento de la sucesión $\{u_k\}_{k \geq 0}$ es acotado y son funciones no negativas y s.c.i. Para ello, vemos primero que

$$\mathcal{T}u_k \leq c(x, a) + \alpha(x, a) \int_{\Delta} u_k(\mathbf{F}(x, a, \varrho)) \mu(d\varrho),$$

lo cual da $\|u_{k+1}\| = \|\mathcal{T}u_k\| \leq \|c\| + \|u_k\|$, de aquí sigue que la función es acotada. Por otro lado, como $c \geq 0$, es fácil ver que $u_{k+1} = \mathcal{T}u_k$ es no negativa y finalmente, por la Proposición 3.7, se concluye que $u_{k+1} \in \mathbb{L}^+(X)$ para cada $k \geq 0$.

(b) Vamos a demostrar que la sucesión $\{u_k\}_{k \geq 0}$ es monótona no decreciente. A saber, dado que la función de costo c es no negativa, tenemos que $u_0 \leq u_1$. Razonando de manera inductiva, supongamos que para $k \geq 1$ se cumple que $u_{k-1} \leq u_k$, por ser \mathcal{T} monótono, tenemos que $u_k \leq u_{k+1}$, así $\{u_k\}_{k \geq 0}$ es monótona no decreciente, por lo tanto converge a alguna función medible y no negativa u^* . Resta demostrar que $u^*(x)$ es finita para toda $x \in X$ y que $u^* \in \mathbb{L}^+(X)$. En efecto, de la Proposición 3.7 tenemos que para cada $k \geq 0$ existe un selector medible $\mathbf{f}_k \in \mathbb{F}$ tal que

$$u_{k+1}(x) = c(x, \mathbf{f}_k(x)) + \alpha(x, \mathbf{f}_k(x)) \int_{\Delta} u_k(\mathbf{F}(x, \mathbf{f}_k(x), \varrho)) \mu(d\varrho) \quad \text{para todo } x \in X. \quad (3.12)$$

Por otro lado, fijo $N \in \mathbb{N}$, nos concentraremos en la decisiones que se toman para la evolución del proceso en transiciones $\{0, \dots, N\}$. Podemos asumir sin pérdida de generalidad que cada política $\pi \in \Pi$ es vista como una política de control de problemas con horizonte de planeación de 0 a N , de la siguiente manera: $\pi = \{\pi_0, \pi_1, \dots, \pi_N\}$. Ahora fijemos $0 \leq t \leq N$ y $\pi \in \Pi$, de tal manera que para cada t -historia $h_t = (x_0, a_0, \xi_0, \dots, x_t) \in H_t$, podemos definir un costo total descontado de t a N , como sigue:

$$J_{t,N}(h_t, \pi) = E_{x_0}^{\pi} \left[\sum_{k=t}^N c(x_k, a_k) \prod_{j=t}^{k-1} \alpha(x_j, a_j) \middle| h_t \right].$$

Es obvio que $J_{t,N}(h_t, \pi)$ depende de la política π en los tiempos $\{t, \dots, N\}$; esto es, la función depende del conjunto $\{\pi_t, \pi_{t+1}, \dots, \pi_N\}$. Definamos ahora

$$J_{t,N}(h_t) := \inf_{\pi \in \Pi} J_{t,N}(h_t, \pi) \quad \text{para } h_t \in H_t \quad (3.13)$$

y la política $\pi_N^* = \{\mathbf{f}_0, \dots, \mathbf{f}_N\} \in \Pi$ con \mathbf{f}_i obtenida como (3.12) en el rango $\{0, \dots, N\}$. Nuestro objetivo ahora es mostrar que, fijando un $N \geq 0$, se tendrá que

$$J_{t,N}(h_t) = J_{t,N}(h_t, \pi_N^*) = u_{N+1-t}(x_t) \quad \text{para cada } 0 \leq t \leq N \quad \text{y } h_t \in H_t \quad (3.14)$$

Demostremos este hecho por inducción matemática hacia atrás en t :

Si $t = N$ tenemos que para cada $\pi \in \Pi$ y $h_N \in H_N$

$$J_{N,N}(h_N, \pi) = E_{x_0}^{\pi} [c(x_N, a_N) | h_N] = c(x_N, \pi_N(h_N)),$$

y así

$$J_{N,N}(h_N) = \min_{a \in A(x_N)} c(x_N, a) = c(x_N, \mathbf{f}_0(x_N)) = J_{N,N}(h_N, \pi_N^*) = u_1(x_N)$$

Ahora supongamos que (3.14) cumple para $t+1$ y demostremos que se cumple para t . Tomemos $\pi \in \Pi$ y $h_t \in H_t$, entonces

$$\begin{aligned} J_{t,N}(h_t, \pi) &= E_{x_0}^{\pi} \left[E_{x_0}^{\pi} \left[\sum_{k=t}^N c(x_k, a_k) \prod_{j=t}^{k-1} \alpha(x_j, a_j) \middle| h_{t+1} \right] \middle| h_t \right] \\ &= E_{x_0}^{\pi} \left[c(x_t, a_t) + \alpha(x_t, a_t) E_{x_0}^{\pi} \left[\sum_{k=t+1}^N c(x_k, a_k) \prod_{j=t+1}^{k-1} \alpha(x_j, a_j) \middle| h_{t+1} \right] \middle| h_t \right] \\ &= E_{x_0}^{\pi} \left[c(x_t, a_t) + \alpha(x_t, a_t) J_{t+1,N}(h_{t+1}, \pi) \middle| h_t \right] \\ &\geq E_{x_0}^{\pi} \left[c(x_t, a_t) + \alpha(x_t, a_t) u_{N-t}(x_{t+1}) \middle| h_t \right] \\ &= \left[c(x_t, \pi_t(h_t)) + \alpha(x_t, \pi_t(h_t)) \int_{\Delta} u_{N-t}(\mathbf{F}(x_t, \pi_t(h_t), \varrho)) \mu(d\varrho) \right] \\ &\geq u_{N+1-t}(x_t) \end{aligned}$$

y se cumple la igualdad cuando $\pi = \pi_N^*$. Así hemos demostrado que para $t = 0$, $N \geq 0$ y $x \in X$.

$$J_{0,N}(x) = J_{0,N}(x_0, \pi_N^*) = u_{N+1}(x),$$

y puesto que la función costo por etapa es no negativa, esto implica que, para cada $x \in X$, $N \geq 0$ y $\pi \in \Pi$.

$$u_{N+1}(x) \leq J_{0,N}(x, \pi) \leq J(x, \pi) \quad (3.15)$$

y así $u_{N+1}(x) \leq J^*(x)$, la cual por la Proposición 2.5, es finita, esto muestra que el limite $u^*(x) \leq J^*(x)$, es finita en X , y así concluimos que $u^* \in \mathbb{L}^+(X)$.

(c) Para todo $k \geq 0$, tenemos que $u_{k+1} := \mathcal{T}u_k \leq \mathcal{T}u^*$ y así $u^* \leq \mathcal{T}u^*$, Ahora bien para la desigualdad inversa, podemos tomar $\mathbf{f}_k \in \mathbb{F}$ para $k \geq 0$ tal que

$$u_{k+1}(x) = c(x, \mathbf{f}_k(x)) + \alpha(x, \mathbf{f}_k(x)) \int_{\Delta} u_k(\mathbf{F}(x, \mathbf{f}_k(x), \varrho)) \mu(d\varrho) \quad \text{para todo } x \in X. \quad (3.16)$$

Ahora bien fijemos $x \in X$ arbitrario. Ya que $A(x)$ es compacto, existe una sub-sucesión $\{k_n\}_{n \geq 0}$, con $f_{k_n}(x) = a_{k_n} \rightarrow a^*$ para algún $a^* \in A(x)$. Fijando algún n_0 y suponiendo que $n \geq n_0$, tenemos lo siguiente

$$\int_{\Delta} u_{k_n}(\mathbf{F}(x, a_{k_n}, \varrho)) \mu(d\varrho) \geq \int_{\Delta} u_{k_{n_0}}(\mathbf{F}(x, a_{k_n}, \varrho)) \mu(d\varrho). \quad (3.17)$$

La función $u_{k_{n_0}} \in \mathbb{L}^+(X)$, por lo tanto tenemos que $\int_{\Delta} u_{k_{n_0}}(\mathbf{F}(\cdot, \cdot, \varrho)) \mu(d\varrho)$ es semi-continua inferior en \mathbb{K} . Entonces,

$$\liminf_{n \rightarrow \infty} \int_{\Delta} u_{k_n}(\mathbf{F}(x, a_{k_n}, \varrho)) \mu(d\varrho) \geq \int_{\Delta} u_{k_{n_0}}(\mathbf{F}(x, a^*, \varrho)) \mu(d\varrho). \quad (3.18)$$

Usando el teorema de convergencia monótona cuando $n_0 \rightarrow \infty$,

$$\liminf_{n \rightarrow \infty} \int_{\Delta} u_{k_n}(\mathbf{F}(x, a_{k_n}, \varrho)) \mu(d\varrho) \geq \int_{\Delta} u^*(\mathbf{F}(x, a^*, \varrho)) \mu(d\varrho) \quad (3.19)$$

ahora por la semi-continuidad de c y α , obtenemos

$$\begin{aligned} u^*(x) &= \liminf_{n \rightarrow \infty} u_{k_n+1}(x) \\ &\geq \liminf_{n \rightarrow \infty} c(x, a_{k_n}) + \liminf_{n \rightarrow \infty} \left[\alpha(x, a_{k_n}) \int_{\Delta} u_{k_n}(\mathbf{F}(x, a_{k_n}, \varrho)) \mu(d\varrho) \right] \\ &\geq c(x, a^*) + \alpha(x, a^*) \int_{\Delta} u^*(\mathbf{F}(x, a^*, \varrho)) \mu(d\varrho) \geq \mathcal{T}u^*(x). \end{aligned}$$

por lo tanto $u^* = \mathcal{T}u^*$. □

Una vez que hemos probado la existencia de soluciones de la ecuación de programación dinámica (EPD) en cuestión, haremos un puente entre esta solución y la función de valor J^* de nuestro problema de control híbrido. Esta y otras propiedades están dadas en el siguiente resultado.

Teorema 3.9. *Bajo las hipótesis del Teorema 3.8, se tiene:*

(a) *La función óptima de costo descontado J^* es igual a la función limite u^* obtenida en el Teorema 3.8, de hecho J^* resulta ser la solución mínima en $L^+(X)$ de la EPD (3.3).*

(b) Existe $\mathbf{f} \in \mathbb{F}$ para la cual la ecuación óptima $J^* = \mathcal{T}J^*$ alcanza el mínimo en \mathbf{f} ; ésta política es óptima para el problema de control híbrido.

(c) Si, además, $J^* \in \mathbb{B}^+(X)$, entonces J^* es la única solución de la EPD (3.3) en $\mathbb{L}^+(X) \cap \mathbb{B}^+(X)$.

