Técnicas de Operadores Contractivos en estabilidad de procesos controlados de Markov.
Versión Γ

Asesor: Dr. Evgueni Gordienko
Doctorando: M. en C. Enrique Lemus-Rodríguez

12 de abril de 2012

El presente texto es la última versión de la tesis doctoral que bajo la dirección del Prof. E. Gordienko se presentará para optar por el título de Doctor. Consta de un capítulo introductorio sobre el tema que se espera sea accesible a un amplio público, un capítulo que recapitula el proyecto doctoral, y los capítulos intermedios que constituyen el cuerpo de la tesis y se refieren a los tres artículos ya publicados con el Prof. E. Gordienko a lo largo del trabajo doctoral. Esta versión se remata con un capítulo de conclusiones en las que se presentan algunos problemas propuestos surgidos tanto de los tres artículos antes mencionados como de investigaciones recientes publicadas junto con el Dr. José Raúl Montes de Oca.

Índice general

1. Introducción **5**
1.1. De la mejor manera 5
1.2. Operadores contractivos. 8
1.3. Métricas Probabilísticas 9
1.4. Estabilidad en Control 10
1.5. Agradecimientos 11
1.6. Dedicatoria 11

2. Operadores Contractivos y Estabilidad **13**
2.1. Introducción 13
2.2. Convergencia de Procesos de Markov a tiempo discreto y métricas probabilísticas. 13
2.3. Métricas, contracciones y estabilidad de procesos de Markov controlados. 14
2.4. Objetivo 14
2.5. Problemas Propuestos 15
2.5.1. Problema 1 15
2.5.2. Problema 2 16
2.5.3. Problema 3 17
2.5.4. Problema 4 17
2.6. Comentarios Generales. 18

3. Caso Costo Descontado **19**
3.1. Planteamiento del Problema 20
3.2. Hipótesis y Resultados 23
3.3. Ejemplos de aplicación de las Desigualdades de Estabilidad . 29
3.3.1. Modelo a tiempo discreto de un problema de optimización de portafolio de consumo en inversión 29

3.3.2. Un sistema de inventarios 33
3.4. Demostraciones 34
3.4.1. Demostración del Teorema 1 34
3.4.2. Demostración del Teorema 2. 38
3.4.3. Demostración del Teorema 3 40

4. Caso Costo Promedio 45
4.1. El caso promedio 46
4.2. Condiciones y resultado principal 49
4.3. Contraejemplos 52
4.4. Estabilidad en Control Autoregresivo 54
4.5. Demostración del Teorema 59

5. Conclusiones 71
5.1. La Métrica de Kantorovich 71
5.2. Consideraciones Finales 73

Capítulo 1

Introducción

1.1. De la mejor manera

El uso de los métodos matemáticos de optimízación se remonta a tiempos de los griegos, cuando Herón analizó sencillos pero significativos problemas de optimización en relación con sus estudios de óptica, tema que retomó Fermat. A partir de entonces, el uso de métodos variacionales en la física se hizo natural a través del Principio de la Mínima Acción y justificó el gran interés y auge del asi llamado Cálculo de Variaciones. Pero en el siglo XX, cuando parecía agotado el interés de los matemáticos en el cálculo de variaciones, la necesidad de aplicar las matemáticas más allá de la física, a problemas de optimización en la industria, la aeronáutica y la economía llevó a la renovación de la disciplina con la creación de nuevos campos de las matemáticas como la Programación Lineal, la Programación Convexa, la Programación Dinámica y el Control Óptimo. En esta nueva fase del desarrollo de la optimización, el énfasis pasó del estudio de la manera en la que la naturaleza espontáneamente llegaba a un óptimo (por ejemplo, cuando, según el Principio de Fermat, la luz recorre las trayectorias de tiempo mínimo) hasta el caso en el que es un agente, un decisor, el que tiene que llegar a determinar la secuencia de decisiones que le permitan alcanzar el óptimo. Por ejemplo, lograr el alunizaje de un módulo lunar con el mínimo gasto de combustible, realizar un recorrido de un punto de partida a un punto de llegada en tiempo mínimo, invertir de manera que se logre el mejor consumo a lo largo de un período dado de tiempo, etc.

Realizando un escueto esbozo histórico, quizás el inicio de este cambio de modo de pensar se dió tanto por las terribles condiciones en la dos guerras mundiales del siglo XX, durante las cuales era imperioso sacar el máximo provecho de los escasos recursos, y el desarrollo de ramas de las matemáticas abstractas como el análisis funcional, el cual permitió unificar en un esquema abstracto las muy diversas técnicas que estaban dispersas en ramas diversas como el cálculo (regla de Fermat, Teoremas de la Función Inversa y la Función Implícita, Multiplicadores de Lagrange, Método de Newton), las ecuaciones diferenciales, el álgebra matricial, el cálculo de variaciones, etc. Esta combinación permitíó el surgimiento a mediados del siglo XX de la Programación Lineal, de manera paralela en la entonces Unión Soviética, con la obra de Kantoróvich (quien resultará un personaje muy importante en esta tesis) y en los Estados Unidos con la obra de Dantzig. El método Símplex de Dantzig tuvo un impacto inmediato en el cálculo práctico de los programas lineales óptimos, lo que obscureción en Occidente las aportaciones de Kantorovich (en particular, al replanteamiento de los Problemaas de Transporte de Masa), que ahora reciben cada vez más atención fuera de Rusia.

En épocas y por razones similares, R. Bellman abandonó una carrera en Teoría Analítica de los Números, para atacar problemas de optimización por medio del método que él bautizó por interesantes razones Programación Dinámica (en particular, para lograr apoyo de autoridades estadounidenses adversas a las matemáticas). Aprovechando el desarrollo de las computadoras digitales, logró plantear un extenso programa de investigación alrededor de dicha técnica, analizando muy diversos problemas en control. En particular, logró formular el problema de manera abstracta introduciendo una ecuación funcional para una función asociada al problema, que de ser conocidad, lo permite resolver (la función de Valor), logrando así dar una formulación matemática precisa a la técnica de inducción retrógrada.

Motivado por Seminarios a los cuales se invitaban a ingenieros, Pontryagin y su escuela lograron plantear una generalización del Cálculo clásico de Variaciones, que en cierto sentido es un método infinito-dimensional de la familiar regla de los Multiplicadores de Lagrange, y que ahora se conoce como Principio del Máximo de Pontryagin, que se aplica a problemas de control óptimo. Una aplicación clásica es la determinación de la manera óptima $u(t)$ de utilizar el cohete inferior de un módulo lunar, de manera que este alunize suavemente con un gasto mínimo de combustible.

Vale la pena señalar que el primer problema de Cálculo de Variaciones que recibió atención general fue el problema de la Braquistocrona de Bernoulli, problema que ya había planteado Galileo al analizar planos inclinados: ¿cuál es la trayectoria $C(u)$ (un tobogán) de un punto A a un punto B a un nivel inferior, tal que permita a una canica que inicia en A llegar a B en un tiempo mínimo? Diversos matemáticos, como Newton y Leibniz lograron resolver el problema. Un método, por cierto, consistía en relacionar este problema con el Principio de Fermat en óptica. Posteriormente Euler y Lagrange plantearon métodos generales que permitieron iniciar el Cálculo de Variaciones propiamente dicho. Sin embargo, fue cuestión de tiempo descubrir que el Principio del Máximo también se aplica a este problema, por lo cual a veces se le considera con el primer problema de Control Óptimo propiamente dicho.

Una familia natural de problemas de optimización analiza aquellos casos en los que interviene el azar, dando origen a diversas variantes hermanadas en conceptos, pero diferenciadas por detalles de modelo o por las comunidades que los analizaron: Procesos de Decisión Markovianos, Procesos controlados de Markov, Control Óptimo de Difusiones, etc. En el caso en el que el modelo considera tiempo discreto es posible utilizar directamente el método de Programación Dinámica si el número de etapas es finito. Cuando el proceso es a tiempo discreto, pero con un número infinito de etapa, modelo que tiene mucho sentido en la práctica en varios campos (por ejemplo, computación, economía) donde el número de etapas es mucho muy grande, estrictamente hablando no se puede aplicar Programación Dinámica, que al ser un tipo de inducción retrógrada, precisamente comienza etapa por etapa, desde la etapa final hasta llegar a la primer etapa. En ese contexto, varios investigadores propusieron métodos iterativos inspirados en Programación Dinámica, conocidos como Iteración de Valores e Iteración de Políticas, que resultaron variantes de el método usual de solución de ecuaciones por iteración y el método de Newton. Ambos métodos ya era conocido se podían formalizar a partir de los métodos análogos en dimensión finita a través del Teorema del Punto Fijo de Banach. Efectivamente, este enfoque, el de Operadores Contractivos ha tenido mucho impacto en esta área y es una herramienta básica en la presente tesis.

Dada la novedad y gran variedad de problemas en economía e ingeniería en los cuales se pueden aplicar las técnicas de Programación Dinámica para sistemas modelados a tiempo discreto, esta disciplina creció de manera acelerada. En particular, surgieron diversos enfoques que pertenecían a diversas comunidades interesadas en el tema con sus propias terminologías: matemáticos, ingenieros, economistas y más recientemente computólogos, lo que atestigua por un lado la vitalidad del tema y por el otro representa cierta exhuberancia que en un futuro requerirá un trabajo crítico de reordenamiento: en esta tesis analizamos un aspecto de dicha labor, la cuestión de la estabilidad del modelo que se utiliza. El estudio de esta problemática se remonta a los trabajos de Lyapunov en sistemas dinámicos, al analizar la sensibilidad de la solución de una ecuación diferencial a las condiciones iniciales. La estabilidad de un modelo matemático determina en cierto sentido su validez como modelo de una situación real. Con la formalización matemática de la Teoría de la Probabilidad y el desarrollo subsecuente de la Teoría de los Procesos Estocásticos y su utilización como modelos en diversos campos aplicados, la cuestión de la estabilidad se presentó de manera natura. Matemáticos de la talla de Prójorov y Kalashnikov analizaron este tema, planteando un esquema por medio de métricas probabilísticas: medidas numéricas del parecido entre modelos estocásticos distintos.

La presente tesis analiza la aplicación de estas ideas (el uso de métricas probabilísticas) para analizar la estabilidad de procesos controlados. De esta manera se prosigue el programa de investigación que planteó el Prof. E. Gordienko desde los años ochentas, utilizando en diversas ocasiones la herramienta de operadores contractivos.

Para finalizar este capítulo haremos un breve recorrido por los temas del Teorema del Punto Fijo de Banach, las métricas probabilísticas, el concepto de estabilidad que nos interesa en procesos controlados. La sección final describe la estructura de la tesis.

1.2. Operadores contractivos.

Uno de los más resultados más poderosos y a la vez sencillos en matemáticas es el Teorema del Punto Fijo de Banach. Durante años diversos matemáticos como Newton, Peano, entre otros habían utilizado métodos de aproximaciones sucesivas. Con el desarrollo del Análisis Funcional y la noción de Espacio Métrico Completo fue posible que Banach formalizará de manera muy sencillo todos esos resultados dispersos en la literatura.

Recordemos que el Teorema dice que si X es un espacio métrico completo con métrica d y T es un operador de X a X, entonces, bajo la hipótesis de que existe α positiva menor que uno tal que $d(T(x),T(y)) \leq \alpha d(x,y)$ para cualquier $x, y \in X$ tenemos:

- Existe un único punto $x^* \in X$ tal que $T(x^*) = x^*$.