Demostración. (a) Sea $v \in \mathbb{L}^+(X)$ solución de la ecuación de programación dinámica $v = \mathcal{T}v$. Por la Proposición 3.7, existe $\mathbf{f} \in \mathbb{F}$ con:

$$v(x) = c(x, \mathbf{f}(x)) + \alpha(x, \mathbf{f}(x)) \int_{\Delta} v(\mathbf{F}(x, \mathbf{f}(x), \varrho)) \mu(d\varrho) \quad \text{para todo } x \in X \quad (3.20)$$

Interando la ecuación anterior, obtenemos que para cada $N \geq 0$ y $x \in X$

$$\begin{aligned} v(x) &= E_x^\pi \left[\sum_{t=0}^N c(x_t, \mathbf{f}(x_t)) \prod_{i=0}^{t-1} \alpha(x_i, \mathbf{f}(x_i)) \right] + E_x^\pi \left[v(x_{N+1}) \prod_{i=0}^N \alpha(x_i, \mathbf{f}(x_i)) \right] \\ &\geq E_x^\pi \left[\sum_{t=0}^N c(x_t, \mathbf{f}(x_t)) \prod_{i=0}^{t-1} \alpha(x_i, \mathbf{f}(x_i)) \right] \end{aligned}$$

Así cuando $N \rightarrow \infty$ y usando el teorema de convergencia monótona obtenemos:

$$v(x) \geq \lim_{N \rightarrow \infty} E_x^\pi \left[\sum_{t=0}^N c(x_t, \mathbf{f}(x_t)) \prod_{i=0}^{t-1} \alpha(x_i, \mathbf{f}(x_i)) \right] = J(x, \mathbf{f}) \geq J^*(x). \quad (3.21)$$

Entonces podemos observar que si $v \in \mathbb{L}^+(X)$ es una solución de la EPD, entonces $v \geq J^*$. De otro lado, en la Proposición 3.8 mostramos que $u^* \in \mathbb{L}^+(X)$ es una solución de la ecuación de optimización con $u^* \leq J^*$; así $u^* = J^*$, siendo esta última la mínima solución en $\mathbb{L}^+(X)$.

(b) Si aplicamos el mismo razonamiento de la parte (a) para $v = J^*$, de (3.21) obtenemos que:

$$J^*(x) \geq J(x, \mathbf{f}) \geq J^*(x);$$

luego, $J^*(x) = J(x, \mathbf{f})$, por lo tanto necesariamente $\mathbf{f} \in \mathbb{F}$ debe ser óptima.

(c) Ahora supongamos a $v \in \mathbb{L}^+(X)$ y acotada, la cual resuelve $v = \mathcal{T}v$. De la parte (a) tenemos que $v \geq J^*$. Ahora demostraremos que $v \leq J^*$. Fijemos $x \in X$ y $\pi \in \Pi$. Claramente, si $J(x, \pi) = \infty$, entonces $v(x) \leq J(x, \pi)$. Por lo tanto en lo que sigue, supondremos que, $J(x, \pi) < \infty$

$$v(x) \leq c(x, a) + \alpha(x, a) \int_{\Delta} v(\mathbf{F}(x, a, \varrho)) \mu(d\varrho) \quad \text{para todo } (x, a) \in \mathbb{K},$$

Ahora bien, para todo $t \geq 0$, tenemos

$$\begin{aligned} E_x^\pi \left[v(x_{t+1}) \prod_{i=0}^t \alpha(x_i, a_i) \middle| h_t, a_t \right] &= \prod_{i=0}^t \alpha(x_i, a_i) \int_{\Delta} v(\mathbf{F}(x_t, a_t, \varrho)) \mu(d\varrho) \\ &= \prod_{i=0}^{t-1} \alpha(x_i, a_i) \left[c(x_t, a_t) + \alpha(x_t, a_t) \int_{\Delta} v(\mathbf{F}(x_t, a_t, \varrho)) \mu(d\varrho) - c(x_t, a_t) \right] \\ &\geq \prod_{i=0}^{t-1} \alpha(x_i, a_i) [v(x_t) - c(x_t, a_t)] \end{aligned}$$

ahora tomando esperanza E_x^π en ambos lados de la anterior desigualdad, obtenemos

$$\begin{aligned} E_x^\pi \left[c(x_t, a_t) \prod_{i=0}^{t-1} \alpha(x_i, a_i) \right] &\geq E_x^\pi \left[v(x_t) \prod_{i=0}^{t-1} \alpha(x_i, a_i) \right] \\ &\quad - E_x^\pi \left[v(x_{t+1}) \prod_{i=0}^t \alpha(x_i, a_i) \right] \end{aligned}$$

sumando la anterior desigualdad para $t = 0, 1, 2, \dots, N$, tenemos

$$E_x^\pi \left[\sum_{t=0}^N c(x_t, a_t) \prod_{i=0}^{t-1} \alpha(x_i, a_i) \right] \geq v(x) - E_x^\pi \left[v(x_{N+1}) \prod_{i=0}^N \alpha(x_i, a_i) \right]$$

aplicando limites a ambos lados,

$$\lim_{N \rightarrow \infty} E_x^\pi \left[\sum_{t=0}^N c(x_t, a_t) \prod_{i=0}^{t-1} \alpha(x_i, a_i) \right] \geq v(x) - \lim_{N \rightarrow \infty} E_x^\pi \left[v(x_{N+1}) \prod_{i=0}^N \alpha(x_i, a_i) \right]$$

ahora por teorema de convergencia monótona,

$$J(x, \pi) \geq v(x) - \lim_{N \rightarrow \infty} E_x^\pi \left[v(x_{N+1}) \prod_{i=0}^N \alpha(x_i, a_i) \right],$$

bastaría demostrar que $\lim_{N \rightarrow \infty} E_x^\pi \left[v(x_{N+1}) \prod_{i=0}^N \alpha(x_i, a_i) \right] = 0$, para obtener el resultado deseado. Puesto que $v \in \mathbb{L}^+(X)$ y acotada, se sigue

$$\lim_{N \rightarrow \infty} E_x^\pi \left[v(x_{N+1}) \prod_{i=0}^N \alpha(x_i, a_i) \right] \leq \lim_{N \rightarrow \infty} E_x^\pi \left[\prod_{i=0}^N \alpha(x_i, a_i) \right] \|v\|$$

demostramos que $\lim_{N \rightarrow \infty} E_x^\pi \left[\prod_{i=0}^N \alpha(x_i, a_i) \right] = 0$. Suponiendo que se cumple la Hipótesis 2.1 y que $J(x, \pi)$ es finito, consideremos a $x_0 = x$ y los siguientes conjuntos medibles

$$\begin{aligned} A_t &= \{(x_0, a_0, \xi_0, x_1, a_1, \xi_1, \dots) \in H_\infty : a_t \in A^s\} \quad \text{para } t \geq 0, \\ B &= \{(x_0, a_0, \xi_0, x_1, a_1, \xi_1, \dots) \in H_\infty : \alpha(x_t, a_t) > 0\} \quad \text{para todo } t \geq 0, \end{aligned}$$

escojamos un historia

$$(x_0, a_0, \xi_0, x_1, a_1, \xi_1, \dots) \in B \cap \liminf A_k.$$

Si para algún t_0 , $a_t \in A^s$ para todo $t \geq t_0$, por la Hipótesis 2.1 tenemos que $\alpha(x_t, a_t) = 1$ y $c(x_t, a_t) = \ell(x_t, a_t) \geq \ell_0$. Ya que $(x_0, a_0, \xi_0, x_1, a_1, \xi_1, \dots) \in B$, se tiene

$$\prod_{i=0}^{t-1} \alpha(x_i, a_i) = \prod_{i=0}^{t_0-1} \alpha(x_i, a_i) > 0, \quad \forall t \geq t_0$$

y así

$$\sum_{t=0}^{\infty} c(x_t, a_t) \prod_{i=0}^{t-1} \alpha(x_i, a_i) = \infty.$$

Por lo anterior, el costo total descontado en las trayectorias que pertenecen a $B \cap \liminf A_t$ es infinito. Pero por hipótesis $J(x, \pi)$ es finita, entonces necesariamente $P_x^\pi(B \cap \liminf A_t) = 0$, o equivalentemente

$$P_x^\pi(B^c \cup \limsup A_t^c) = 1.$$

Tomemos ahora una historia $(x_0, a_0, \xi_0, x_1, a_1, \xi_1, \dots) \in B^c$. Tenemos que $\prod_{i=0}^t \alpha(x_i, a_i)$ converge a cero pues por lo menos un factor de descuento se anula. Por otro lado, tomando un historia $(x_0, a_0, \xi_0, x_1, a_1, \xi_1, \dots) \in \limsup A_t^c$, lo cual significa que $a_t \in A^f$ para un número infinito de t , y así $\prod_{i=0}^t \alpha(x_i, a_i)$ también converge a cero, puesto que $\alpha(x_t, a_t) \leq \alpha_0$ para un número infinito de t . En resumen, hemos mostrado que $\prod_{i=0}^t \alpha(x_i, a_i) \rightarrow 0$ P_x^π -casi seguramente. Aplicando el teorema de convergencia dominada, concluimos entonces que

$$\lim_{N \rightarrow \infty} E_x^\pi \left[\prod_{i=0}^N \alpha(x_i, a_i) \right] = 0$$

que es justamente lo que deseábamos demostrar. Por lo tanto, $J(x, \pi) \geq v(x)$, teniendo así $J^* \geq v$. En conclusión J^* es la única solución de $v = \mathcal{T}v$ en $\mathbb{L}^+(X)$. \square

3.1. Conjunto de contacto y región de continuación

En los Teoremas 3.9 y 3.8, dimos condiciones suficientes bajo las cuales la *función de valor* (2.11) llega a ser solución de la EPD (3.3) y mostramos además cómo obtener una política óptima estacionaria a partir de la ecuación de punto fijo $J^* = \mathcal{T}J^*$. Definiendo a $x := (x^f, x^s)$ y recordando los operadores de programación dinámica (3.1) y (3.2), definidos anteriormente, obtenemos

$$J^*(x) = \begin{cases} \mathcal{M}J^*(x), & \text{si } x \in D^\wedge, \\ \min \{ \mathcal{M}J^*(x), \mathcal{H}J^*(x) \}, & \text{si } x \in D^\vee \setminus D^\wedge, \\ \mathcal{H}J^*(x), & \text{si } x \in X \setminus D^\vee. \end{cases}$$

Además, considerando a $f \in \mathbb{F}$ un selector en el cual la ecuación de punto fijo $J^* = \mathcal{T}J^*$ alcanza su mínimo, entonces $f \in \mathbb{F}$ es una política óptima. De esto último, es fácil ver que si el estado del sistema x se encuentra ubicado en $D^\vee \setminus D^\wedge$, obtendríamos las siguientes situaciones:

- Si $\mathcal{M}J^*(x) > \mathcal{H}J^*(x)$ entonces la acción óptima $f(x)$ esta en A^f ;
- Si $\mathcal{M}J^*(x) < \mathcal{H}J^*(x)$ entonces la acción óptima $f(x)$ esta en A^s ;
- Si $\mathcal{M}J^*(x) = \mathcal{H}J^*(x)$ entonces la acción óptima $f(x)$ puede ser escogida de A^s o A^f .

Claramente, cuando el estado x esta en D^\wedge o $X \setminus D^\vee$, la acción óptima necesariamente debe ser escogida de A^s o A^f , respectivamente.