- Dado cualquier punto $x_0 \in X$, la sucesión de iteraciones sucesivas $x_{n+1} = T(x_n)$ converge a x^*.

Este esquema es sumamente flexible, como lo veremos en el capítulo siguiente. Baste mencionar ahora que por ejemplo, el Método de Newton se puede plantear como un problema de operadores contractivos y a partir de ese planteamiento se pueden enunciar condiciones suficientes para que permita construir una sucesión convergente a la raíz de la ecuación original.

En esta tesis nuestro interés particular radica en la posibilidad de utilzar el método de iteraciones sucesivas para resolver la ecuación funcional de Bellman para encontrar la ecuación de Valor del problema de control. Este método se conoce como Iteración de Valores en la literatura de Procesos Markovianos de Decisión, y tal como hemos comentado, lo analizaremos bajo el contexto del Teorema del Punto Fijo de Banach. Este tipo de unificación matemática en esta ocasión sí resulta de interés a los ingenieros, economistas y computólogos, porque al contrario de muchas de la generalizaciones matemáticas, además de garantizar la existencia de la solución nos ofrece un método para aproximarla.

1.3. Métricas Probabilísticas

Cuando analizamos modelos determinísticos, a veces basta comparar trayectorias del modelo utilizando alguna métrica sencilla. Por ejemplo, podríamos analizar qué tan cercanas están dos trayectorias $x(t)$ y $y(t)$ en función la máxima discrepancia posible:

$$d(x, y) = \sup\{|x(t) - y(t)| \colon t \in [0, \tau]\}$$

donde $[0, \tau]$ es el intervalo de tiempo durante el cual se analiza la trayectoria.

Métricas de este tipo son cómodas para estudiar la estabilidad de trayectorias determinísticas. Sin embargo, en nuestro contexto, en vez de trabajar con trayectorias determinísticas, trabajaremos con variables y procesos aleatorios. En este caso la situación se complica enormemente. Dado que el objetivo de este capítulo es esbozar el esquema de la tesis, nos limitaremos a realizar las siguientes observaciones. El campo de las métricas probabilísticas es muy vasto y confuso. En algunos casos ni siquiera se está hablando de métricas en el sentido de espacios métricos. A veces se analiza una métrica sobre el espacio de variables aleatorias sobre Ω, otras sobre el espacio de medidas de probabilidad sobre Ω. Tenemos métricas como las de Kolmogórov, Levy, Ky Fan. En nuestro caso utilizaremos la métrica ponderada del supremo de la discrepancia esperada, cuando analicemos difusiones controlados, la métrica de variación total y la métrica de variación total ponderada y finalmente utilizaremos la métrica de Kantoróvich. Los detalles técnicos de dichas métricas los referiremos a la literatura correspondiente. Las razones por las cuales se decidió utilizar cada métrica en particular se analizarán en la sección correspondiente al problema de estabilidad.

1.4. Estabilidad en Control

Deseamos controlar un sistema Σ. Cada vez que elegimos una estrategia de control u incurrimos en un costo $c(u)$. Bajo condiciones adecuadas, existe una estrategia u^* tal que

$$c(u^*) = \text{ínf}\{c(u) \colon u \in U\}$$

donde U denota el conjunto de todos los controles admisibles. Dicho control u^* queda determinado implícitamente por medio de la función de valor V del problema.

Nuestro problema radica que en la práctica no se conoce el sistema Σ, sino una aproximación $\tilde{\Sigma}$. Podemos por ejemplo encontrar la función de valor $\tilde{V}$ del problema. Incluso es posible que V y $\tilde{V}$ sean funciones muy cercanas. Podemos hallar también la estrategia óptima de control $\tilde{u}^*$. Sin embargo, la cuestión real en la práctica es que esa estrategia $\tilde{u}^*$ no la vamos a utilizar en $\tilde{\Sigma}$ sino en el sistema original Σ. No sólo es muy improbable que $\tilde{u}^*$ sea igual a u^* sino que es posible que aunque se le parezca, al utilizarla en el sistema original, su costo $c(\tilde{u}^*)$ sea muy alto.

Nuestro reto consiste en

- Elegir una métrica d que nos permita estimar qué tan cercanos se encuentar los sistemas Σ y $\tilde{\Sigma}$.

- Determinar una cota de la diferencia

  $$\Delta = c(\tilde{u}^*) - c(u^*)$$

  en términos de $d(\Sigma, \tilde{\Sigma})$.

- Determinar si $\Delta \to 0$ cuando $d(\Sigma, \tilde{\Sigma}) \to 0$

En este caso diremos que el sistema es estable en el sentido de controles, control-estable o c-estable, para distinguir el concepto de otros tipos de estabilidad.

En particular, el estudio de la c-estabilidad en términos de operadores contractivos es el tema de la presente tesis.

Vale la pena señalar que apesar de su gran importancia, este tema aún plantea una gran riqueza de problemas abiertos que pueden ser atacados, por lo que esperamos que esta tesis sirva también para divulgar esta rama de la investigación y lograr motivar a alumnos de postgrado a considerla como un opción interesante como tema de investigación.

1.5. Agradecimientos

Como es la regla, cualquier labor humana es un fruto de una enorme red de seres humanos. Esta tesis lo es mucho más que el promedio.

Así que con gusto agradezco, en primer lugar, al Dr. E. Gordienko, por su inmensa generosidad humana e intelectual, y por su solidaridad, al tender una mano amiga en el momento que más lo necesitaba.

A Pablo Padilla, J.A. Christen, Raúl Montes de Oca, soberbios modelos en lo académico y grandes amigos en lo personal.

A Carlos Castaño y Miguel Ángel Pizaña, por haber mantenido siempre vivo el interés en las matemáticas.

A mis profesores, Onésimo Hernández Lerma, José Antonio de la Peña Mena, Ana Irene Ramírez Galarza, Guillermo Torres (q.e.d.), Leonardo Salmerón.

1.6. Dedicatoria

Dedico esta tesis a:

Enrique Lemus Rodríguez, mi padre, que en paz descanse.

A Amélie, Daniela, Sofía, mis hijas.

A Minelly, esposa devota.

A Esperanza Rodríguez de Lemus, gran compañera e inspiración constante.

Gerardo y Alexandra, hermanos queridos tan lejanos y tan cercanos.

Josefina, abuelita y ángel de mis hijas.

A Ernesto Lemus Noguez, mi abuelo que en paz descanse, gracias por Julio Verne.

Y a todos aquellos a los que les debo todo, desde Tales, Euclides, Arquímedes, Euler, Riemann, Hilbert, Poincaré, Kolmogórov ..., pasando por Julio Verne, Bach, Michael Nyman, The Beatles,...,Perelman, Martin Gardner, Yaglom, ..., Kurosawa, Eisenstein, Chaplin, los hermanos Marx, ..., llegando ahora con mis hijas a Michael Ende, Cornelia Funke, Carlo Collodi, a todos los que todos debemos tanto.

Capítulo 2

Operadores Contractivos y Estabilidad

2.1. Introducción

Esta sección recapitula en anteproyecto original que se presentó. Sirve para retomar el contexto original y los conceptos que se detallan posteriormente. Los siguientes dos capítulos describen los dos trabajos realizados y publicados para esta tesis.

2.2. Convergencia de Procesos de Markov a tiempo discreto y métricas probabilísticas.

Un proceso de Markov a tiempo discreto es un modelo matemático de un sistema dinámico sujeto a perturbaciones aleatorias periódicas. La teoría de la convergencia de dichos procesos a su distribución estacionaria (incluyendo la estimación de la tasa de convergencia) se encuentra bien desarrollada en el caso de métricas probabilísticas fuertes (como la de variación total, y la de variación total ponderada, ver por ejemplo [9]). En algunos problemas aplicados relacionados con procesos de Markov es importante estimar la tasa de convergencia de distribuciones del proceso respecto a métricas débiles (como la de Prokhorov o la métrica de Kantorovich). Sobre este tema existen resultados aislados (ver [17][12]). Por ejemplo, en contraste con el caso de las métricas fuertes, no existe para las débiles un criterio general de determinación de tasa de convergencia exponencial a las distribuciones estacionarias.

2.3. Métricas, contracciones y estabilidad de procesos de Markov controlados.

Si la evaluación de un proceso de Markov a tiempo discreto depende de algunas acciones que pudiera llevar a cabo un controlador externo estamos hablando de un proceso controlado. El controlador debe contar con una función objetivo que le permita evaluar la calidad de sus acciones. En este contexto, dos de los criterios más utilizados en la teoría y las aplicaciones son el costo total descontado y el costo promedio esperado. (ver [3]). El problema de optimización de un sistema controlado bajo alguno de estos criterios es un relativamente viejo y cuenta con una extensa teoría. Más reciente es el problema de estimaci´ón de la estabilidad de los controles óptimos cuando nos enfrentamos a la incertidumbre en la determinación de los parámtros que determinan el proceso. Los resultados con los que contamos en etes área (estimación del índice de estabilidad) utilizan métricas fuertes. (ver por ejemplo [1], [2], [7], [10], [18]). Por otro lado, la estabilidad (y su estimación cuantitativa) es de una gran importancia práctica para un clase muy amplia de problemas (colas, riesgo, finanzas), cuando se utilizan métodos estadísticos para evaluar los parámetros del proceso real [19]. En particular, en el caso del uso de distribuciones empíricas, salvo para algunos casos particulares ([13] y [16]), no existe una teoría en esta dirección. En los trabajos antes mencionados los métodos de métricas probabilísticas ([5] y [6]) se combinaron exitosamente con la técnica de los operadores contractivos (ver también [8] y [11]). Podemos ser optimistas y conjeturar que un enfoque análogo podría tener éxito para algunas clases generales de procesos de Markov controlados.

2.4. Objetivo

De acuerdo a los comentarios anteriores, planteamos los objetivos del Proyecto de la siguiente manera.

Utilizando los métodos de métricas probabilísticas junto con los métodos de operadores contractivos en espacios métricos y las nuevas estimaciones de la tasa de convergencia de medidas empíricas se busca determinar por medio de la combinación de estas técnicas, las condiciones generales de convergencia exponencial y de momentos de distribuciones a las distribuciones invariantes (con respecto de las métricas de Kantorovich y Dudley, [6], [13]) de procesos de Markov a tiempo discreto. En una siguiente etapa se espera poder aplicar estos resultados junto con la técnica del artículo [2] para establecer las desigualdades para el índice de estabilidad para procesos controlados con costo total descontado y con costo promedio. Finalmente, se planea considerar aplicaciones de los resultados obtenidos en el caso de aproximación por distribuciones empíricas para estudiar la estabilidad de procesos de consumo e inversión y de procesos controlados de colas.

2.5. Problemas Propuestos

2.5.1. Problema 1

Estimaciones cuantitativas de estabilidad de procesos de Markov controlados, a tiempo discreto con respecto a métricas débiles (de Kantorovich y sus extensiones, [13]).

En el planteamiento de este problema se consideran dos procesos de Markov controlados, el proceso real P_r, del cual no conocemos todas las características y el procesos aproximado P_a, del cual conocemos todas sus características. Sean $Q(\pi)$ y $\widetilde{Q}(\pi)$ los valores del criterio de optimización de los procesos real y aproximado respectivamente, es decir, $Q(\pi)$ y $\widetilde{Q}(\pi)$ denotan el costo total descontado o el costo promedio esperado que se obtiene al aplicar la política π tanto a P_r como P_a. Sean π^* y $\widetilde{\pi}^*$ las políticas óptimas para ambos procesos, es decir:

$$Q(\pi^*) = \inf_{\pi} Q(\pi) \tag{2.1}$$

y

$$\widetilde{Q}(\widetilde{\pi}^*) = \inf_{\pi} \widetilde{Q}(\pi). \tag{2.2}$$

Si utilizamos la política "aproximada" $\widetilde{\pi}^*$ (es decir, la óptima para el proceso aproximado) en el proceso real P_r entonces posiblemente no alcanzaremos el costo mínimo, quedando quizás por encima. A dicha diferencia la denominaremos índice de estabilidad:

$$\Delta := Q(\widetilde{\pi}^*) - \widetilde{Q}(\pi^*). \tag{2.3}$$

Nuestro objetivo consiste en poder establecer desigualdades de estabilidad del tipo:

$$\Delta \leq d(P_r, P_a),$$

donde d es una distancia entre las probabilidades de transición de P_r y P_a expresada en términos de métricas probabilísticas de tipo Kantorovich [13]. Aquí el paso clave consiste en establecer algún tipo de "suavidad" de las soluciones de las ecuaciones de estabilidad relacionadas con el costo en cuestión [3].

2.5.2. Problema 2

Estimación de la tasa de convergencia del índice de estabilidad en el caso de aproximación por distribuciones empíricas.

Supongamos la siguiente situación que se presenta frecuentemente en la práctica: el proceso aproximado P_a se obtiene a partir del real P_r a través de un cambio de su distribución μ por una distribución empírica μ_n, originada a partir de n observaciones independiente de correspondiente vectores aleatorios. Se puede esperar poder combinar desigualdades de estabilidad con las estimaciones conocidas respecto a la tasa de convergencia de las medidas empíricas μ_n a μ cuando $n \to \infty$ respecto a ciertas métricas probabilísticas ([4], [15], [14]). Ésto nos permitiría evaluar errores en el problema de optimización en términos del volumen de información estadística disponible. Básicamente, los resultados conocidos para tasa de convergencia de distribuciones empíricas se pueden ser utilizados cuando la función de costo (solución de la ecuación de optimalidad) satisface la condición de Lipschitz. En el caso más general (por ejemplo, cuando la función satisface la condición de Hölder) el problema de evaluar la tasa de convergencia de distribuciones empíricas respecto a métricas débilies generadas por clase de funciones Hölder se encuentra aún fundamentalmente sin resolver. Este problema tiene interés no sólo en el estudio de la estabilidad de procesos controlados.

2.5.3. Problema 3

Aplicaciones respecto a la estabilidad de procesos de optimización de consumo e inversión.

Existen clases de procesos de consumo e inversión para los cuales las políticas óptimas respecto al criterio de costo descontado son conocidas de forma explícita. Sería muy interesante hasta qué grado éstas expresiones explícitas funcionan cuando los parámetros del proceso han sido perturbados. Es claro que disponer de desigualdades de estabilidad sería un paso en la resolución del problema.

2.5.4. Problema 4

Estimación respecto a métricas débiles de la tasa de convergencia de procesos de Markov a tiempo discreto a distribuciones estacionarias.

Sea μ la distribución inicial de un proceso de Markov con probabilidad de transición P. Supongamos que existe una distribución estacionaria invariante π, es decir, que cumple:

$$P\pi = \pi.$$

(ver por ejemplo [17] para resultados recientes sobre la existencia de π). Denotemos con $t = 1, 2, \ldots$, un índice discreto de tiempo. Sea $\mu_t = P^t\mu$, es decir, μ_t es la distribución del proceso al tiempo t. Existen resultados del tipo

$$\rho(\mu_t, \pi) \to 0, t \to \infty, \tag{2.4}$$

$$\rho(\mu_t, \pi) \leq c\alpha^t, \alpha < 1, t = 1, 2, 3, \ldots \tag{2.5}$$

donde ρ es una distancia de variación total o ponderada, bajo condiciones adecuadas [9]. Cuando ρ es una métrica débil el problema de convergencia se complica y existe poco investigación realizada al respecto. Existen ejemplos de procesos (ver [12] y [17]) para los cuales

$$\rho(\mu_t, \pi) \geq c/t, c > 0, t = 1, 2, 3, \ldots \tag{2.6}$$

donde ρ es una métrica débil (la métrica de Dudley). Pero para dicho proceso existe una métrica equivalente ρ_e para la cual si se cumple:

$$\rho(\mu_t, \pi) \leq c\alpha^t, \alpha < 1, t = 1, 2, 3, \ldots \tag{2.7}$$

Un problema abierto interesante consiste en establecer condiciones sobre la probabilidad de transición P bajo las cuales se cumplan 2.6 y 2.7 para métricas débiles como las de Dudley o Kantorovich. Es posible que el uso de métodos de operadores contractivos permita dilucidar dicha cuestión. Desigualdades como la 2.7 serán de gran utilidad en el estudio de la estabilidad de procesos de Markov, controlados o no.

2.6. Comentarios Generales.

La UAM Unidad Iztapalapa ofreció la infraestructura necesaria para desarrollar la parte del proyecto anteriormente descrito que constituye el objeto de esta tesis doctoral y que se describe en los dos antepenúltimos capítulos. Tal como se comenta en el capítulo final, además de los problemas de este proyecto que quedaron pendientes, hay nuevos problemas que surgieron durante el trabajo doctoral.

Capítulo 3

Caso Costo Descontado

El objetivo de este trabajo es determinar desigualdades que estimen la estabilidad (robustez) en problemas de optimízación respecto al costo descontado de Procesos Markovianos controlados en espacios de estados de Borel. En particular, nos interesa que el costo por etapa pueda ser no-acotado, para poder cubrir casos importantes en la práctica como el problema de control lineal con costo cuadrático.

Se ha desarrollado mucho trabajo al respecto, pero en esta tesis trataremos el caso en el cual la perturbación sobre las probabilidades de transición se estudia midiéndola respecto a la métrica de Kantorovich, relacionada íntimamente con la convergencia débil. En particular, esto abre la posibilidad de estimar la tasa en la que se va a cero el índice de estabilidad cuando la aproximación se realiza vía medidas empíricas, caso que tiene interés estadístico. En estudios anteriores, cuando el índice de estabilidad se analizaba vía métricas fuertes como la de variación total, el uso de medidas empíricas no era posible.

El contenido de este capítulo corresponde al artículo **Discounte cost optimality problem: stability with respecto to weak metrics** publicado en 2008 por Math. Meth. Oper. Res., vol. 68: 77-96, escrito por Evgueni Gordienko, Enrique Lemus-Rodríguez (doctorando) y Raúl Montes-de-Oca.

3.1. Planteamiento del Problema

Para beneficio del lector, detallaremos el prolbema a tiempo discreto que nos concierne (para más detalles, se puede consultar [4], [9]). Nos interesa primero analizar el criterio de costo total descontado. En el capítulo siguiente analizaremos el criterio de costo promedio.

Supondremos que el proceso de control está dado por la siguiente ecuación estocástica en diferencias:

$$x_t = F(x_{t-1}, a_t, \xi_t), \qquad t = 1, 2, \dots, \tag{1,1}$$

donde F es una función medible dada;

$$x_{t-1}, x_t \in X, \;\; a_t \in A(x_{t-1}) \subset A, \;\; t \geq 1,$$

y $\xi_1, \xi_2, \dots$ es una sucesión i.i.d. de vectores aleatorios que toman valores en un espacio de Borel S. Denotaremos por ξ a un vector aleatorio (v.a.) genérico para $\xi_1, \xi_2, \dots$. El espacio de estados X y el espacio de acciones$A$ son de Borel también: a cada estado $x \in X$ le corresponde un conjunto $A(x)$ de acciones admisibles que supondremos compacto, además $\mathbb{K} := \{(x, a) : x \in X,\ a \in A(x)\}$ es medible en $X \times A$.

Supondremos que la función que determina el costo por etapa,$c(x, a)$, $(x, a) \in \mathbb{K}$, es medible. Sean $x \in X$ y π respectivamente, un estado inicial del sistema dado y una*política de control* (ver, [4], [9] para las definiciones). Sea $\alpha \in (0, 1)$ un factor fijo de descuento. Por medio de E_x^π denotaremos el operador esperanza que corresponde a la medida de probabilidad inducida sobre el espacio de trayectorias del proceso bajo la política π cuando el estado inicial es x. El costo total descontado $V(x, \pi)$ se define como es usual:

$$V(x, \pi) := E_x^\pi \sum_{t=1}^{\infty} \alpha^{t-1} c(x_{t-1}, a_t). \tag{1,2}$$

Posteriormente será necesario especificar condiciones que garanticen que sea finita *la función de valor*:

$$V_*(x) := \inf_{\pi \in \Pi} V(x, \pi) \;\; x \in X, \tag{1,3}$$

y que *un política óptima* π_* exista, de tal manera que

$$V(x, \pi_*) = V_*(x) \quad \text{para toda} \quad x \in X. \tag{1,4}$$

De aquí en adelante $\sqcap$ denota la clase de todas las políticas de control.

Para poder plantear el problema de estabilidad (o robustez) en el que estamos interesados, supondremos que la distribución D_ξ del v.a. ξ es desconocida para el controlador, quien sin embargo dispone de una aproximación $D_{\widetilde{\xi}}$ to D_ξ. Aquí $\widetilde{\xi}$ es un vector aleatorio genérico en la sucesión de v.a.'s i.i.d. $\widetilde{\xi}_1, \widetilde{\xi}_2, \ldots$ con distribución común $D_{\widetilde{\xi}}$. Este esquema modela una situación que se presenta con cierta frecuencia en las aplicaciones.

En particular, en lugar del proceso x_t en (1.1) el controlador analiza su correspondiente aproximación $\widetilde{x}_t$ dada por la siguiente ecuación:

$$\widetilde{x}_t = F(\widetilde{x}_{t-1}, \widetilde{a}_t, \widetilde{\xi}_t), \quad t = 1, 2, \ldots \tag{1,5}$$

Supondremos que la única diferencia entre los dos problemas de control (descritos en (1.1) en (1.5)) radica en la diferencia de las correspondientes distribuciones de ξ y de $\widetilde{\xi}$, cuando la segunda está disponible. De esa manera, en vez de la política óptima π_* (que satisface (1.4)) y es de hecho inaccesible, se debe utilizar en el sistema la política $\widetilde{\pi}_*$, óptima para (1.5), como una aproximación de π_*. La política $\widetilde{\pi}_*$ (en caso de que exista) debe satisfacer:

$$\widetilde{V}(x, \widetilde{\pi}_*) = \widetilde{V}_*(x) := \inf_{\pi \in \sqcap} \widetilde{V}(x, \pi), \quad x \in X, \tag{1,6}$$

donde

$$\widetilde{V}(x, \pi) := E_x^\pi \sum_{t=1}^{\infty} \alpha^{t-1} c(\widetilde{x}_{t-1}, \widetilde{a}_t). \tag{1,7}$$

Mediremos la *calidad* de tal aproximación por medio del costo en exceso de V_* e (1.3), i.e., por medio de la diferencia

$$\Delta(x) := V(x, \widetilde{\pi}_*) - V(x, \pi_*) \geq 0, \tag{1,8}$$

a la que denominamos *índice de estabilidad* (ver [6], [8]).