Definamos ahora del conjunto D^* como sigue:

$$D^* = \{x \in D^\vee : J^*(x) = \mathcal{M}J^*(x)\},$$

A este conjunto se le denomina *conjunto de contacto*, éste es considerado como la región óptima, en el sentido que fuera de su cerradura, la regla óptima es aplicar una acción en A^f , y como consecuencia la sub-dinámica estándar es activada y una vez que el sistema alcanza el interior de D^* , la regla óptima es aplicar una acción en A^s , luego entonces se activaría la sub-dinámica especial. La siguiente proposición resumirá un poco lo mencionado anteriormente.

Proposición 3.10. *Si para cierto estado x , éste esta fuera de $\overline{D^*}$, la acción óptima necesariamente debe ser escogida en A^f , mientras que si el estado x , esta en el interior de D^* , la acción óptima debe ser tomada en A^s .*

Demostración. En primer lugar, notemos que la regla óptima para escoger una política de control en la región $X \setminus D^\vee$, ya fue dada anteriormente durante el desarrollo de este documento, necesariamente debemos aplicar una acción $a \in A^f$. También, la ecuación de programación dinámica (3.3) muestra que cuando un estado $x \in D^\vee$, pero, $x \notin \overline{D^*}$,

$$J^*(x) = \min \{ \mathcal{M}J^*(x), \mathcal{H}J^*(x) \}, \quad (3.22)$$

entonces necesariamente tenemos los siguientes casos: (a) $J^*(x) > \mathcal{M}J^*(x)$ o (b) $J^*(x) < \mathcal{M}J^*(x)$. Mostraremos que la regla óptima, necesariamente será el aplicar una acción de A^f . En efecto, por (3.22) y considerando el caso (a) primero, tenemos que $\min \{ \mathcal{M}J^*(x), \mathcal{H}J^*(x) \} > \mathcal{M}J^*(x)$, lo cual es una contradicción. Por otro lado combinando la opción (b) con (3.22), obtenemos que $\min \{ \mathcal{M}J^*(x), \mathcal{H}J^*(x) \} < \mathcal{M}J^*(x)$, con esto, $\mathcal{H}J^*(x) < \mathcal{M}J^*(x)$, entonces, en virtud de las condiciones dadas al inicio de esta sección, tenemos que la regla óptima sería aplicar el control $f(x) \in A^f$.

Ahora bien, si x esta en el interior de D^* , también obtendremos dos casos: (a) $x \in D^\wedge$ o (b) $x \in D^\vee \setminus D^\wedge$. Considerando el caso (a), tenemos que por definición de las políticas admisibles, necesariamente debemos tomar un control en A^s . Ahora en el caso (b), notemos que la ecuación (3.22) se cumple en el interior de D^* , por lo tanto,

$$\min \{ \mathcal{M}J^*(x), \mathcal{H}J^*(x) \} = \mathcal{M}J^*(x) \leq \mathcal{H}J^*(x).$$

Entonces, teniendo en cuenta las condiciones dadas al principio de la sección, necesariamente se debe cumplir que $f(x) \in A^s$ cuando $\mathcal{M}J^*(x) < \mathcal{H}J^*(x)$, y cuando $\mathcal{M}J^*(x) = \mathcal{H}J^*(x)$ tenemos la opción de escoger a $f(x) \in A^s$, con esto obtenemos que si x esta fuera de $\overline{D^*}$, entonces debemos aplicar un control de A^f , mientras que si x esta en el interior de D^* , debemos aplicar un control de A^s , que es justamente lo que deseábamos demostrar. \square

Las reglas anteriores, sirven para saber cuándo aplicar una acción óptima, estas son claras cuando nuestro estado x , esta dentro o fuera de D^* ; Sin embargo, nuestras hipótesis son insuficientes para dar una regla de elección cuando nuestro estado x , este en la frontera de D^* , (∂D^*) . Una condición suficiente es extender la proposición anterior sobre todo el espacio de estados X y asumir que el conjunto D^* es cerrado (abierto); Otra condición suficiente puede darse en el espíritu de la continuidad $\mathcal{M}J^*$ y J^* . La siguiente proposición nos dará la regla de elección de nuestra política óptima en ∂D^* .

Proposición 3.11. *Asumamos que J^* y $\mathcal{M}J^*$ son funciones continuas en X . Entonces el conjunto de contacto D^* es cerrado, lo cual implica que la acción óptima a escoger en ∂D^* debe pertenecer a A^s .*

Demostración. Tomemos una sucesión $\{x_k\}_{k \geq 0}$ en D^* , tal que $x_k \rightarrow x$, para algún $x \in X$. Entonces, para cada k , tenemos

$$J^*(x_k) = \mathcal{M}J^*(x_k), \quad k \geq 1.$$

Tomando límites a ambos lados de la expresión anterior, y usando la continuidad de J^* y $\mathcal{M}J^*$, tenemos $J^*(x) = \mathcal{M}J^*(x)$, así $x \in D^*$, lo que implica que D^* es cerrado.

En efecto la acción óptima que debemos escoger en ∂D^* , debe estar en A^s , esto se sigue de la proposición anterior y al tener que $D^* = \overline{D^*}$ por ser cerrado. \square

El conjunto de contacto D^* , nos permite definir un conjunto que llamaremos región de continuación, ésta es muy común en los problemas de control impulsivos. La región de continuación simplemente es el complemento del conjunto de contacto D^* .

Los resultados obtenidos a lo largo de este documento nos dan bases sólidas para poder entender de forma correcta lo que es un modelo de control híbrido, desde su planteamiento hasta su desarrollo. En esta tesis nos basamos en las ecuaciones de programación dinámica para caracterizar la función de valor y mostrar la existencia de políticas de control óptimas. Sin embargo, creemos muy conveniente mostrar algunos ejemplos, los cuales serán desarrollados en la siguiente sección.

4. Aplicaciones

En esta sección nos ocuparemos en dar dos aplicaciones relacionadas con modelos de control híbrido. El objetivo es mostrar de forma explícita la ecuación de programación dinámica y dar soluciones aproximadas usando *mathematica* como herramienta para el análisis numérico en cada ejemplo.

4.1. Problema de Inventario y Producción

Consideremos un sistema de inventario y producción en el que una compañía produce un solo tipo de artículo. La producción es efectuada por una máquina, la cual tiene una única configuración predefinida. La variable aleatoria x_t^f representa el inventario del artículo en el tiempo $t \geq 0$, la cual toma valores en \mathbb{N}_0 . Asumiremos que la variable x_t^s es constante igual a Λ , en consecuencia el conjunto $X^s = \{\Lambda\}$; esto quiere decir que sólo estaremos considerando la configuración de que la(s) máquina(s) de producción siempre están en funcionamiento. En cada periodo de tiempo t , la variable de control a_t puede ser alguna de las siguientes opciones: (a) el artículo producido por la compañía o (b) el artículo que compra la compañía a competidores externos. Asumamos que esta última tiene una capacidad de almacenamiento $C < \infty$, de esta manera el conjunto de acciones y el de acciones admisibles estarán dados por $A = \{1, \dots, C\}_p \cup \{1, \dots, C\}_e$ y $A(x^f, x^s) = \{1, \dots, C - x^f\}_p \cup \{1, \dots, C - x^f\}_e$ respectivamente, donde los conjuntos $\{1, \dots, C\}_p$ y $\{1, \dots, C\}_e$ (resp. $\{1, \dots, C - x^f\}_p$ y $\{1, \dots, C - x^f\}_e$), son los artículos que son producidos por la empresa y los que son comprados a los competidores.

Denotaremos el conjunto donde tomaremos acciones especiales por $D^\wedge := \{\emptyset\}$, es decir, no existe una acción especial obligatoria para el controlador o en otras palabras, la empresa no esta obligada a comprarle artículos a los competidores. Por otro lado nuestro $D^\vee = X$, donde $X = \{0, \dots, C\} \times \{\Lambda\}$. Como se mencionó anteriormente, hemos asumido que x_t^s para cada $t \geq 0$, es una constante igual a Λ , así que podremos asumir que la variable X^s es irrelevante para nuestro problema. Por lo tanto la dinámica del sistema se rige de la siguiente manera.

$$\underbrace{x_{t+1}^f = x_t^f + a_t - \xi_t}_{\text{sub-dinámica estándar}} \quad \text{si } a_t \in \{1, \dots, C - x_t^f\}_p, \quad (4.1)$$

sub-dinámica estándar

$$\underbrace{x_{t+1}^f = x_t^f + a_t}_{\text{sub-dinámica especial}} \quad \text{si } a_t \in \{1, \dots, C - x_t^f\}_e, \quad (4.2)$$

sub-dinámica especial

donde $\{\xi_t\}_t$ es una sucesión de vv.aa. i.i.d definida en el espacio de probabilidad $\mathcal{E} = (\Omega, \mathcal{F}, P)$, con distribución común uniforme μ la cual toma valores en el conjunto $\Delta = [0, C]$ y representa el nivel de demanda del producto.

Asumiremos que la demanda cumple lo siguiente:

$$P(\xi_0 \neq x^f) > 0, \quad \forall x^f \in \mathbb{N}_0, \quad (4.3)$$

esto implica que la propiedad anterior se cumple para cualquier ξ_t pues la sucesión $\{\xi_t\}_t$ es idénticamente distribuida. Note también que la condición anterior (4.3) no es nada fuerte, pues de cumplirse lo contrario, se seguiría que ξ_t es constante (con valor x^f) para toda $t \geq 0$, lo cual trivializa el problema.

Entonces, resumiendo, nuestros elementos quedan de la siguiente manera: $X := \{0, \dots, C\} \times \{\Lambda\}$ (o simplemente $X = \{0, \dots, C\}$), $A = \{1, \dots, C\}_p \cup \{1, \dots, C\}_e$ y $A(x^f, x^s) = \{1, \dots, C - x^f\}_p \cup$

$\{1, \dots, C - x^f\}_e$, $D^\wedge = \{\emptyset\}$, y $D^\vee = X$. De esto último, podemos notar que $X \setminus D^\vee = \emptyset$, es decir, no existe la posibilidad de producir el producto o de comprarlo de forma obligatoria, así en cada instante de tiempo, la empresa tendrá la opción de comprar el artículo o producirlo dependiendo de lo que más le convenga.

Definamos ahora el costo de inventario-producción l , dado por:

$$l(x^f, a) := l_1(x^f)a + l_2x^f, \quad (4.4)$$

Donde l_1 es una función continua no-negativa que representa el costo de producción que depende de la cantidad de artículos en inventario, mientras que $l_2 > 0$ denota el costo de inventario por unidad de producto. Por otra parte, nuestro costo de impulso está definido como sigue:

$$\ell(x^f, a) := \rho(x^f)a, \quad (4.5)$$

donde $\rho(x^f) > 0$ es una función continua que representa el costo del producto a comprar un artículo, este dependerá de la cantidad de artículos se tengan en inventario.

El operador de programación dinámica \mathcal{T} , para cualquier función $u \in \mathbb{L}^+(X)$ y $x^f \in X$, está dado por:

$$\mathcal{T}u(x^f) = \min \{ \mathcal{M}u(x^f), \mathcal{H}u(x^f) \}, \quad \text{para } x^f \in \mathbb{N}_0.$$

donde \mathcal{M} y \mathcal{H} , están definidos como en (3.1) y (3.2) respectivamente. De manera explícita y extendida, se tiene que estos operadores tienen la siguiente expresión

$$\mathcal{M}u(x^f) = \inf_{a \in \{1, \dots, C - x^f\}_e} \{ \rho(x^f)a + u(x^f + a) \} \quad (4.6)$$

$$\mathcal{H}u(x^f) = \inf_{a \in \{1, \dots, C - x^f\}_p} \left\{ l_1(x^f)a + l_2x^f + \alpha(x^f, a) \int_{\Delta} u(x^f + a - \varrho) \mu(d\varrho) \right\}. \quad (4.7)$$

Ahora bien, el problema de control híbrido consiste en hallar una política π^* , que este conformada por una serie de decisiones de producción y compra, tal que (2.11) se satisfaga en este contexto de inventario-producción.

Lo que sigue es verificar que las hipótesis expuestas en esta tesis se satisfacen para este ejemplo. A saber, para $x^f \in X$ y $a \in A(x^f)$, consideremos los costos (4.4) y (4.5), verifiquemos que estas funciones cumplen la Hipótesis 2.1. Por definición, los costos anteriores son no-negativos. Por otro lado, tomando en cuenta que el inventario tiene capacidad máxima igual a C , se puede ver que las funciones de costo están acotadas como sigue

$$l(x^f, a) \leq (l_1(x^f) + l_2)C \leq (M_1 + l_2)C \quad \text{y} \quad \ell(x^f, a) \leq \rho(x^f)C \leq M_2C,$$

En lo anterior hemos usado el hecho que no podemos producir o comprar más artículos una vez el inventario nivel de inventario alcance su capacidad total C , además de considerar a $M_1 = \max_{X^f} l_1(x^f)$ y $M_2 = \max_{X^f} \rho(x^f)$, estos máximos son finitos debido a la continuidad de las funciones en el espacio compacto X^f . Gracias a esto tenemos (4.4) y (4.5) son no-negativas y acotadas.

Por ultimo, definamos,

$$\ell_0 := \min_{X^f} \rho(x^f)$$

Así, se puede deducir fácilmente que, $\ell(x^f, a) \geq \ell_0 > 0$. Por lo tanto las funciones (4.4) y (4.5), verifican la Hipótesis 2.1. Respecto al factor de descuento $\alpha : \mathbb{K} \rightarrow [0, \infty)$, podemos seleccionarlo de tal forma que también cumpla la Hipótesis 2.1(c) y de esta forma, la Hipótesis 2.1 se cumple en su totalidad.

A continuación vamos a verificar que se cumple la Hipótesis 2.3, para ello usaremos la política estacionaria

$$f(x^f) = 0 \in \{1, \dots, C - x_t^f\}_p \quad (4.8)$$

Dada la política anterior, es claro que $D = D^\wedge = \emptyset$, donde D es el conjunto considerado en la Hipótesis 2.3, este último hecho es debido a que no existe $x \in X^f \times \{\Lambda\}$ para la política anteriormente escogida tal que, $f(x^f) \in X^s$. Así pues, se sigue que, existe $m = 1$, $f \in \mathbb{F}$ como en (4.8) y $\varepsilon > 0$ tal que $\mathbf{Q}(D|x^f, f) \leq 1 - \varepsilon$, lo cual garantiza la Hipótesis 2.3.

Por último, verifiquemos que se cumple la Hipótesis 3.1. Demostremos primero que la función costo por etapa $c(\cdot)$, definida por:

$$c(x^f, a) = l(x^f, a)\mathbf{1}_{\{a \in \{0, \dots, C - x^f\}_p\}} + \ell(x^f, a)\mathbf{1}_{\{a \in \{1, \dots, C - x^f\}_e\}}, \quad (4.9)$$

es semi-continua inferior; ésto se reduce a demostrar que tanto l como ℓ son funciones s.c.i en \mathbb{K} . Consideremos $\{(x_n^f, a_n)\}_{n \geq 0}$ una sucesión en \mathbb{K} , la cual converge, es decir, $(x_n^f, a_n) \rightarrow (x^f, a)$ en \mathbb{K} . Entonces

$$l_1(x_n^f)a_n + l_2x_n^f \rightarrow l_1(x^f)a + l_2x^f,$$

Así cuando $n \rightarrow \infty$, $l(x_n^f, a_n) \rightarrow l(x^f, a)$. Como consecuencia, se concluye que l es s.c.i. Un razonamiento análogo, nos llevara a concluir que ℓ es semi-continua inferior. Por lo tanto la función costo por etapa (4.9) es también s.c.i. Por otra parte, podemos considerar de inicio que nuestro factor de descuento $\alpha : \mathbb{K} \rightarrow [0, \infty)$ sea una función s.c.i. y que cumpla las hipótesis correspondientes expuestas en este documento.

Para finalizar, $F(x^f, a, \xi) = x^f + a - \xi$ y $G(x^f, a) = x^f + a$ son funciones continuas en $\mathbb{N}_0 \times A$ para cualquier $\xi \in \Delta$. Bajo estas condiciones y gracias a la Proposición 3.2, las leyes de transición Q^f y Q^s son débilmente continuas.

Una vez que nuestras hipótesis han sido verificadas, los Teoremas 3.8 y 3.9 garantizan la existencia de una política óptima estacionaria $f^* \in \mathbb{F}$ que minimiza el costo total descontado

$$J(x^f, \pi) = E_{x^f}^\pi \sum_{t=0}^{\infty} \left[\left\{ (l_1(x_t^f)a_t + l_2x_t^f)\mathbf{1}_{\{a_t \in \{0, \dots, C - x_t^f\}_p\}} + \rho(x_t^f)a_t\mathbf{1}_{\{a_t \in \{1, \dots, C - x_t^f\}_e\}} \right\} \prod_{i=0}^{t-1} \alpha(x_i^f, a_i) \right].$$

Además el costo óptimo (función de valor) asociado a la esperanza de arriba, es solución de la EPD (3.3).

Para concluir, plantearemos las ecuaciones de programación dinámica para este ejemplo, además de presentar un algoritmo, el cual nos ayudará de forma aproximada a encontrar la soluciones deseadas. Recordemos que para encontrar la solución, en este caso, debemos plantear el siguiente sistema con condición inicial a resolver:

$$u_0(x^f) = 0$$

y

$$u_n(x^f) = \min \{ \mathcal{M}u_{n-1}(x^f), \mathcal{H}u_{n-1}(x^f) \} \quad \text{para } n \geq 1,$$

donde \mathcal{M} y \mathcal{H} estarían definidas como en (4.6) y (4.7). Recordemos que nuestro objetivo, es calcular $\lim_{n \rightarrow \infty} u_n(x^f)$, ya que, por los Teoremas 3.8 y 3.9 este límite es igual a $J^*(x^f)$, que es justo lo que estamos buscando.

Desarrollaremos ahora el algoritmo en el software *Mathematica*, el cual, mediante métodos numéricos nos ayudara a dar una idea de la solución explícita al el sistema anteriormente planteado, dándonos una aproximación de la función de valor.

Con esto en mente y con la finalidad de ilustrar los cálculos, se tomarán como valores para la capacidad de inventario $C = 6$ y el inventario con valores $x^f \in \{0, 1, \dots, 6\}$. Tomaremos también valores fijos para los costos de compra, costos de producción, el costo de inventario y factor de descuento, éstos serán los siguientes: $\rho(x^f) = \{\rho(0), \rho(1), \rho(2), \rho(3), \rho(4), \rho(5), \rho(6)\} = \{5, 4, 3, 2, 1, 1, 0.8\}$, análogamente tendremos a $l_1(x^f) = \{l_1(0), l_1(1), l_1(2), l_1(3), l_1(4), l_1(5), l_1(6)\} = \{5, 4.5, 4, 3, 2, 1, 0.8\}$, $l_2 = 3$ y $\alpha(x^f, a) = \frac{1}{9}$ para todo $(x^f, a) \in \mathbb{K}$.

A continuación se presentara el código correspondiente, con los datos iniciales anteriormente mencionados:

```
Module[{x, Lista},
Lista = {};
rho={5,4,3,2,1,1,0.8};
ele={5,4.5,4,3,2,1,0.8};
For[x = 0, x <= 6, x++, Print[" Para x = ", x, " tenemos que : "];
Inicialslow [j_]:= 0;
Inicialfast [j_]:= 0;
M[j] = Inicialslow [j];
H[j] = Inicialfast [j];
g[j] = Min [M[j],H[j]];
u[j] = g[j];
For[i = 1, i <= 10, i++,
M[j] = MinValue[{(rho[[x]]*a) + N[u[x+a]]; 1 <= a <= 6 - x },{a}];
H[j] = MinValue[{(ele[[x]]*a) + 3x + 1/9 N[(Expectation[u[x+a-\xi],\xi
- UniformDistribution[{0,6}])]}], 1<= a <= 6-x},{a}];
u[j] = Min[M[j],H[j]];
Print[" El operador M[x], en la etapa ", i, " es: ", N[M[j]]];
Print[" El operador H[x], en la etapa ", i, " es: ", N[H[j]]];
Print[" El operador u[x], en la etapa ", i, " es : ", N[u[j]]];
];
AppendTo[Lista,N[u[j]]]; Lista
]; Print[Lista]
].
```

Este algoritmo es de convergencia rápida, ya que, a lo más en la octava iteración se alcanza la convergencia para cada $x \in \{1, \dots, 6\}$. Aunado a que dicho algoritmo nos da el costo óptimo, también logramos obtener las acciones óptimas, esto es gracias a la arreglo Print, el cual en cada iteración nos imprime el valor de los operadores $M[x]$ y $H[x]$ para saber el origen del mínimo. Se presentarán algunas tablas donde se pueden observar los cálculos realizados, éstos estarán basados en la variación de la cantidad de inventario. Se hace notar que en la última columna se especifica el vector óptimo del tipo de dinámica.

Solución Ecuación Programación Dinámica, $x^f = 0$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(0)$	Dinámica óptima
n = 1	5	5	5	f/s
n = 2	10	5.55556	5.55556	(f/s, f)
n = 3	10.5556	5.61728	5.61728	(f/s, f, f)
n = 4	10.6173	5.62414	5.62414	(f/s, f, f, f)
n = 5	10.6241	5.6249	5.6249	(f/s, f, f, f, f)
n = 7	10.6249	5.62499	5.62499	(f/s, f, f, f, f, f)
n = 8	10.625	5.625	5.625	(f/s, f, f, f, f, f, f)
n = 9	10.625	5.625	5.625	(f/s, f, f, f, f, f, f, f)

Tabla 1: Ningún artículo en inventario.

Solución Ecuación Programación Dinámica, $x^f = 1$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(1)$	Dinámica óptima
n = 1	4	4.5	4	s
n = 2	8	4.94444	4.94444	(s, f)
n = 3	8.94444	5.04938	5.04938	(s, f, f)
n = 4	9.04938	5.06104	5.06104	(s, f, f, f)
n = 5	9.06104	5.06234	5.06234	(s, f, f, f, f)
n = 6	9.06234	5.06248	5.06248	(s, f, f, f, f, f)
n = 7	9.06248	5.0625	5.0625	(s, f, f, f, f, f, f)
n = 8	9.0625	5.0625	5.0625	(s, f, f, f, f, f, f, f)
n = 9	9.0625	5.0625	5.0625	(s, f, f, f, f, f, f, f, f)

Tabla 2: Un Artículo en inventario.