Tal como se estudia en [6], [8] definimos en el contexto de esta tesis al problema de estimación de estabilidad como el de encontrar cotas, a las que nos referiremos como desigualdades de estabilidad, de la siguiente forma:

$$\Delta(x) \leq C(x) \psi\big[\mu(D_\xi, D_{\widetilde{\xi}})\big], \quad x \in X, \tag{1,9}$$

donde $\psi(s) \to 0$ cuando $s \to 0$, y μ es una métrica definida sobre el espacio de medidas de probabilidad. Aunque suponemos que D_ξ es desconocida, en ocasiones se pueden encontrar cotas superiores para $\mu(D_\xi, D_{\widetilde{\xi}})$. Por ejemplo, esto ocurre cuando $D_{\widetilde{\xi}} = \widehat{D}(\xi_1, \xi_2, \ldots, \xi_n)$ es la distribución empírica y μ es una métrica "débil.ªpropiada (i.e. que metriza la convergencia débil, quizás bajo alguna condición adicional).

Los resultados en los artículos [6], [7], [8], [11], [12] presentan diversas desigualdades tipo (1.9) para $\psi(s) = s^\gamma$ $(0 < \gamma \leq 1)$ bajo "*métricas fuertes*"; tales como

- *la métrica de variación total:*

$$\mu(D_\xi, D_{\widetilde{\xi}}) := \sup \left\{ \left| E[\varphi(\xi) - \varphi(\widetilde{\xi})] \right| : \ \|\varphi\|_\infty \leq 1 \right\};$$

- *la métrica w-ponderada de variación total:*

$$\mu_w(D_\xi, D_{\widetilde{\xi}}) := \sup \left\{ \left| E[\varphi(\xi) - \varphi(\widetilde{\xi})] \right| : \sup_{s \in S} \frac{|\varphi(s)|}{w(s)} \leq 1 \right\}.$$

El lector puede consultar [2], [15] para revisar resultados relacionados. Sin embargo, estos resultados no permiten analizar el problema de estabilidad cuando utilizamos una médida empírica $\widehat{D}(\xi_1, \ldots, \xi_n)$ para aproximar D_ξ: en general, la variación total entre una distribución con soporte finito y una que no tenga soporte finito está acotada para abajo por una constante positiva y por lo tanto no se puede ir a cero (por ejemplo, la variación total entre la medida uniforme sobre el intervalo [0,1] y cualquier medida empírica siempre es 1. Dos de los resultados importantes presentados en esta tesis se refieren al planteamiento de desigualdades de estabilidad en las cuales se utilice una métrica débil. Bajo métricas débiles, si es posible que una medida empírica se aproxime arbitrariamente a la medida dada.

Presentaremos en este capítulo un sistema de condiciones bajo olas cuales los modelos de control bajo consideración obedecen la desigualdad (1.9) para $\psi(s) = s$ y una métrica "débil"μ muy importante. Por ejemplo, el Teorema 2 de la siguiente sección asegura que bajo condiciones adecuadas

$$\Delta(x) \leq K(x)\ell(D_\xi, D_{\widetilde{\xi}}), \quad x \in X, \tag{1,10}$$

donde

$$\ell(D_\xi, D_{\widetilde{\xi}}) := \sup \left\{ |E[\varphi(\xi) - \varphi(\widetilde{\xi})]| : \varphi \text{ such that } |\varphi(s) - \varphi(s')| \leq r(s, s'), \ s, s' \in S \right\}, \tag{1,11}$$

y r es la métrica sobre el espacio S.

Utilizaremos, tal como se mencionó en el primer capítulo, la métrica de Kantorovich que denotaremos por medio de ℓ . Una sucesión de vectores aleatorios ℓ-convergente automáticamente converge débilmente. El siguiente resultado importante (ver, por ejemplo, [14]) ilustra la importancia de la métrica de Kantorovich: si $S = \mathbb{R}^k$ entonces $\ell(D_{\xi_n}, D_\xi) \to 0$ si y sólo si $\xi_n \Rightarrow \xi$ (converge débilmente) y $E|\xi_n| \to E|\xi|$.

Cuando se utilizan medidas empíricas $\widehat{D}(\xi_1, \ldots, \xi_n)$ para aproximar la distribución desconocida D_ξ on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ se demostrará que

$$E\Delta(x) \leq \widetilde{K}(x) n^{-\delta(k)} \ln n, \quad n = 2, 3, \ldots,\ x \in X, \tag{1,12}$$

sujeto a $Ee^{\gamma|\xi|} < \infty$ para alguna $\gamma > 0$. En (1.12) el exponente $\delta(k)$ se comporta como $1/k$. Vale la pena mencionar que existen ejemplos de procesos markovianos controlados (ver [6]) con costo total descontado finito que son inestables respecto a este criterio. Para tales procesos, desigualdades tipo (1.10) y (1.12) son imposibles.

En la tercera sección del capítulo daremos dos ejemplos de procesos markovianos de control para los cuales es posible dar estimaciones tipo (1.10) y (1.12). Uno de ellos es un modelo de inventarios o modelo de sistema de colas con tasa de servicio controlada. El otro es un problema de optimzación de portafolios a tiempo discreto [10].

3.2. Hipótesis y Resultados

Recordemos: denotamos por X el espacio de estados, por A el espacio de acciones y por S el espacio del ruido (el espacio en el cual las perturbaciones aleatorias ξ_n toman valores), todos son espacios métrics con sus respectivas métricas $(X, \rho), (A, d), (S, r)$.

En lo sucesivo utilizaremos en el planteamiento de las hipótesis el concepto de distancia o métrica de Hausdorff, que se define como el real $h(B, C)$ dados conjuntos compactos B, C contenidos en A determinado por la expresión:

$$h(B, C) := \max\{\sup_{x \in B} d(x, C), \sup_{y \in C} d(y, B)\}.$$

Separaremos a las condiciones técnicas en varios bloques. El primero, que es conocido en el contexto de los procesos markovianos de control (ver [9]) se requiere para asegurar la existencia de minimizadores en las correspondientes ecuaciones de optimalidad.

Bloque 1 de Condiciones.

1. *El conjunto $A(x)$ es compacto para cada $x \in X$ y el mapeo que toma como valores conjuntos $x \mapsto A(x)$ es superiormente semicontinuo con respecto a la métrica de Hausdorff.*

2. *La función de costo $c : \mathbb{K} \mapsto \mathbb{R}$ es inferiormente semicontinua.*

3. *Para cada función continua acotada $u : X \mapsto \mathbb{R}$ las funciones*

$$u'(x,a) := \mathbb{E}u[F(x,a,\xi)]$$

y

$$u''(x,a) := \mathbb{E}u[F(x,a,\widetilde{\xi})]$$

son continuas sobre $\mathbb{K}$.

El segundo conjunto de condiciones garantiza la existencia de soluciones a la ecuación de optimalidad y permite que las soluciones posean algunas propiedades útiles para nuestros propósitos (ver [6], [9] y la demostración del Teorema 1 en la sección 4).

Bloque 2 de Condiciones. *Existe una función continua $W : X \mapsto [1,\infty)$ tal que:*

1. *$|c(x,a)| \leq W(x)$ para cada $(x,a) \in \mathbb{K}$.* (2.1)

2. *Existe una constante $\beta \in (\alpha, 1)$ tal que*

$$\mathbb{E}W[F(x,a,\xi)] \leq \frac{\beta}{\alpha} W(x), \tag{2,2}$$

$$\mathbb{E}W[F(x,a,\widetilde{\xi})] \leq \frac{\beta}{\alpha} W(x), \tag{2,3}$$

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

3. *Las funciones*

$$w'(x,a) := \mathbb{E}W[F(x,a,\xi)]$$

y

$$w''(x,a) := \mathbb{E}W[F(x,a,\widetilde{\xi})]$$

son continuas sobre $\mathbb{K}$.

Observación 1. *Se sabe (ver [9],[13]) que las seguientes condiciones son suficientes para (2.1) - (2.3).*

Existen una función continua $W_1 : X \to [1, \infty)$ *y constantes* γ, $\bar{c}$ *y* b *tales que*

- $\alpha\gamma < 1$;
- $|c(x,a)| \leq \bar{c}W_1(x), (x,a) \in \mathbb{K};$ (2.4)
- $EW_1[F(x,a,\xi)] \leq \gamma W_1(x) + b,$ (2.5)
- $EW_1[F(x,a,\widetilde{\xi})] \leq \gamma W_1(x) + b, \quad (x,a) \in \mathbb{K}.$ (2.6)

Denotaremos por B_w el espacio de Banach de todas las funciones medibles $u : X \mapsto \mathbb{R}$ para las cuales la norma:

$$\|u\|_w := \sup\{|u(x)|/W(x) : \quad x \in X\}$$

es finita.

Del teorema 8.3.6 [9], página 47, se sigue que bajos los Bloques 1 y 2 de condiciones, ambos procesos de control (1.1) y (1.5) poseen sus respectivas políticas óptimas estacionarias:

$$\pi_* = \{f_*, f_*, \dots\}, a_t = f_*(x_{t-1}), t \geq 1;$$

$$\widetilde{\pi}_* = \{\widetilde{f}_*, \widetilde{f}_*, \dots\}, \widetilde{a}_t = \widetilde{f}_*(\widetilde{x}_{t-1}), t \geq 1;$$

con correspondientes funciones de valor inferiormente semicontinuas $V_*(x) = V(x, \pi_*), \widetilde{V}_*(x) = \widetilde{V}(x, \widetilde{\pi}_*) \in B_w$, y que las esperanzas $EV_*[F(x,a,\xi)]$, $EV_*[F(x,a,\widetilde{\xi})]$ existen para toda $(x,a) \in \mathbb{K}$. Denotemos por M el conjunto de todas las distribuciones de ξ y $\widetilde{\xi}$ para las cuales las esperanzas anteriores existen para cualquier $(x,a) \in \mathbb{K}$.

Definamos la siguiente pseudo-métrica μ sobre M:

$$\mu(D_\xi, D_{\widetilde{\xi}}) := \sup\{|\mathbb{E}V_*[F(x,a,\xi)] - EV_*[F(x,a,\widetilde{\xi})]| : (x,a) \in \mathbb{K}\}. \quad (2,7)$$

Esta pseudo-métrica toma valores en $[0, \infty]$. Es importante tomar en cuenta que $\mu(D_\xi, D_{\widetilde{\xi}})$ puede ser cero incluso para distribuciones distintas $D_\xi, D_{\widetilde{\xi}}$.

Teorema 1. *Bajo los Bloques 1 y 2 de condiciones se cumple:*

$$\Delta(x) \leq 2\alpha[(1-\alpha)^{-1} + \alpha(1-\beta)^{-2}W(x)]\mu(D_\xi, D_{\widetilde{\xi}}), \quad x \in X. \qquad (2,8)$$

El siguiente paso es acotar la pseudométrica μ por medio de la métrica probabilística débil que nos interesa.

Bloque 3 de Condiciones. *Existen una constante L_0 y una función medible $L_1 : S \mapsto [0, \infty)$ tales que:*

1. $|c(x,a) - c(y,a)| \leq L_0 \rho(x,y)$ *para toda* $(x,a),(y,a) \in \mathbb{K}$.
2. $\rho[F(x,a,\xi), F(y,a,\xi)] \leq L_1(\xi)\rho(x,y)$ *para toda* $(x,a),(y,a) \in \mathbb{K}$ *y* $L_1 := EL_1(\xi) \leq 1$.
3. *El conjunto A es compacto y $A(x) = A$ para cada $x \in X$.*

Teorema 2. *Bajo los Bloques 1, 2 y 3 de condiciones, tenemos:*

$$\Delta(x) \leq 2\alpha \, \frac{L_0}{1 - \alpha L_1} \left[\frac{1}{1-\alpha} + \frac{\alpha}{(1-\beta)^2} \, W(x) \right] \ell(D_\xi, D_{\widetilde{\xi}}), \quad x \in X, \qquad (2{,}9)$$

donde ℓ es la métrica de Kantorovich definida en (1.11).

Observación 2. *Para $S = \mathbb{R}$,*

$$\ell(D_\xi, D_{\widetilde{\xi}}) = \int_{-\infty}^{\infty} \left| F_\xi(x) - F_{\widetilde{\xi}}(x) \right| dx,$$

donde $F_\xi, F_{\widetilde{\xi}}$ son las correspondientes funciones de distribución (see [14]). Sería interesante determinar las expresiones correspondientes la métrica de Kantorovich sobre probabilidades en $\mathbb{R}^n$.

Consideremos la aproximación de D_ξ por medio de medidas empíricas. Sea $n \geq 1$ fija, $\xi_1, \xi_2, \ldots \xi_n$ v.a.'s i.i.d.con distrubución D_ξ y

$$\widehat{D}_n = \widehat{D}(\xi_1, \ldots, \xi_n) := \frac{1}{n} \sum_{i=1}^{n} \delta_{\xi_i}$$

la medida empírica (es decir, una medida aleatoria sobre $(S, \mathcal{B}(S))$. Denotemos por medio de $\widetilde{\xi}^{(n)}$ un vector aleatorio con distribución $\widehat{D}_n$ y sean $\widetilde{\xi}_t^{(n)}$, $t = 1, 2, \ldots$ copias independientes de $\widetilde{\xi}^{(n)}$.

Supongamos que $\widetilde{\xi}_t \equiv \widetilde{\xi}_t^{(n)}$, $t = 1, 2, \ldots$ en el proceso aproximante(1.5). Entonces el índice de estabilidad en(1.8) es aleatorio. Para evitar problemas de medibilidad la esperanza considerada en (2.11) y (2.12) y abajo se puede tomar como la integral exterior, tal como se define en [16]. En este contexto se plantea de manera natural el problema de control adaptado: ¿qué ocurriría si en cada etapa modificamos la medida empírica utilizando la nueva información disponible y actualizamos la política óptima del nuevo problema aproximante?¿Se puede adaptar este proceso al contexto Bayesiano? Esperamos en un futuro cercano poder analizar estas cuestiones, pero por lo pronto, escapan al marco de esta tesis, salvo su mención como problemas abiertos en el capítulo de conclusiones.

Corolario. *Sea*

$$S = \mathbb{R}^k, \quad Ee^{\gamma|\xi|} < \infty \quad \textit{para algún} \quad \gamma > 0, \tag{2,10}$$

y supongamos que los tres Bloques de Condiciones se cumplen. Entonces, para cada $\epsilon \in (0, \frac{1}{2})$ *existe una función finita* $K(x) = K\big(\epsilon, \alpha, \beta, L_0, L_1, \gamma, k, W(x)\big)$ tal que

$$E\Delta(x) \leq K(x) n^{-\delta(\epsilon,k)} \ln(n), \;\; n = 2, 3, \ldots, \;\; x \in X. \tag{2,11}$$

donde

$$\delta(\epsilon, k) := \begin{cases} 1/2, & if \quad k = 1, \\ 1/2 - \epsilon, & if \quad k = 2, \\ 1/k, & if \quad k \geq 3. \end{cases}$$

Este corolario sigue de (2.9) y de la proposición que le sigueque se deriva de los resultados presentes en [3] y tiene interés propio. Nótese que el exponente $\delta(\epsilon, k) = 1/k$ y el conjunto $K(x)$ en (2.11) no dependen de ϵ para $k \neq 2$. Es plausible que para $k = 2$ in (2.11) también se puede tomar $\epsilon = 0$.

Proposición. *Sea* ξ *un vector aleatorio en* $\mathbb{R}^k$, *para alguna* $\gamma > 0$, $Ee^{\gamma|\xi|} < \infty$, *y* $\xi_1, \xi_2, \ldots \xi_n$ *v.a.s i.i.d. distribuidas con*ξ.

Sea $Lip := \{\varphi : \mathbb{R}^k \to \mathbb{R} : |\varphi(x) - \varphi(x')| \leq |x - x'|, \; x, x' \in \mathbb{R}^k\}$.
Entonces existe una constante$M = M(\gamma, k)$ *tal que*

$$E \sup_{\varphi \in Lip} \left| E\varphi(\xi) - \frac{1}{n} \sum_{i=1}^{n} \varphi(\xi_i) \right| \leq M n^{-\delta(\epsilon,k)} \ln(n), \;\; n = 2, 3, \ldots \tag{2,12}$$

Observación 3. *De los resultados en [3] se sigue que en general para* $k \neq 2$ *el exponente* $\min\left(\frac{1}{2}, \frac{1}{k}\right)$ *en (2.12) no se puede mejorar.*

Observación 4. *(Un problema abierto).*

Al demostrar las desigualdades de estabilidad (2.9), (2.11) de demostró primero que la función de valor V_* involucrada en (2.7) satisface la condición de Lipschitz. Esto se debe al Bloque 3 de Condiciones, particularmente al hecho de que la función de costo c es por hipótesis de Lipschitz. Existe una amplia clase de procesos markovianos controlados en las aplicaciones, para las cuales se asume que

$$c(x) = bx^{\alpha}, \quad \alpha < 1 \quad (X = [0, \infty)),$$

y se prueba que V_* satisface la condición de Hölder para la misma α: sin embargo, en este caso V_* no es de Lipschit. Por ejemplo, esto ocurre en procesos de distribución de recursos entre consumo y producción ([4], Ch. 6, §11) o en problemas de optimización de portafolios ([10]). Para tales procesos algunas versiones de la desigualdad (2.9) se pueden demostrar substituyendo funciones Lipschitz en la definición (1.11) por funciones Hölder. Sin embargo, hasta donde sabemos, constituye un problema abierto estableces la tasa de convergencia a cero de la expresión:

$$E \sup_{\varphi \in H\ddot{o}\ell} \left| E\varphi(\xi) - \frac{1}{n} \sum_{i=1}^{n} \varphi(\xi_i) \right|,$$

where

$$\text{Höl} := \{\varphi : \mathbb{R}^k \to \mathbb{R} : |\varphi(x) - \varphi(x')| \leq |x - x'|^{\alpha}, \quad x, x' \in \mathbb{R}^k\}.$$

Para concluir, presentaremos un Bloque de condiciones suficientes para que $\Delta(x) \to 0$ bajo cualquier aproximación de D_ξ respecto a la topología débil.

Bloque 4 de Condiciones.

1. *La función $c(x, a)$ de costo por etapa es acotada y uniformemente continua en $\mathbb{K}$.*

2. *El mapeo $x \to A(x)$ (que a cada x le asigna el conjunto $A(x)$) es uniformemente continuo respecto a la métrica de Hausdorff.*

3. *La función $F(x, a, s)$ es uniformemente continua sobre $\mathbb{K} \times S$.*

Tal como en la sección 1, sea ξ un vector aleatorio genérico para la sucesión $\xi_1, \xi_2, \ldots$ involucrada en (1.1), y tambi´´en para $n = 1, 2, \ldots$ sea $\widetilde{\xi}^{(n)}$ un vector aleatorio genérico para la sucesión i.i.d. $\widetilde{\xi}_1^{(n)}, \widetilde{\xi}_2^{(n)}, \ldots$ la cual para cada n define un proceso de control aproximante tal como en (1.5),

$$\widetilde{x}_t^{(n)} = F\left(\widetilde{x}_{t-1}^{(n)}, \widetilde{a}_t^{(n)}, \widetilde{\xi}_t^{(n)}\right), \quad t = 1, 2, \ldots. \tag{2,13}$$

Se sabe (ver, por ejemplo [4]) que el Bloque 4, Condiciones (1) y (2) garantizan la existencia de políticas óptimas estacionarias $\pi_* = \{f_*, f_*, \ldots\}$ para el proceso (1.1) y $\widetilde{\pi}_*^{(n)} = \{\widetilde{f}_*^{(n)}, \widetilde{f}_*^{(n)}, \ldots\}$ para su $n-$ésima aproximación (2.13). De manera similar a (1.8), definimos la sucesión de índices de estabilidad :

$$\Delta_n(x) := V(x, \widetilde{\pi}_*^{(n)}) - V(x, \pi_*) \geq 0 \qquad x \in X, \;\; n \geq 1.$$

Teorema 3. *Supongamos que también se cumple el Bloque 4 de Condiciones. Si $\widetilde{\xi}^{(n)} \Rightarrow \xi$ (converge débilmente), entonces, para cada $x \in X$*

$$\Delta_n(x) \to 0 \qquad \text{as} \qquad n \to \infty.$$

3.3. Ejemplos de aplicación de las Desigualdades de Estabilidad

En esta sección consideraremos dos ejemplos de procesos markovianos aplicados de control para los cuales lo Bloques 1, 2 y 3 de condiciones se cumplen y por lo tanto, las desigualdades (2.9), (2.11) también.

3.3.1. Modelo a tiempo discreto de un problema de optimización de portafolio de consumo en inversión

Específicamente, se considerará una versión a tiempo discreto de problema de consumo de vida (ver [10]), donde se desea maximizar la utilidad α-descontada:

$$U(x, \pi) = E_x^\pi \sum_{t=1}^{\infty} \alpha^{t-1} u\big(x_{t-1}, a_t\big). \tag{3,1}$$

Aquí $x_{t-1} \in X = (0, \infty)$ denota la riqueza presente, mientras el control a_t designa la proporción de dicha riqueza x_{t-1} que se decidió consumir, y describe en el tiempo $t-1$ la asignación del resto en un portafolio de activos.

Consideraremos m activos en el mercado, indizdos por $k = 1, 2, \ldots, m$.

En esta versión del modelo, los conjuntos de acciones no dependen de x , es decir, $A(x) = A$ para cada $x \in X$, donde A es el siguiente subconjunto compacto de $\mathbb{R}^{m+1}$:

$$A := \left\{ a = (a_0, a_1, \ldots, a_m) : \sum_{k=0}^{m} a_k = 1;\ a_k \geq 0,\ k = 1, \ldots, m;\ \ \delta_0 \leq a_0 \leq \frac{1}{1+\delta} \right\}$$

y δ_0, $\delta > 0$ son constantes dadas tales que $\delta_0 \leq \frac{1}{1+\delta}$. (Se les puede tomar muy cercanas a cero.)

La dinámica del sistema la determinan las ecuaciones:

$$x_t = (1 - a_0) x_{t-1} \sum_{k=1}^{m} a_k \zeta_t^{(k)}, \quad t = 1, 2, \ldots, \tag{3,2}$$

donde los vectores aleatorios $\xi_t = \left(\zeta_t^{(1)}, \ldots, \zeta_t^{(m)}\right)$, $t = 1, 2, \ldots$ son independientes e idénticamente distribuidos. La variable aleatoria $\zeta_t^{(k)}$ estrictamente positiva describe el movimiento del precio del k-ésimo activo en el período t, $(k = 1, 2, \ldots, m)$.