Solución Ecuación Programación Dinámica, $x^f = 2$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(2)$	Dinámica óptima
n = 1	3	4	3	s
n = 2	6	4.33333	4.33333	(s, f)
n = 3	7.33333	4.48148	4.48148	(s, f, f)
n = 4	7.48148	4.49794	4.49794	(s, f, f, f)
n = 5	7.49794	4.49977	4.49977	(s, f, f, f, f)
n = 6	7.49977	4.49997	4.49997	(s, f, f, f, f, f)
n = 7	7.49997	4.5	4.5	(s, f, f, f, f, f, f)
n = 8	7.5	4.5	4.5	(s, f, f, f, f, f, f, f)
n = 9	7.5	4.5	4.5	(s, f, f, f, f, f, f, f, f)

Tabla 3: Dos Artículos en inventario.

Solución Ecuación Programación Dinámica, $x^f = 3$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(3)$	Dinámica óptima
n = 1	2	3	2	s
n = 2	4	3.22222	3.22222	(s, f)
n = 3	5.22222	3.35802	3.35802	(s, f, f)
n = 4	5.35802	3.37311	3.37311	(s, f, f, f)
n = 5	5.37311	3.37479	3.37479	(s, f, f, f, f)
n = 6	5.37479	3.37498	3.37498	(s, f, f, f, f, f)
n = 7	5.37498	3.375	3.375	(s, f, f, f, f, f, f)
n = 8	5.375	3.375	3.375	(s, f, f, f, f, f, f, f)
n = 9	5.375	3.375	3.375	(s, f, f, f, f, f, f, f, f)

Tabla 4: Tres Artículos en inventario.

Solución Ecuación Programación Dinámica, $x^f = 4$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(4)$	Dinámica óptima
n = 1	1	2	2	s
n = 2	2	2.11111	2	(s, s)
n = 3	3	2.22222	2.22222	(s, s, f)
n = 4	3.22222	2.24691	2.24691	(s, s, f, f)
n = 5	3.24691	2.24966	2.24966	(s, s, f, f, f)
n = 6	3.24966	2.24996	2.24996	(s, s, f, f, f, f)
n = 7	3.24996	2.25	2.25	(s, s, f, f, f, f, f)
n = 8	3.25	2.25	2.25	(s, s, f, f, f, f, f, f)
n = 9	3.25	2.25	2.25	(s, s, f, f, f, f, f, f, f)

Tabla 5: Cuatro Artículos en inventario.

Solución Ecuación Programación Dinámica, $x^f = 5$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(5)$	Dinámica óptima
n = 1	1	1	1	f/s
n = 2	2	1.11111	1.11111	(f/s, f)
n = 3	2.11111	1.12346	1.12346	(f/s, f, f)
n = 4	2.12346	1.12483	1.12483	(f/s, f, f, f)
n = 5	2.12483	1.12498	1.12498	(f/s, f, f, f, f)
n = 6	2.12498	1.125	1.125	(f/s, f, f, f, f, f)
n = 7	2.125	1.125	1.125	(f/s, f, f, f, f, f, f)
n = 8	2.125	1.125	1.125	(f/s, f, f, f, f, f, f, f)
n = 9	2.125	1.125	1.125	(f/s, f, f, f, f, f, f, f, f)

Tabla 6: Cinco Artículos en inventario.

Solución Ecuación Programación Dinámica, $x^f = 6$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(6)$	Dinámica óptima
n = 1	0.8	0.8	0.8	f/s
n = 2	1.6	0.888889	0.888889	(f/s, f)
n = 3	1.68889	0.898765	0.898765	(f/s, f, f)
n = 4	1.69877	0.899863	0.899863	(f/s, f, f, f)
n = 5	1.69986	0.899985	21.3704	(f/s, f, f, f, f)
n = 6	1.69998	0.899998	21.3745	(f/s, f, f, f, f, f)
n = 7	1.7	0.9	0.9	(f/s, f, f, f, f, f, f)
n = 8	1.7	0.9	0.9	(f/s, f, f, f, f, f, f, f)
n = 9	1.7	0.95	0.9	(f/s, f, f, f, f, f, f, f, f)

Tabla 7: Seis Artículos en inventario.

Para ser más explícitos en los resultados obtenidos tenemos en las primeras columnas los operadores de programación dinámica y en la última la dinámica óptima, esta será vista como vector, donde la componente f indica una acción rápida, es decir, que el mínimo en esa etapa se alcanza en el operador \mathcal{H} , por otro lado, s , indica una acción lenta y entonces el mínimo se alcanzara en el operador \mathcal{M} .

Basándonos en los cálculos anteriores, para los valores de inventario de $0, 1, \dots, 6$, se obtuvieron las siguientes aproximaciones de u :

$$\{u[0], u[1], u[2], u[3], u[4], u[5], u[6]\} = \{5.625, 5.0625, 4.5, 3.375, 2.25, 1.125, 0.9\},$$

la cual es nuestra función de costo óptimo para diferentes niveles de inventario.

La gráfica de esta función se puede ver en la Figura 1 que aparece abajo.

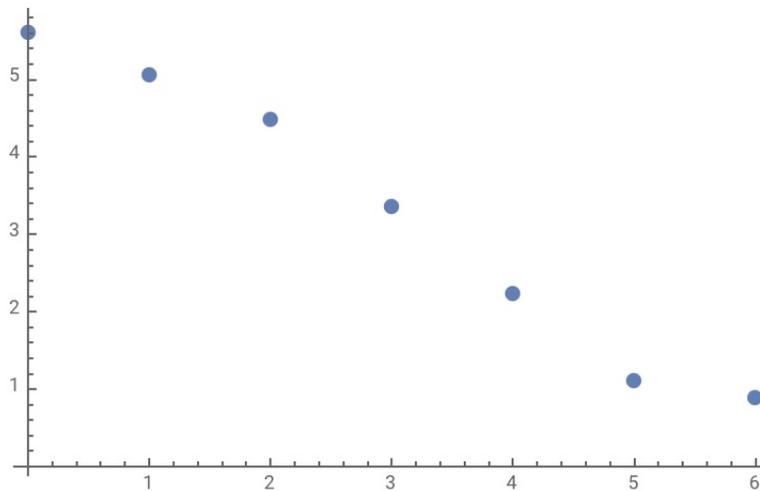


Figura 1: Aproximación de la función $u[x]$, para una capacidad de inventario de 6 artículos

De los resultados obtenidos, podemos deducir que en las primeras etapas del proceso, por lo regular es mejor comprarle a los competidores el artículo, en algunos casos con los datos iniciales dados, tendremos la oportunidad de escoger entre comprar al competidor o producir el artículo, como es el caso de cuando la condición inicial es $x^f = 3$. Una vez que pasen las primeras etapas del proceso (un máximo de tres), lo más conveniente será producir el artículo, ya que éste genera el costo mínimo para ciertas iteraciones.

4.2. Un problema de acumulación de contaminación

Suponga que la economía de un país consume un bien o producto específico y como consecuencia de este consumo se genera contaminación. Vamos a denotar por x_t^f a la cantidad de contaminación generada en el tiempo $t \geq 0$, cuya evolución está representada de la siguiente forma:

$$x_{t+1}^f = x_t^f + \psi(a_t) - g(x_t^s)x_t^f + \xi_t \quad t \geq 0.$$

En la ecuación de arriba, la variable x_t^f , como ya se mencionó, representa la cantidad (o cúmulo) de contaminación en el tiempo t , la cual toma valores en el conjunto $X^f = [0, \infty)$. Por otro lado, la variable x_t^s representa los niveles (o modos) de contingencia ambiental decididos por el gobierno al tiempo t , misma que toma valores en $X^s = \{1, \dots, l\}$. También, consideremos la cantidad $a_t \geq 0$ que denota la tasa de consumo del bien (o producto) en la etapa t cuyo rango es $[0, \gamma(x^s)]$, donde $\gamma(x^s) > 0$ es una constante generalmente impuesta por protocolos internacionales, mientras que la función continua $\psi(a)$ denota la cantidad de contaminación derivada de la tasa de consumo a , donde $\psi : [0, \infty) \mapsto [0, \infty)$. Además, para cada $i \in X^s$, existe una tasa de disminución de la contaminación (tasa de limpieza) asociada al nivel de contingencia i , la cual está representada por la función continua $0 \leq g(i) < 1$. Finalmente en la dinámica de arriba, asumiremos la existencia de una sucesión de variables aleatorias $\{\xi_t\}_t$, que son i.i.d con valores en $\Delta = [0, \infty)$ las cuales miden eventos externos que no se predicen en el modelo. Estas variables tienen distribución uniforme común denotada por μ y cumplen la condición (4.3).

El gobierno también es capaz de cambiar entre los diferentes modos de contingencia ambiental; tales cambios son instantáneos en el tiempo y se aplican a través de una acción de control de un modo $x^s = i$ a un modo $a = j$.

De lo anteriormente mencionado, en cada etapa se tiene la posibilidad de aplicar ya sean acciones de consumo (como se ilustra en la dinámica de arriba) o se pueden aplicar acciones de cambio de modo de contingencia. Todas estas acciones se encuentran en los siguientes conjuntos denominados conjunto de acciones y conjunto de acciones admisibles, respectivamente, los cuales se muestran a continuación. $A = [0, \infty) \cup X^s$ y $A(x^f, x^s) = [0, \gamma(x^s)] \cup [X^s \setminus \{x^s\}]$.

De forma unificada, la dinámica se rige de la siguiente manera:

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = (x_t^f + \psi(a_t) - g(x_t^s)x_t^f + \xi_t, x_t^s)}_{\text{sub-dinámica estándar}} \quad \text{si } a_t \in [0, \gamma(x_t^s)], \quad (4.10)$$

o

$$\underbrace{(x_{t+1}^f, x_{t+1}^s) = (x_t^f, a_t)}_{\text{sub-dinámica especial}} \quad \text{si } a_t \in X^s \setminus \{x_t^s\}, \quad t \geq 0 \quad (4.11)$$

Como ya se mencionó antes, el gobierno también es capaz de cambiar entre los diferentes modos de contingencia ambiental; tales cambios producen un costo de cambiar entre los modos $x^s = i$ a $a = j$, el cual es denotado por la función continua y no-negativa $\ell(i, j)$. También existe para cada nivel i una aplicación continua de desutilidad $D_i(x^f)$ derivada del cúmulo de contaminación x^f . Adicionalmente, se considera un factor de descuento constante para cada nivel $i \in X^s$, éste es denotado por α_i , tal que $\alpha_i < 1$. Así, el factor de descuento no constante está dado de la siguiente manera

$$\alpha(x^f, x^s, a) = \alpha_{x^s} \mathbf{1}_{\{a \in [0, \gamma(x^s)]\}} + \mathbf{1}_{\{a \in X^s \setminus \{x^s\}\}} \quad (4.12)$$

El objetivo es encontrar un política de consumo-cambio de contingencia $\pi = \{a_t\}$ con valores en $[0, \gamma(x^s)] \cup [X^s \setminus \{x^s\}]$, que minimice el costo del tipo (2.9) asociado al costo unitario

$$c(x^f, x^s, a) = D_{x^s}(x^f) \mathbf{1}_{\{a \in [0, \gamma(x^s)]\}} + \ell(x^s, a) \mathbf{1}_{\{a \in X^s \setminus \{x^s\}\}}$$

y factor de descuento (4.12).