Observación 5. *En (3.2) algunas de las variables aleatorias $\zeta_t^{(k)}$ pueden ser degeneradas, es decir, pueden corresponder a inversiones sin riesgo. Un problema de control similar es considerado en [4] (utilizando un terminología diferente).*

Para atacar el problema de minimización planteado, es conveniente introducir la función de costo por etapa c definiéndola como $c(x, a) := -u(x, a)$, $(x, a) \in \mathbb{K}$. Entonces (ver (3.1)) la magnitud a minimizar está dada en (1.2).

Condición 3.1. *Existen una función $C = \mathbb{R} \times A \to \mathbb{R}$ continuas respecto a a y una constante L_0 tales que*

(a) $c(x, a) = C[\ln(a_0 x), a]$, *(utilidad logarítmica)*;

(b) $|C(y, a) - C(y', a)| \leq L_0 |y - y'|$ *para toda* $a \in A$; $y, y' \in \mathbb{R}$.

Para cumplir las Condiciones 2 y 3 definimos $y_t = \ln x_t$ y pasamos de (3.2) al siguiente procesos:

$$y_t = y_{t-1} + \eta_t(a), \quad t \geq 1, \tag{3,3}$$

donde

$$\eta_t = \ln(1 - a_0) + \ln\left[\sum_{k=1}^{m} a_k \zeta_t^k\right]. \tag{3,4}$$

El proceso de control (3.3) está definido en el espacio de estados $Y = \mathbb{R}$; involucra el mismo espacio de controles A (con $A(y) = A, y \in Y$). Su función de costo por etapa $\bar{c}$ corresponde a $\bar{c}(y,a) = C(y + \ln a_0, a)$ (ver Condición 3.1, (a)).

Condición 3.2. *La función*

$$\varphi(a) := E \ln\left|\sum_{k=1}^{m} a_k \zeta_1^{(k)}\right|$$

es finita y continua sobre A.

Procedamos a revisar que se satisfacen las Condiciones 1-3 para el proceso (3.3). La Condición 1 sigue inmediatamente de las Condiciones 3.1, 3.2 y el teorema de convergencia acotada. Para verificar (2.1) y (2.2) es mejor verificar (2.4) y (2.5) en la Observación 1. Con este propósito hagamos $\gamma = 1$ y $W_1(y) = 1 + |y|$.

La desigualdad (2.4) se sigue de la Condición 3.1, (b)ya que la función continua $C\big(\ln(a_0), a\big)$ es acotada en el conjunto A.

Condición (2.5) and Condición 2, item 3 se satisfacen por (3.3) y la Condición 3.2.

Finalmente, la Condición 3 se satisface (con $L_1 = 1$) dada la Condición 3.1, (b) y la estructura de la ecuación (3.3).

Aplicando el Teorema 8.3.6 en [9] observamos que existe política óptima estacionaria $\pi_* = \{f_*, f_*, \dots\}$, $f_* : Y \to A$ que minimiza el costo total descontado

$$V(y,\pi) = E_y^{\pi} \sum_{t=1}^{\infty} \alpha^{t-1} \bar{c}\big(y_{t-1}, a_t\big) \tag{3,5}$$

para el proceso (3.3). Dado que el conjunto $(0,\infty)$ se mapea biyectivamente sobre el conjunto $(-\infty,\infty)$ bajo la función $\ln(x)$, la política estacionaria para el proceso controlado (3.2) definida como $f_*(x) := f_*\big(\ln(x)\big)$ is óptima para el último respecto al criterio:

$$V(x,\pi) = E_x^{\pi} \sum_{t=1}^{\infty} \alpha^{t-1} C\big[\ln(a_0 x), a\big].$$

Sean ahora

$$\widetilde{\xi}_t = \big(\widetilde{\zeta}_t^{(1)}, \ldots, \widetilde{\zeta}_t^{(m)}\big), \quad t \geq 1$$

los vectores aleatorios utilizados para definier el proceso aproximante (3.2). En particular, sea

$$\widetilde{x}_t = (1-a_0)\widetilde{x}_{t-1} \sum_{k=1}^{m} a_k \widetilde{\zeta}_t^{(k)}, \quad t = 1, 2, \ldots.$$

Supongamos que las Condiciones 3.1-3.2 se cumplen para este proceso, tomando logaritmos obtenemos el proceso aproximante

$$\widetilde{y}_t = \widetilde{y}_{t-1} + \widetilde{\eta}_t(a) \tag{3,6}$$

(3.3). $\big(\widetilde{\eta}_t(a)$ se define de manera similar a (3.4)$\big)$.

Sea $\widetilde{\pi}_* = \{\widetilde{f}_*, \widetilde{f}_*, \ldots\}$ una política óptima estacionaria para el proceso (3.6). Entonces, tomando en cuenta el comentario anterior sobre la correspondencia entre las políticas óptimas $\widetilde{f}_*(y)$ y $\widetilde{f}_*(\ln(x))$, podemos reformular el índice de estabilidad en (1.8) de la siguiente manera

$$\Delta(x) = V(y, \widetilde{\pi}_*) - V(y, \pi_*),$$

donde $y = \ln x$ y el costo total descontado V fue definido en (3.5) para el proceso (3.3).

De esta manera hemos verificado que se cumplen las hipótesis del Teorema 2. En consecuencia, obtenemos la siguiente desigualdad de estabilidad para el problema de consumo e inversión óptimo que surge de (3.2):

$$\Delta(x) \leq M(x)\ell(D_{\xi_1}, D_{\tilde{\xi}_1}), \tag{3,7}$$

donde ℓ es la métrica de Kantorovich y $M(x)$ es una función finita cuyos valores se pueden calcular explícitamente en términos de $\alpha, \delta, \delta_0, L_0, \bar{c}$ en (2.4) y $\sup_{a\in A} \varphi(a)$.

En en el caso en el que $D_{\tilde{\xi}_1}$ es la medida empírica, se puede suponer (2.10) y aplicar la desigualdad (2.11) para estimar en este ejemplo la tasa de convergencia a cero del índice de estabilidad.

Observación 6. Para calcular la constante $M(x)$ en (3.7) (para una x dada) se puede utilizar (2.9) con $L_1 = 1$ y evaluar β y $W(x)$ en términos de las constantes presentes en (2.4), (2.5) (con $\gamma = 1$). La manera de lograrlo se detalla en la Sección 8 de [9].

3.3.2. Un sistema de inventarios

En esta ocasión considerarremos un proceso markoviano de control definido por las equaciones:

$$x_t = (x_{t+1} + a_t - \xi_t)^+, \qquad t = 0, 1, \ldots, \tag{3,8}$$

donde $x_t \in X := [0, \infty)$, $a_t \in A(x) \equiv A$ y A es un subconjunto compacto de un intervalo $(0, \theta]$, $\theta \in A$, para $\xi_1, \xi_2, \ldots$ una sucesión de variables aleatorias no-negativas i.i.d.

El proceso de control (3.8) modela la evolución del nivel de la mercancía en la teoría de inventarios (véase e.g. [4], [9]). Por otro lado (3.8) describe los tiempos de espera en el sistema de colas $GI|D|1|\infty$ con tasa controlable de servicio a_t. En el contexto del problema de estabilidad, el proceso (3.8) se estudia en [6] (utilizando la métrica de la variación total).

El proceso aproximante

$$\widetilde{x}_t = \left(\widetilde{x}_{t-1} + \widetilde{a}_t - \widetilde{\xi}_t\right)^+, \quad t = 0, 1, \ldots \tag{3,9}$$

difiere de (3.8) sólo por la distribución de $\widetilde{\xi}_t$.

Denotemos por ξ y $\widetilde{\xi}$ variables aleatorias genéricas respectivamente para ξ_t y para $\widetilde{\xi}_t$.

Condición 3.3

(a) $E\xi > \theta$ y $E\widetilde{\xi} > \theta$;

(b) las variables aleatorias ξ y $\widetilde{\xi}$ poseen densidades acotadas y continuas sobre $[0, \infty)$;

(c) para cada $x \in [0, \infty)$ la función de costo por etapa $c(x, a)$ es continua respecto a a, admite una constante L_0 tal que

$$|c(x,a) - c(y,a)| \leq L_0 |x - y| \tag{3,10}$$

para toda $x, y \in [0, \infty), \qquad a \in A.$

En los artículos [5], [6] se establece que bajo la Condición 3.3 tanto la Condición 1 como la Condición 2, item 3 se cumplen. Sea $\beta \in (\alpha, 1)$ arbitraria fija y $\gamma = (\beta/\alpha - 1)^{-1}$. Como se sigue de lo establecido en los artículos antes mencionados se sigue que la Condición 3.3 implica la existencia de una constante positiva q tal que (2.2) y (2.3) se satisfacen con $W(x) = b(e^{qx} + \gamma)$, donde b es otra constante positiva arbitraria. Sea $d := \sup_{a \in A} |c(0, a)|$. Por (3.10) tenemos que $|c(x, a)| \leq L_0 x + d$. Haciendo $b = \max \left\{ \dfrac{d}{1+\gamma}, L_0/q \right\}$ aseguramos que se cumpla la desigualdad (2.1). Finalmente, la Condición 3, item 2 se sigue de la estructura de las ecuaciones (3.8), (3.9), donde podemos tomar $L_1 = 1$.

Se han establecido todas las hipótesis del Teorema 2. De esa manera, definiendo como en (1.8) índice de estabilidad del proceso controlado (3.8) obtenemos la siguiente desigualdad de estabilidad:

$$\Delta(x) \leq \frac{2\alpha L_0}{1-\alpha} \left[\frac{1}{1-\alpha} + \frac{\alpha b}{(1-\beta)^2} \big(e^{qx} + \gamma \big) \right] \ell\big(D_\xi, D_{\tilde{\xi}}\big), \quad x \in [0, \infty).$$

3.4. Demostraciones

3.4.1. Demostración del Teorema 1

Simplificando la notación de la Sección 2, sean $\pi = \{f, f, \dots\}$, $\widetilde{\pi} = \{\widetilde{f}, \widetilde{f}, \dots\}$ las políticas óptimas estacionarias de los procesos (1.1) y (1.5), respectivamente, y $V_*, \widetilde{V}_*$ las correspondientes funciones de valor (véase (1.3), (1.6)).

Entonces (ver [9], Cap. 8) $V_*, \widetilde{V}_*$ y $f, \widetilde{f}$ satisfacen sus correspondientes ecuaciones de optimalidad (moreover, donde V_* $\widetilde{V}_*$ son las únicas soluciones):

$$V_*(x) = \inf_{a \in A(x)} \big\{ c(x,a) + \alpha E V_* \big[F(x,a,\xi) \big] \big\} = c\big(x, f(x)\big) + \alpha E V_* \big[F(x, f(x), \xi) \big], \tag{4,1}$$

$$\widetilde{V}_*(x) = \inf_{a\in A(x)} \Big\{c(x,a) + \alpha E\widetilde{V}_*\big[F(x,a,\widetilde{\xi})\big]\Big\} = c\big(x,\widetilde{f}(x)\big) + \alpha E\widetilde{V}_*\big[F(x,\widetilde{f}(x),\widetilde{\xi})\big], \tag{4,2}$$

Para $(x,a)\in\mathbb{K}$ denotaremos:

$$H(x,a) := c(x,a) + \alpha EV_*\big[F(x,a,\xi)\big],\ \widetilde{H}(x,a) := c(x,a) + \alpha E\widetilde{V}_*\big[F(x,a,\widetilde{\xi})\big], \tag{4,3}$$

y denotemos con $\Gamma_t = \{x, a_1, x_1, a_2, \dots, x_{t-1}, a_t\}$, $(t \geq 1)$la parte inicial de la trayectoria de (1.1)hasta $t-1$ bajo la política de control $\widetilde{\pi} = \{\widetilde{f}, \widetilde{f}, \dots\}$.