Para simplificar las cosas, consideraremos los conjuntos $D^\wedge = \emptyset$ y $D^\vee = [0, \infty) \times X^s$. Antes de comenzar el desarrollo de la ecuación de programación dinámica y verificar las hipótesis expuestas para este ejemplo, nos vemos en la necesidad de hacer la siguiente aclaración. Puesto que $D^\wedge = \emptyset$ y $D^\vee = [0, \infty) \times X^s$, tenemos que el controlador tiene la libertad de escoger la acción a su conveniencia en todo momento. Así, el operador de programación dinámica \mathcal{T} , para cualquier función $u \in \mathbb{L}^+(X)$ y $x = (x^f, x^s) \in X$, está dado por:

$$\mathcal{T}u(x) = \min \{ \mathcal{M}u(x), \mathcal{H}u(x) \}, \quad \text{para } x \in [0, \infty) \times X^s.$$

Donde \mathcal{M} y \mathcal{H} , están definidos como en (3.1) y (3.2) respectivamente. De manera explícita y extendida, se tiene que estos operadores tienen la siguiente expresión

$$\mathcal{M}u(x^f, x^s) = \inf_{a \in X^s \setminus \{x^s\}} \{ \ell(x^s, a) + u(x, a) \},$$

$$\mathcal{H}u(x^f, x^s) = \inf_{a \in [0, \gamma(x^s)]} \left\{ D_{x^s}(x^f) + \alpha_{x^s} \int_{\Delta} u((x^f + \psi(a) - g(x^s)x^f + \varrho, x^s)) \mu(d\varrho) \right\}.$$

A continuación vamos a verificar las hipótesis dadas en esta tesis. Primero veremos que, para $x = (x^f, x^s) \in X$ y $a \in A(x)$, los costos $D_{x^s}(x^f)$ y $\ell(x, a)$, deben satisfacer la Hipótesis 2.1. En efecto, por definición podemos considerar de inicio que las funciones de costo $D_i(x^f)$ y $\ell(x, a)$ son no-negativas, además al ser $\ell(x, a) > 0$ y $X^s (= A^s)$ finito, es fácil ver que existe una constante $0 < \ell_0 = \min \{ \ell(i, j) | i, j \in \{1, \dots, l\} \}$, tal que,

$$\ell(x, a) \geq \ell_0 \quad \forall (x, a) \in X^s \times A^s. \quad (4.13)$$

Ahora, por continuidad de $\ell(x, a)$ en $X^s \times X^s$, ésta es acotada, debido a la compacidad del conjunto anterior, pues este último es finito. Por otro lado, se puede asumir que $D_i(x^f)$ es acotada en X^f , ya que en términos prácticos lo es (i.e, en el mundo real no se tienen costos infinitos de desutilidad). Adicionalmente, el factor de descuento definido como en (4.12) satisface la Hipótesis 2.1, ya que que si $a \in X^s \setminus \{x^s\}$ entonces $\alpha(x^f, x^s, a) = 1$, mientras que si $a \in [0, \gamma(x^s)]$, $\alpha(x^f, x^s, a) = \alpha_{x^s} < 1$, de esta forma, se satisface la Hipótesis 2.1 en su totalidad.

Vamos a verificar que se cumple la Hipótesis 2.3, utilizando una idea parecida al ejemplo anterior, así, consideremos la política estacionaria

$$\mathbf{f}(x^f, x^s) = 0 \quad \text{para todo } x \in X. \quad (4.14)$$

Se sigue fácilmente que $D = \emptyset$, el cual es el conjunto considerado en la Hipótesis 2.3. Trivialmente, tenemos que, existe $m = 1$, $\mathbf{f} \in \mathbb{F}$ como en (4.14) y $\varepsilon > 0$ tal que $\mathbf{Q}(D|x, \mathbf{f}) \leq 1 - \varepsilon$, lo cual garantiza la Hipótesis 2.3.

Por último, verifiquemos que se cumple la Hipótesis 3.1. Demostremos primero que la función costo por etapa $c(\cdot)$, definida por:

$$c(x^f, x^s, a) = D_{x^s}(x^f) \mathbf{1}_{\{a \in [0, \gamma(x^s)]\}} + \ell(x^s, a) \mathbf{1}_{\{a \in X^s \setminus x^s\}}$$

es semi-continua inferior; análogamente al ejemplo anterior, basta demostrar que tanto $D_{x^s}(x^f)$ como $\ell(x^s, a)$ son funciones s.c.i en \mathbb{K} . Consideremos $\{(x_n, a_n)\}_{n \geq 0}$ una sucesión en \mathbb{K} , la cual converge, es decir, $(x_n, a_n) \rightarrow (x, a)$ en \mathbb{K} . Entonces por continuidad

$$\ell(x_n^s, a_n) \rightarrow \ell(x^s, a),$$

Así, se concluye que $\ell(x^s, a)$ es s.c.i. Un razonamiento parecido al anterior, acompañado de la continuidad de $D_{x^s}(x^f)$, nos llevara a concluir que $D_{x^s}(x^f)$ es semi-continua inferior. Por lo tanto, la función costo por etapa $c(\cdot)$ definida anteriormente es también s.c.i. Por otra parte, nuestro factor de descuento α es una función s.c.i. por como se definió, ya que, es la suma de funciones constantes.

Para finalizar, $F(x, a, \xi) = x^f + \psi(a) - g(x^s)x^f + \xi$ y $G(x, a) = (x^f, a)$ son funciones continuas en $[0, \infty) \times X^s \times A(x^f, x^s)$ para cualquier $\xi \in \Delta$. Bajo estas condiciones y gracias a la Proposición 3.2, la ley de transición \mathbf{Q} es débilmente continua.

Una vez que nuestras hipótesis han sido verificadas, los Teoremas 3.8 y 3.9 garantizan la existencia de una política óptima estacionaria $\mathbf{f}^* \in \mathbb{F}$, que es justamente lo que se desea para resolver el problema de control híbrido anteriormente expuesto.

Ahora, procederemos a la aproximación de las correspondientes ecuaciones de programación dinámica. Para simplificar los cálculos, vamos a considerar el costo de cambio de contingencia de la siguiente forma:

$$\ell(x^s, a) = \begin{cases} a & \text{si } x^s = 1 \\ 1 & \text{si } x^s = 2 \\ 1/a & \text{si } x^s = 3 \end{cases}$$

La función anteriormente definida, nos dice lo siguiente: Estar en el modo de contingencia más bajo $x^s = 1$, el costo de cambiar de modo de contingencia es lineal, ya que, aumenta proporcionalmente a medida que cambiamos a niveles de contingencia mayores; en otras palabras, el costo será más alto mientras el modo de contingencia sea más alto. Ahora, si nos encontramos en el modo de contingencia $x^s = 2$ nos generará el mismo costo cambiar a cualquier modo de contingencia; es decir, el costo es constante (en este caso lo consideraremos como 1). Para finalizar, si nos encontramos en el modo de contingencia $x^s = 3$, cambiar a los modos de contingencia anteriores nos costara más cada vez que nos acercamos el primer modo. No está de más resaltar que la anterior función de costo es continua, ya que, en el conjunto donde está definida, es un conjunto discreto y con la topología discreta lograríamos la continuidad de la función, cumpliendo así, las hipótesis anteriormente mencionadas para la misma.

Por otro lado, asumiremos que la función de desutilidad $D_{x^s}(x^f) = a \cdot x^f$; esto quiere decir que para cada nivel de contingencia i , estamos asumiendo que la función de desutilidad D_i es directamente proporcional al nivel de contaminación x^f existente en el momento y a manera de simplificar las cosas, estamos asumiendo que esta desutilidad es la misma en todo nivel de contingencia.

Suponderemos también que algunos elementos de la dinámica (4.10) tienen naturaleza específica. A saber, nuestra tasa de limpieza $g(i)$ estará dada de la siguiente manera: $g(i) = 1 - 1/i$ y la contaminación ψ derivada del consumo a , estará dada por $\psi(a) = ca$, donde c es una constante dada. Con los ingredientes previos, vamos a proponer un algoritmo desarrollado en el software *Mathematica* que nos dará una idea aproximada sobre la solución a las ecuaciones de programación dinámica. Para efectos prácticos, consideraremos una capacidad de contaminación finita, es decir, $0.25 \leq x^f \leq 3$. En este caso consideramos que la contaminación es estrictamente positiva. Definiremos $\gamma(x^s) = 10$ para todo $x^s \in \{1, 2, 3\}$, $c = 3$ y consideraremos que los factores externos ξ tienen distribución uniforme en $[0, 10]$. Para el desarrollo del algoritmo en *Mathematica* discretizaremos el intervalo $[0.25, 3]$ como sigue:

$$int = \{0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, 2.75, 3\};$$

es decir, partiremos la unidad de manera uniforme en cuartos. Cabe resaltar que la partición se puede efectuar de múltiples formas pero en este caso escogeremos esta. Resumiendo lo anteriormente dicho, presentaremos el siguiente algoritmo en donde, $l(x^s, a) = \text{elexsa}$, $\alpha(x^s) = \text{alpha}$.

```
Module [{ x, Lista },
Lista = { };
elexsa = {a, 1, 1/a};
alpha = {0.5, 0.4, 0.7};
int = {0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, 2.75, 3};
For [xs = 1 , xs <= 3, xs++, Print[" Nivel de contingencia ", xs , " Se tiene: "];
Axs = Table[xs, {xs, 1, 3}];
y = Delete[Axs, xs];
For[ x = 1 , x <= 12 , x++ , Print[" Para xf = ", int[[x]] , " tenemos que : "];
Inicialslow [j_]:= 0;
Inicialfast [j_]:= 0;
M[j] = Inicialslow [j];
H[j] = Inicialfast [j];
g[j] = Min [M[j],H[j]];
u[j] = g[j];
For[ k = 1, k <= 10 , i++,
M[j] = MinValue[{elexsa[[xs]] + N[u[int[[x]],a]]; y[[1]] <= a <= y[[2]]},{a}];
H[j] = MinValue[{(a*int[[x]]) + (alpha[[xs]) N[(Expectation[u[int[[x]]+3a+(1-1/i)
int[[x]]+\xi],\xi- UniformDistribution[{0,10}]]), 1<= a <= 10},{a}];
u[j] = Min[M[j],H[j]];
Print[" El operador M[x], en la etapa ", k , " es: ", N[M[j]]];
Print[" El operador H[x], en la etapa ", k , " es: ", N[H[j]]];
Print[" El operador u[x], en la etapa ", k , " es : ", N[u[j]]];
];
AppendTo[Lista,N[u[j]]]; Lista
]]; Print[Lista]
].
```

La convergencia del algoritmo es un poco más lenta, ya que, máximo en 50 iteraciones se llega al objetivo deseado para ciertos casos de contaminación y modos de contingencia.. Sin embargo, a medida que aumentemos la capacidad máxima de contaminación permitida por consumir un bien o servicio por un país, el número de iteraciones necesarias para llegar a la función objetivo va aumentando, aún así, la convergencia es obtenida en no más de 50 iteraciones. En ambos ejemplo buscamos entender el comportamiento de los operadores asociados a la programación dinámica y estos algoritmos aunque no sean exactos nos permiten dar una intuición sobre la convergencia y la naturaleza de solución a las ecuaciones. Al tener una cantidad de datos relativamente alta, las tablas presentarán de forma resumida los cálculos obtenidos en el algoritmo, para ciertas aproximaciones. Al igual que el ejemplo anterior, en las primeras columnas de las tablas se mostrarán los operadores de programación dinámica y en la ultima se presentará la dinámica óptima, esta será vista como vector, donde la componente f indica una acción rápida, es decir, que el mínimo en esa etapa se alcanza en el operador \mathcal{H} , en cuanto a s , indica una acción lenta y el mínimo se alcanzara en el operador \mathcal{M} .