Por la propiedad de Markov del proceso

$$\begin{aligned} \zeta_t := \quad & E^{\tilde{\pi}}\big[\alpha V_*(x_t)|\Gamma_t\big] = H(x_{t-1},a_t) - c(x_{t-1},a_t) \\ & - \inf\nolimits_{a\in A(x_{t-1})} H(x_{t-1},a) + \inf\nolimits_{a\in A(x_{t-1})} H(x_{t-1},a). \end{aligned}$$

Dados (4.1) y (4.3) tenemos:

$$\begin{aligned} \zeta_t = \quad & H(x_{t-1},a_t) - \inf\nolimits_{a\in A(x_{t-1})} H(x_{t-1},a) - c(x_{t-1},a_t) + V_*(x_{t-1}) \\ = \quad & \Lambda_t - c(x_{t-1},a_t) + V_*(x_{t-1}), \end{aligned}$$

donde

$$\Lambda_t := H(x_{t-1},a_t) - \inf_{a\in A(x_{t-1})} H(x_{t-1},a). \tag{4,4}$$

De esa manera

$$E_x^{\tilde{\pi}}\alpha V_*(x_t) = E_x^{\tilde{\pi}}\zeta_t = E_x^{\tilde{\pi}}V_*(x_{t-1}) - E_x^{\tilde{\pi}}c(x_{t-1},a_t) + E_x^{\tilde{\pi}}\Lambda_t. \tag{4,5}$$

Sumando (4.5) con pesos α^{t-1} se obtiene:

$$\sum_{t=1}^{n}\alpha^{t-1}E_x^{\tilde{\pi}}c(x_{t-1},a_t) = \sum_{t=1}^{n}\alpha^{t-1}\big[E_x^{\tilde{\pi}}V_*(x_{t-1}) - \alpha E_x^{\tilde{\pi}}V_*(x_t)\big] + \sum_{t=1}^{n}\alpha^{t-1}E_x^{\tilde{\pi}}\Lambda_t,$$

o

$$\sum_{t=1}^{n}\alpha^{t-1}E_x^{\tilde{\pi}}c(x_{t-1},a_t) - V_*(x) = -\alpha^n E_x^{\tilde{\pi}}V_*(x_n) + \sum_{t=1}^{n}\alpha^{t-1}E_x^{\tilde{\pi}}\Lambda_t. \tag{4,6}$$

Dado que $V_*\in B_w$, por la Condición 2 cumple (véase [9], p. 52) $\lim\limits_{n\to\infty}\alpha^n E_x^{\tilde{\pi}}V_*(x_n) = 0$.

Por lo tanto, pasando al límite $n \to \infty$ in (4.6) obtenemos(ver (1.8)):

$$\Delta(x) = \sum_{t=1}^{\infty} \alpha^{t-1} E_x^{\tilde{\pi}} c(x_{t-1}, a_t) - V_*(x) = \sum_{t=1}^{\infty} \alpha^{t-1} E_x^{\tilde{\pi}} \Lambda_t. \tag{4,7}$$

Por la definición de Λ_t en (4.4) y por (4.2) obtenemos:

$$\Lambda_t = H(x_{t-1}, a_t) - \widetilde{H}(x_{t-1}, a_t) + \inf_{a \in A(x_{t-1})} \widetilde{H}(x_{t-1}, a) - \inf_{a \in A(x_{t-1})} H(x_{t-1}, a).$$

Thus,

$$|\Lambda_t| \leq 2 \sup_{a \in A(x_{t-1})} |H(x_{t-1}, a) - \widetilde{H}(x_{t-1}, a)| \leq 2\alpha \sup_{a \in A(x_{t-1})}$$
$$\big| EV_*\big[F(x_{t-1}, a, \xi)\big] - E\widetilde{V}_*\big[F(x_{t-1}, a, \widetilde{\xi})\big]\big|,$$

donde la esperanza en el último término se toma con respecto a las v.a.'s ξ y $\widetilde{\xi}$ manteniendo x_{t-1} fijo. De la última desigualdad tenemos:

$$\begin{aligned} |\Lambda_t| &\leq 2\alpha \sup_{a \in A(x_{t-1})} \big| EV_*\big[F(x_{t-1}, a, \xi)\big] - EV_*\big[F(x_{t-1}, a, \widetilde{\xi})\big]\big| \\ &+ 2\alpha \sup_{a \in A(x_{t-1})} \big| EV_*\big[F(x_{t-1}, a, \widetilde{\xi})\big] - E\widetilde{V}_*\big[F(x_{t-1}, a, \widetilde{\xi})\big]\big|. \end{aligned} \tag{4,8}$$

Dado (2.7), el primer sumando del lado derecho de (4.8) está acotado por $2\alpha\mu$, donde denotamos $\mu := \mu(D_\xi, D_{\tilde{\xi}})$.

El segundo sumando es inferior a

$$\begin{aligned} &2\alpha \sup_{a \in A(x_{t-1})} E \left\{ \frac{W[F(x_{t-1}, a, \widetilde{\xi})]}{W[F(x_{t-1}, a, \widetilde{\xi})]} \big| V_*\big[F(x_{t-1}, a, \widetilde{\xi})\big] - \widetilde{V}_*\big[F(x_{t-1}, a, \widetilde{\xi})\big]\big| \right\} \\ &\leq 2\alpha \|V_* - \widetilde{V}_*\|_w \sup_{a \in A(x_{t-1})} EW\big[F(x_{t-1}, a, \widetilde{\xi})\big]. \end{aligned} \tag{4,9}$$

Ahora necesitamos acotar el promedio del último factor del lado derecho de (4.9). Utilizando la notación $E_{\tilde{\xi}}$ en vez de E y la desigualdad (2.3) obtenemos para cualquier x_{t-1}, a fija, que

$$EW[F(x_{t-1}, a, \widetilde{\xi})] \equiv E_{\tilde{\xi}} W[F(x_{t-1}, a, \widetilde{\xi})] \leq \frac{\beta}{\alpha} W(x_{t-1}).$$

Así, sacando provecho del que la la sucesión $\{x_{t-1}, a_t\}$ aparece como resultado de la aplicación de la política $\tilde{\pi}$ al proceso (1.1), y tomando en cuenta (2.2) obtenemos la siguiente cota.

$$E_x^{\tilde{\pi}} \sup_{a\in A(x_{t-1})} E_{\tilde{\xi}} W[F(x_{t-1}, a, \widetilde{\xi})] \le \frac{\beta}{\alpha} E_x^{\tilde{\pi}} W(x_{t-1}) = \frac{\beta}{\alpha} E_x^{\tilde{\pi}} W[F(x_{t-2}, a_{t-1}, \xi_{t-1})]$$
$$= \frac{\beta}{\alpha} E_x^{\tilde{\pi}} \big\{ EW[F(x_{t-2}, a_{t-1}, \xi_{t-1})] \bigm| \Gamma_{t-1} \big\} \le \left(\frac{\beta}{\alpha}\right)^2 E_x^{\tilde{\pi}} \left\{ W(x_{t-2}) \bigm| \Gamma_{t-1} \right\} =$$
$$\left(\frac{\beta}{\alpha}\right)^2 E_x^{\tilde{\pi}} \left\{ W\big[F(x_{t-3}, a_{t-2}, \xi_{t-2})\big] \bigm| \Gamma_{t-1} \right\} = \left(\frac{\beta}{\alpha}\right)^2 E_x^{\tilde{\pi}} W \left[F\big(x_{t-3}, a_{t-2}, \xi_{t-2}\big)\right].$$

Procediendo por inducción tenemos las desigualdades:

$$E_x^{\tilde{\pi}} \sup_{a\in A(x_{t-1})} E_{\tilde{\xi}} W[F(x_{t-1}, a, \widetilde{\xi})] \le \left(\frac{\beta}{\alpha}\right)^{t-1} W(x), \quad t = 1, 2, \ldots \qquad (4{,}10)$$

En [6], [9] se establece que los siguientes operadores

$$Tu(x) := \operatorname*{ínf}_{a\in A(x)} \big\{ c(x,a) + \alpha E u\big[F(x,a,\xi)\big] \big\},$$
$$\widetilde{T}u(x) := \operatorname*{ínf}_{a\in A(x)} \big\{ c(x,a) + \alpha E u\big[F(x,a,\widetilde{\xi})\big] \big\}$$

son contracciones B_w con modulo β.

Dado que V_* y $\widetilde{V}_*$ son puntos fijos de dichos operadores tenemos

$$\|V_* - \widetilde{V}_*\|_w = \|TV_* - \widetilde{T}V_*\|_w + \|\widetilde{T}V_* - \widetilde{T}\widetilde{V}_*\|_w,$$

lo que implica:

$$\|V_* - \widetilde{V}_*\|_w \le (1-\beta)^{-1} \|TV_* - \widetilde{T}V_*\|_w \le \alpha(1-\beta)^{-1} \sup_{x\in X} W^{-1}(x) \sup_{a\in A(x)}$$
$$\big| EV_*\big[F(x,a,\xi)\big] - EV_*\big[F(x,a,\widetilde{\xi})\big] \big| \le \alpha(1-\beta)^{-1}\mu. \qquad (4{,}11)$$

Combinando las desigualdades (4.8)-(4.11) se llega a:

$$E_x^{\tilde{\pi}} |\Lambda_t| \le 2\alpha \left[1 + \frac{\alpha}{1-\beta} \left(\frac{\beta}{\alpha}\right)^{t-1} W(x) \right] \mu,$$

y finalmente, dado (4.7) se establece la desigualdad (2.8).

∎

3.4.2. Demostración del Teorema 2.

Dada la definición (1.11) de la métrica de Kantorovich ℓ y la definición (2.7) de la métrica μ, la desigualdad (2.9) sigue de(2.8) si establecemos que para cada $(x, a) \in \mathbb{K}$ la función

$$\varphi(\cdot) := V_*\big[F(x, a, \cdot)\big] : S \to \mathbb{R}$$

satisface la condición de Lipschitz con constante $L_0/(1 + \alpha L_1)$ (véase la Condición 3 para la definición de L_0 y L_1).

Sea $V_0 := 0$ y para $n = 1, 2, \ldots$

$$V_n(x) := \inf_{a\in A} \big\{c(x, a) + \alpha E V_{n-1}\big[F(x, a, \xi)\big]\big\}, \quad x \in X.$$

Como se demuestra en [9], p.47 para cada $x \in X \quad V_n(x) \to V_*(x) \;\; as \to \infty$. Por la Condición 3, items 1 and 3 tenemos que $V_1 \in Lip(L_0)$ (i.e. satisface la condición de Lipschitz con la constante L_0).