Se tomarán algunos casos específicos para las tablas presentadas, mismos que se eligieron de tal forma que se pueda ver una evolución “coherente” del proceso en cada iteración. Algunas de las siguientes tablas, presentarán los datos cada dos o cinco iteraciones dependiendo del caso, debido a que, para ciertos valores de x^f y x^s la convergencia del proceso varía. Sin más preámbulo a continuación se prestarán los resultados obtenidos.

Nivel de contingencia, $x^s = 1, x^f = 0.25$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(0.25, 1)$	Dinámica óptima
n = 1	2	0.25	0.25	f
n = 2	2.25	0.375	0.375	(f, f)
n = 4	2.4375	0.46875	0.499512	(f, \dots, f)
n = 6	2.48438	0.492188	0.492188	(f, f, \dots, f)
n = 8	2.49609	0.498047	0.498047	(f, f, f, \dots, f)
n = 10	2.49902	0.499512	0.499512	(f, f, f, f, \dots, f)
n = 12	2.49976	0.499878	0.499878	$(f, f, f, f, f, \dots, f)$
n = 14	2.49994	0.499969	0.499969	$(f, f, f, f, f, f, \dots, f)$
n = 16	2.49998	0.499992	0.499992	$(f, f, f, f, f, f, f, \dots, f)$
n = 18	2.5	0.499998	0.499998	$(f, f, f, f, f, f, f, f, \dots, f)$
n = 20	2.5	0.5	0.5	$(f, f, f, f, f, f, f, f, f, \dots, f)$

Tabla 8: Nivel de contingencia 1.

Nivel de contingencia, $x^s = 1, x^f = 1$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(1, 1)$	Dinámica óptima
n = 1	2	1	1	f
n = 2	3	1.5	1.5	(f, f)
n = 4	3.75	1.875	1.875	(f, \dots, f)
n = 6	3.9375	1.96875	1.96875	(f, f, \dots, f)
n = 8	3.98438	1.99219	1.99219	(f, f, f, \dots, f)
n = 10	3.99609	1.99805	1.99805	(f, f, f, f, \dots, f)
n = 12	3.99902	1.99951	1.99951	$(f, f, f, f, f, \dots, f)$
n = 14	3.99976	1.99988	1.99988	$(f, f, f, f, f, f, \dots, f)$
n = 16	3.99994	1.99997	1.99997	$(f, f, f, f, f, f, f, \dots, f)$
n = 18	3.99998	1.99999	1.99999	$(f, f, f, f, f, f, f, f, \dots, f)$
n = 20	4	2	2	$(f, f, f, f, f, f, f, f, f, \dots, f)$

Tabla 9: Nivel de contingencia 1.

Nivel de contingencia, $x^s = 1, x^f = 2$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(2, 1)$	Dinámica óptima
n = 1	2	2	2	f/s
n = 2	4	3.	3.	$(f/s, f)$
n = 4	5.5	3.75	3.75	$(f/s, \dots, f)$
n = 6	5.875	3.9375	3.9375	$(f/s, f, \dots, f)$
n = 8	5.96875	3.98438	3.98438	$(f/s, f, f, \dots, f)$
n = 10	5.99219	3.99609	3.99609	$(f/s, f, f, f, \dots, f)$
n = 12	5.99805	3.99902	3.99902	$(f/s, f, f, f, f, \dots, f)$
n = 14	5.99951	3.99976	3.99976	$(f/s, f, f, f, f, f, \dots, f)$
n = 16	5.99988	3.99994	3.99994	$(f/s, f, f, f, f, f, f, \dots, f)$
n = 18	5.99997	3.99998	3.99998	$(f/s, f, f, f, f, f, f, f, \dots, f)$
n = 20	6	4	4	$(f/s, f, f, f, f, f, f, f, f, \dots, f)$

Tabla 10: Nivel de contingencia 1.

Nivel de contingencia, $x^s = 1, x^f = 3$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(3, 1)$	Dinámica óptima
n = 1	2	3	3	s
n = 2	4	4	4	(s, s)
n = 4	7	5.5	5.5	(s, s, f, f)
n = 6	7.75	5.875	5.875	(s, s, f, f, ..., f)
n = 8	7.9375	5.96875	5.96875	(s, s, f, f, ..., f)
n = 10	7.98438	5.99219	5.99219	(s, s, f, f, f, ..., f)
n = 12	7.99609	5.99805	5.99805	(s, s, f, f, f, ..., f)
n = 14	7.99902	5.99951	5.99951	(s, s, f, f, f, f, ..., f)
n = 16	7.99976	5.99988	5.99988	(s, s, f, f, f, f, f, ..., f)
n = 18	7.99994	5.99997	5.99997	(s, s, f, f, f, f, f, f, ..., f)
n = 20	7.99998	5.99999	5.99999	(s, s, f, f, f, f, f, f, f, ..., f)
n = 22	8	6	6	(s, s, f, f, f, f, f, f, f, f, ..., f)

Tabla 11: Nivel de contingencia 1.

Nivel de contingencia, $x^s = 2, x^f = 0.25$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(0.25, 2)$	Dinámica óptima
n = 1	1	0.25	0.25	f
n = 2	1.25	0.35	0.35	(f, ..., f)
n = 4	1.39	0.406	0.406	(f, ..., f)
n = 6	1.4124	0.41496	0.41496	(f, f, ..., f)
n = 8	1.41598	0.416394	0.416394	(f, f, f, ..., f)
n = 10	1.41656	0.416623	0.416623	(f, f, f, f, ..., f)
n = 12	1.41665	0.41666	0.41666	(f, f, f, f, f, ..., f)
n = 14	1.41666	0.416666	0.416666	(f, f, f, f, f, f, ..., f)
n = 16	1.41667	0.416667	0.416667	(f, f, f, f, f, f, f, ..., f)
n = 18	1.41667	0.416667	0.416667	(f, f, f, f, f, f, f, f, ..., f)
n = 20	1.41667	0.416667	0.416667	(f, f, f, f, f, f, f, f, f, ..., f)

Tabla 12: Nivel de contingencia 2.

Nivel de contingencia, $x^s = 2, x^f = 1$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(1, 2)$	Dinámica óptima
n = 1	1	1	1	f/s
n = 2	2	1.4	1.4	(f/s, f)
n = 4	2.56	1.624	1.624	(f/s, f, f, f)
n = 6	2.6496	1.65984	1.65984	(f/s, f, ..., f)
n = 8	2.66394	1.66557	1.66557	(f/s, f, f, ..., f)
n = 10	2.66623	1.66649	1.66649	(f/s, f, f, f, ..., f)
n = 12	2.6666	1.66664	1.66664	(f/s, f, f, f, f, ..., f)
n = 14	2.66666	1.66666	1.66666	(f/s, f, f, f, f, f, ..., f)
n = 16	2.66666	1.66667	1.66667	(f/s, f, f, f, f, f, f, ..., f)
n = 18	2.66667	1.66667	1.66667	(f/s, f, f, f, f, f, f, f, ..., f)
n = 20	2.66667	1.66667	1.66667	(f/s, f, f, f, f, f, f, f, f, ..., f)

Tabla 13: Nivel de contingencia 2.

Nivel de contingencia, $x^s = 2, x^f = 2$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(2, 2)$	Dinámica óptima
n = 1	1	2	1	s
n = 2	2	2.4	2	(s, s)
n = 4	3.8	3.12	3.12	(s, s, f, f)
n = 6	4.248	3.2992	3.2992	(s, s, f, ..., f)
n = 8	4.31968	3.32787	3.32787	(s, s, f, ..., f)
n = 10	4.33115	3.33246	3.33246	(s, s, f, f, ..., f)
n = 12	4.33298	3.33319	3.33319	(s, s, f, f, f, ..., f)
n = 14	4.33328	3.33331	3.33331	(s, s, f, f, f, f, ..., f)
n = 16	4.33331	3.33332	3.33332	(s, s, f, f, f, f, f, ..., f)
n = 18	4.33333	3.33333	3.33333	(s, s, f, f, f, f, f, f, ..., f)
n = 20	4.33333	3.33333	3.33333	(s, s, f, f, f, f, f, f, f, ..., f)

Tabla 14: Nivel de contingencia 2.

Nivel de contingencia, $x^s = 2, x^f = 3$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(3, 2)$	Dinámica óptima
n = 1	1	3	1	s
n = 2	2	3.4	3.4	(s, s)
n = 4	4	4	4	(s, s, s, f/s)
n = 6	5.6	4.84	4.84	(s, s, s, f/s, f, f)
n = 8	5.936	4.9744	4.9744	(s, s, s, f/s, f, ..., f)
n = 10	5.98976	4.9959	4.9959	(s, s, s, f/s, f, ..., f)
n = 12	5.99836	4.99934	4.99934	(s, s, s, f/s, f, ..., f)
n = 14	5.99974	4.9999	4.9999	(s, s, s, f/s, f, f, ..., f)
n = 16	5.99996	4.99998	4.99998	(s, s, s, f/s, f, f, f, ..., f)
n = 18	5.99999	5	5	(s, s, s, f/s, f, f, f, f, ..., f)
n = 20	6	5	5	(s, s, s, f/s, f, f, f, f, f, ..., f)

Tabla 15: Nivel de contingencia 2.

Nivel de contingencia, $x^s = 3, x^f = 0.25$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(0.25, 3)$	Dinámica óptima
n = 1	0.5	0.25	0.25	f
n = 5	0.75	0.325	0.325	(f, ..., f)
n = 10	1.29971	0.809794	0.809794	(f, ..., f)
n = 15	1.32768	0.829377	0.829377	(f, f, ..., f)
n = 20	1.33238	0.832668	0.832668	(f, f, f, ..., f)
n = 25	1.33317	0.833222	0.833222	(f, f, f, f, ..., f)
n = 30	1.33331	0.833315	0.833315	(f, f, f, f, f, ..., f)
n = 35	1.33333	0.83333	0.83333	(f, f, f, f, f, f, ..., f)
n = 40	1.33333	0.83333	0.83333	(f, f, f, f, f, f, f, ..., f)
n = 45	1.33333	0.83333	0.83333	(f, f, f, f, f, f, f, f, ..., f)
n = 50	1.33333	0.83333	0.83333	(f, f, f, f, f, f, f, f, f, ..., f)

Tabla 16: Nivel de contingencia 3.

Nivel de contingencia, $x^s = 3, x^f = 0.5$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(0.5, 3)$	Dinámica óptima
n = 1	0.5	0.5	0.5	f/s
n = 5	1.7665	1.38655	1.38655	$(f/s, f, \dots, f)$
n = 10	2.09941	1.61959	1.61959	$(f/s, f, \dots, f)$
n = 15	2.15536	1.65875	1.65875	$(f/s, f, \dots, f)$
n = 20	2.16477	1.66534	1.66534	$(f/s, f, f, \dots, f)$
n = 25	2.16635	1.66644	1.66644	$(f/s, f, f, f, \dots, f)$
n = 30	2.16661	1.66663	1.66663	$(f/s, f, f, f, f, \dots, f)$
n = 35	2.16666	1.66666	1.66666	$(f/s, f, f, f, f, f, \dots, f)$
n = 40	2.16667	1.66667	1.66667	$(f/s, f, f, f, f, f, f, \dots, f)$
n = 45	2.16667	1.66667	1.66667	$(f/s, f, f, f, f, f, f, f, \dots, f)$
n = 50	2.16667	1.66667	1.66667	$(f/s, f, f, f, f, f, f, f, f, \dots, f)$

Tabla 17: Nivel de contingencia 3.

Nivel de contingencia, $x^s = 3, x^f = 1.5$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(1.5, 3)$	Dinámica óptima
n = 1	0.5	1.5	0.5	s
n = 5	2.5	2.9	2.5	(s, s, s, s, f)
n = 10	4.765	4.4855	4.4855	$(s, s, s, s, f, \dots, f)$
n = 15	5.37647	4.91353	4.91353	$(s, s, s, s, f, \dots, f)$
n = 20	5.47924	4.98547	4.98547	$(s, s, s, s, f, \dots, f)$
n = 25	5.49651	4.99756	4.99756	$(s, s, s, s, f, \dots, f)$
n = 30	5.49941	4.99959	4.99959	$(s, s, s, s, f, \dots, f)$
n = 35	5.4999	4.99993	4.99993	$(s, s, s, s, f, f, \dots, f)$
n = 40	5.49999	4.99999	4.99999	$(s, s, s, s, f, f, f, \dots, f)$
n = 45	5.5	5	5	$(s, s, s, s, f, f, f, f, \dots, f)$
n = 50	5.5	5	5	$(s, s, s, s, f, f, f, f, f, \dots, f)$

Tabla 18: Nivel de contingencia 3.

Nivel de contingencia, $x^s = 3, x^f = 2.5$				
Cantidad de Iteraciones	$M(x)$	$H(x)$	$u_n(2.5, 3)$	Dinámica óptima
n = 1	0.5	3	0.5	s
n = 5	2.5	4.4	2.5	(s, s, s, s, s)
n = 10	5	6.15	5	$(s, s, s, s, s, \dots, s)$
n = 15	7.5	7.9	7.5	$(s, s, s, s, s, \dots, s)$
n = 20	9.765	9.4855	9.4855	$(s, s, s, s, s, \dots, f/s, f, f)$
n = 25	10.3765	9.91353	9.91353	$(s, s, s, s, s, \dots, f/s, f, \dots, f)$
n = 30	9.98547	9.98547	9.98547	$(s, s, s, s, s, \dots, f/s, f, \dots, f)$
n = 35	10.4965	9.99756	9.99756	$(s, s, s, s, s, \dots, f/s, f, \dots, f)$
n = 40	10.4994	9.99959	9.99959	$(s, s, s, s, s, \dots, f/s, f, \dots, f)$
n = 45	10.4999	9.9999	9.9999	$(s, s, s, s, s, \dots, f/s, f, \dots, f)$
n = 50	10.5	10	10	$(s, s, s, s, s, \dots, f/s, f, \dots, f)$

Tabla 19: Nivel de contingencia 3.

A diferencia del ejemplo anterior donde nuestra solución a la ecuación de programación dinámica no depende de x^s (ya que el espacio X^s tenía un solo elemento), este ejemplo sí depende del espacio X^s , por tanto, la solución aproximada será mostrada en términos matriciales, como se presenta a continuación.

$$u(x^f, x^s) = \begin{pmatrix} 0.5 & 0.416667 & 0.833333 \\ 1 & 0.833333 & 1.66667 \\ 1.5 & 1.25 & 2.5 \\ 2 & 1.66667 & 3.33333 \\ 2.5 & 2.08333 & 4.16667 \\ 3 & 2.5 & 5 \\ 3.5 & 2.91667 & 5.83333 \\ 4 & 3.33333 & 6.66667 \\ 4.5 & 3.75 & 7.5 \\ 5 & 4.16667 & 8.33333 \\ 5.5 & 4.58333 & 9.16678 \\ 6 & 5 & 10 \end{pmatrix}$$

En cuanto al conjunto de acciones óptimas, para ciertos niveles de contaminación, este conjunto se encuentra en la última columna de las tablas anteriores. Como conclusión de los resultados obtenidos, podemos verificar que a medida que aumenta la cantidad de contaminación permitida, en las primeras etapas del proceso, este dirá que debemos cambiar de modo de contingencia para disminuir el nivel de contaminación, una vez ocurra esto podremos consumir el bien o servicio que produce contaminación, hasta que vuelva y el nivel de contaminación aumente. Los resultados obtenidos, nos acercan un poco más a la solución real del problema anteriormente expuesto, la aproximación lograda es buena y de cierta manera nos ayuda a entender cómo poder obtener las soluciones de forma explícita, ésta última es una tarea un poco más compleja, debido a la cantidad de casos en cada iteración que se deben analizar, aun así, este es un buen comienzo, para el desarrollo general de las ecuaciones de programación dinámica para ciertos casos específicos . En este ejemplo se optó por no poner gráficas de la función objetivo, puesto que está compuesta por una matriz. Con estos cálculos finalizamos nuestra tesis. En el siguiente capítulo, daremos las conclusiones correspondientes de nuestro documento.

5. Conclusiones

En esta tesis analizamos un problema general de control estocástico de tipo híbrido en tiempo discreto. El análisis consistió básicamente en proponer el modelo general y dar solución de éste a través de la herramienta de programación dinámica.

El modelo que se consideró involucra una gran cantidad de modelos particulares, ya que los espacios de estados y de acciones fueron definidos de manera general como espacios de Borel y la dinámica aún cuando se definió de manera fuerte a través de una ecuación en diferencias, se hizo de manera general, incluyendo un vasto conjunto de casos especiales. La función objetivo se consideró de tipo descontado en horizonte infinito, misma que tiene bastante uso en diversas aplicaciones.

Se ilustraron también dos aplicaciones: una referente a un problema de inventario-producción y otra que tuvo que ver con un problema de consumo-contaminación en donde existen diferentes tipos de contingencias ambientales.

Este trabajo puede ser de gran utilidad para las personas que desean estudiar por primera vez de una manera rigurosa problemas de control híbrido ya que, por un lado, el tiempo discreto da una gran intuición de lo que está pasando en cada etapa y por el otro, agregamos el ingrediente de tener explícitamente la evolución del estado del sistema mediante una dinámica de manera precisa. También es importante señalar que la teoría fue aterrizada a dos aplicaciones que consideramos interesantes, las cuales ya fueron mencionadas con anterioridad. En resumen, pensamos que el trabajo cubre ambas caras: el rigor matemático que involucró la teoría de la tesis y las aplicaciones que se derivaron de la misma. Respecto a las aplicaciones, se logró ver cómo el algoritmo expuesto logró aproximar de forma satisfactoria las soluciones de las ecuaciones de programación dinámica, además de ayudarnos a dar una idea de la rapidez de convergencia de la solución; aunque no se alcanzó a dar la solución de forma explícita, sí se pudo dar una idea intuitiva y aproximada de la solución real. Se espera que en futuros trabajos podamos crear un algoritmo más preciso con el fin de dar la solución explícita de estas aplicaciones.

Para concluir, cabe mencionar que este trabajo puede ser extendido al mismo análisis cuando se tienen otro tipo de criterios de optimalidad, tales como el criterio promedio o el criterio sensible al riesgo. También se puede extender a problemas cuya dinámica estándar es de tiempo continuo. Algunos de estos trabajos han sido estudiados con anterioridad, pero creemos que se han quedado en resultados teóricos y no se ha hecho algún trabajo que involucre ambas facetas: aspectos teóricos y aplicaciones.

Referencias

- [1] V. Azhmyakov, V. Boltyanski, and A. Poznyak. Optimal control of impulsive hybrid systems. *Non-linear Anal. Hybrid Systems*, 2:1089–1097, 2008.
- [2] G. Barles, S. Dharmatti, and M. Ramaswamy. Unbounded viscosity solutions of hybrid control systems. *ESAIM: Control Optim. Calc. Var.*, 16:176–193, 2010.
- [3] A. Bensoussan. Dynamic programming and inventory control. *IOS Press, Amsterdam*, 2011.
- [4] A. Bensoussan and J. Menaldi. Hybrid control and dynamic programming. *Dynam. Contin. Discrete Impulsive System*, 3:395–442, 1997.
- [5] D. Bertsekas and S. Shreve. Stochastic optimal control : The discrete-time case. *Athena Scientific, Belmont, MA*, 1996.
- [6] V. Borkar, M. Ghosh, and P. Sahay. Optimal control of a stochastic hybrid system with discounted cost. *J. Optim. Theory Appl.*, 101:557–580, 1999.
- [7] M. Branicky, V. Borkar, and S. Mitter. A unified framework of hybrid control: model and optimal control theory. *IEEE Trans. Autom. Control*, 43:31–45, 1998.
- [8] S. Dharmatti and M. Ramaswamy. Hybrid control systems and viscosity solutions. *SIAM J. Control Optim.*, 44:1259–1288, 2005.
- [9] O. Hernández-Lerma and J. Lasserre. Discrete-time markov control processes. *Springer-Verlag, New York, Basic optimality criteria*, 1996.
- [10] C. Himmelberg, T. Parthsarathy, and F. V. Vleck. Optimal plans for dynamic programming problems. *Math. Oper. Res.*, 1:390–394, 1976.
- [11] H. Jasso-Fuentes, J. Menaldi, and T. Prieto-Rumeau. Discrete-time hybrid control in borel spaces. *Appl. Math. Optim.*, pages 1–33, 2018.
- [12] H. Jasso-Fuentes, J. Menaldi, T. Prieto-Rumeau, and M. Robin. Discrete-time hybrid control in borel spaces: Average cost optimality criterion. *J. Math. Anal. Appl.*, 2:1695–1713, 2018.
- [13] K. Kawaguchi and H. Morimoto. Long-run average welfare in a pollution accumulation model. *J. Econom. Dynam. Control*, 31:703–720, 2007.
- [14] J. Lygeros. An overview of hybrid systems control. *in: Handbook of Networked and Embedded Control Systems*, pages 519–537, 2005.
- [15] M. Posa, C. Cantu, and R. Tedrake. A direct method for trajectory optimization of rigid bodies through contact. *Int. J. Robotics Res.*, 33:69–81, 2014.
- [16] M. Puterman. Markov decision processes : Discrete stochastic dynamic programming. *Wiley, New York*, 1994.
- [17] P. F. Riedinger, C. Iung, and F. Kratz. An optimal control approach for hybrid systems. *Eur. J. Control*, 9:449–458, 2003.

- [18] M. S. Shaikh and P. E. Caines. On the hybrid optimal control problem:theory and algorithms. *IEEE Trans. Autom. Control*, 52:1587–1603, 2007.
- [19] M. Soler, A. Olivares, and E. Staffetti. Hybrid optimal control to commercial aircraft trajectory planning. *J. Guidance Control Dynamic*, 33:985–991, 2010.
- [20] L. Tavernini. Differential automata and their discrete simulators. *Nonlinear Anal.*, 11:665–683, 1987.