Aplicando la Condición 3, inciso 2 tenemos:

$$\begin{aligned} |V_2(x) - V_2(y)| &\le \sup_{a\in A} \big\{|c(x, a) - c(y, a)| + \alpha |EV_1\big[F(x, a, \xi)\big] - EV_1\big[F(y, a, \xi)\big]|\big\} \\ &\le L_0 \rho(x, y) + \alpha \sup_{a\in A} E L_0 \rho\big[F(x, a, \xi), F(y, a, \xi)\big] \\ &\le L_0(1 + \alpha L_1)\rho(x, y), \qquad x, y \in X. \end{aligned}$$

Por inducción se establece que $V_n \in Lip\big(L_0/(1 - \alpha L_1)\big)$, $n = 1, 2, \ldots,$ y así

$$V_* \in Lip\big(L_0/(1 - \alpha L_1)\big).$$

∎

Demostración de la Proposición

Fijemos algún $r \ge 1$ y sea

$$B_r := \{x \in \mathbb{R}^k : \; |x| \le r\} \qquad \text{y}$$

$$\xi_r := \xi I_{B_r}(\xi); \xi_{r,i} := \xi_i I_{B_r}(\xi_i), \qquad i = 1, 2, \ldots, n.$$

Denotemos

$$\delta_n := \sup_{\varphi\in Lip} \left| E\varphi(\xi) - \frac{1}{n}\sum_{i=1}^{n} \varphi(\xi_i)\right|,$$

se puede observar que tomando $f(x) = \varphi(x) - \varphi(0)$ se escribe

$$\delta_n := \sup_{f \in L} \left| Ef(\xi_i) - \frac{1}{n} \sum_{i=1}^{n} f(\xi_i) \right|,$$

donde

$$L := \left\{ f \in Lip : f(0) = 0, |f(x)| \leq |x|, \quad x \in \mathbb{R}^k \right\}.$$

Entonces

$$\delta_n \leq \sup_{f \in L} \left| Ef(\xi) - Ef(\xi_r) \right| + \Delta_{n,r} + \frac{1}{n} \sum_{i=1}^{n} \sup_{f \in L} \left| f(\xi_i) - f(\xi_{r,i}) \right|, \qquad (4{,}12)$$

donde

$$\Delta_{n,r} := \sup_{f \in L} \left| Ef(\xi_r) - \frac{1}{n} \sum_{i=1}^{n} f(\xi_{r,i}) \right|.$$

Since $f(0) = 0$ tenemos

$$|f(\xi) - f(\xi_r)| = |f(\xi)| I_{B_r^c}(\xi) \leq |\xi| I_{B_r^c}(\xi)$$

y por las desigualdades de Hölder y de Markov:

$$E\left\{ |\xi| I_{B_r^c}(\xi) \right\} \leq (E\xi^2)^{1/2} \big(P(|\xi| > r) \big)^{1/2} \leq d \exp\left\{ -\frac{\gamma r}{2} \right\},$$

where $d = d(\gamma)$ es una constante finita. De esa manera

$$\left| Ef(\xi) - Ef(\xi_r) \right| \leq d \ \exp\left\{ -\frac{\gamma r}{2} \right\}. \qquad (4{,}13)$$

El mismo valor acota la esperanza del tercer término del lado derecho de (4.12). Por otro lado, dado que las variables aleatorias ξ_r, $\xi_{r,i}$ están acotadas, se sigue del Corolario 3.3, la Proposición 3.4 y la Proposición 6.1 en [3] que

$$E\Delta_{n,r} \leq crn^{-\min(\frac{1}{2}, \frac{1}{k})}, \quad n = 1, 2 \ldots \qquad (4{,}14)$$

for $k \neq 2$ y

$$E\Delta_{n,r} \leq c(\epsilon) r \, n^{-\frac{1}{2} + \epsilon}, \quad n = 1, 2, \ldots,$$

for $k = 2$ con constantes $c, c(\epsilon)$ independenties de $r \geq 1$. Para $s = \delta(\epsilon, k)$ seleccionamos $r = \max\left\{ 1, \frac{2}{\gamma} \ln(n^s) \right\}$. Entonces las desigualdes (4.12)-(4.15) proveen las cotas (2.12) necesarias.

∎

3.4.3. Demostración del Teorema 3

Es fácil verificar que eligiendo la función $W(x)$ como una constante, la Condición 4 implica las Condiciones 1 y 2. Adicionalmente, se sabe que la función de valor en (1.4) es acotada.

De esa manera, aplicando (2.8), utilizando la definición (2.7) y el hecho de que V_* es acotado, en esta definición basta verificar que la familia de funciones

$$\mathcal{F} := \big\{\varphi(\cdot) = V_*\big[F(x,a,\cdot)\big], \;\; x,a \in \mathbb{K}\big\}$$

es equicontinua en cada punto $s \in S$ (ésto último implica que $\mu\big(D_\xi, D_{\tilde{\xi}^{(n)}}\big) \longrightarrow 0$ cuando $\widetilde{\xi}^{(n)} \Rightarrow \xi$, see, e.g.[1].).

A su vez, por la Condición 4, inciso 3, basta probar que la función V_* es uniformemente continua sobre X.

Sean $V_0 \equiv 0$, y $V_n = TV_{n-1}$, $n \geq 1$, donde el operador T se define en el espacio de las funciones medibles acotados$B$ de la siguiente manera

$$Tu(x) := \inf_{a \in A(x)} \big\{c(x,a) + \alpha E u\big[F(x,a,\xi)\big]\big\}.$$

Se sabe que

$$\sup_{x \in X} |V_n(x) - V_*(x)| \to 0 \quad \text{as} \quad n \to \infty.$$

Por lo tanto, para demostrar que V_* es uniformemente continua, basta establecer que cada función V_n es uniformemente continua sobre X. Sea para una n arbitraria pero fija, la función V_{n-1} uniformemente continua. Demostraremos que V_n también es uniformemente.

Por la Condición 4, incisos 1 y 3 la función

$$\varphi(x,a) = c(x,a) + \alpha E V_{n-1}\big[F(x,a,\xi)\big]$$

es uniformemente continua sobre $\mathbb{K}$. Efectivamente, $V_n(x) = \inf_{a \in A(x)} \varphi(x,a)$ es uniformemente continua ya que,para cada $\epsilon > 0$ podemos hacer $\delta > 0$ de tal manera que $\rho(x,y) < \delta$ y entonces sigue que $h\big(A(x), A(y)\big) < \epsilon$, donde h es la métrica de Hausdorff. (Véase la Condición 4, inciso 2). Tomemos $A = A(x,y) := A(x) \cup A(y)$. Entonces, ya que φ es uniformemente continua, entonces existe una función $q(\epsilon)$, $q(\epsilon) \to 0$ as $\epsilon \downarrow 0$, independiente de $x, y \in X$

tal que

$$\left|\sup_{a\in A(x)}\varphi(x,a)-\sup_{a\in A}\varphi(x,a)\right|<q(\epsilon),$$

$$\left|\sup_{a\in A(y)}\varphi(y,a)-\sup_{a\in A}\varphi(y,a)\right|<q(\epsilon).$$

Por otro lado,

$$\left|\sup_{a\in A}\varphi(x,a)-\sup_{a\in A}\varphi(y,a)\right|\leq\sup_{a\in A}\left|\varphi(x,a)-\varphi(y,a)\right|,$$

y el lado derecho de esta desigualdad converge a cero uniformemente sobre x, y cuando $\rho(x,y)\to 0$.

∎

Bibliografía

[1] Billingsley P., Topsϕe F. (1967) Uniformity in weak convergence Z. Wahsch. Verw. Geb. 7: 1-16.

[2] Van Dijk N.M. (1988) Perturbation theory for unbounded Markov reward processes with application to queueing. Adv. Apl. Probab. 20: 91-111.

[3] Dudley, R.M., (1969) The speed of mean Glivenko-Cantelli convergence. Ann. Math. Statist. 40: 40-50.

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

[5] Gordienko E.I., Hernández-Lerma O. (1995) Average cost Markov control processes with weighted norms: value iteration. Appl. Math. 23: 219-237.

[6] Gordienko E.I., Salem F.S. (1998) Robustnes inequalities for Markov control processes with unbounded cost. Systems Control Lett. 33:125-130.

[7] Gordienko E.I., Salem F.S. (2000) Estimates of stability of Markov control processes with unbounded costs. Kybernetika 36:195-210.

[8] Gordienko E.I., Yushkevich A.A. (2003) Stability estimates in the problem of average optimal switching of a Markov chain. Mathematical Methods of Operations Research 57: 345-365.

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

[10] Korn R., Korn E. (2001) Option Pricing and Portfolio Optimization. Modern Methods of Financial Mathematics. American Mathematical Society, Providence.

[11] Montes-de-Oca R., Salem-Silva F. (2005) Estimates for perturbations of average Markov decision process with a minimal state and upper bounded by stochastically ordered Markov chains. Kybernetika, 41: 757-772.

[12] Montes-de-Oca R., Sakhanenko A., Salem-Silva F. (2003) Estimates for perturbations of general discounted Markov control chains. Appl. Math. 30: 287-304.

[13] Van Nunen J.A.E.E., Wessels J. (1978) A note on dynamic programming with unbounded rewards. Manage. Sci. 24: 576-580.

[14] Rachev S.T., Rüschendorf L. (1998) Mass Transportation Problem. Vol. II: Applications. Springer, New York.

[15] Van Dijk N.M., Sladky K. (1999) Error bounds for nonnegative dynamic models. J. Optim. Theory Appl. 101:449-474.

[16] Van der Vaart A.W., Wellner J.A. (1996) Weak Convergence and Empirical Processes, Springer, New York.

Capítulo 4

Caso Costo Promedio

En este capítulo extendemos el enfoque de estudio de las perturbaciones de un Proceso Markoviano de Control a tiempo discreto por medio del método de contracciones al caso en el cual el criterio de optimización es el de costo promedio. Trabajaremos con un espacio general de estados y de nuevo mediremos el tamaño de la perturbación por medio de la métrica de Kantorovich. Asumiremos que existe una política óptima costo promedio para el problema perturbado pero conocido, que es la que se utilizará en el problema real para controlar el sistema. Como antes, nos interesa estimar el deterioro del desempeño.

No se asume que costo por etapa sea acotado. Bajo condiciones tipo Lyapunov se determinan cotas superiores para el costo promedio en exceso cuando se utiliza la política aproximante en vez de la óptima (desconocida e inaccesible por la misma naturaleza del problema). Como una aplicación interesantes de estas desigualdades se analiza la estabilidad de un proceso de control autoregresivo. Sin embargo, en este contexto también es fácil que se presenten fenómenos de inestabilidad, loque se ilustra por medio de algunos ejemplos.

El contenido de este capítulo corresponde al artículo **Average cost Markov control processes: stability with respect to the Kantorovich metric** aceptado y publicado on-line por Math. Meth. Oper. Res. en el 2008, escrito por Evgueni Gordienko, Enrique Lemus-Rodríguez y Raúl Montes-de-Oca.

4.1. El caso promedio

En este capítulo seguiremos en la medida de lo posible el análisis realizado en el capítulo anterior, pasando del caso de costo total descontado al de costo promedio. La técnica fundamental utilizada es la de los operadores contractivos, pero dada la complejidad inherente al caso promedio, para lograr contar con operadores contractivos se requerirá añadir una Condición tipo Lyapunov.

De nuevo compararemos dos Procesos Markovianos de control a tiempo discreto (véanse e.g. [1,7] las definiciones pertinentes), cuya dinámica está definida por las siguientes ecuaciones:

$$x_t = F(x_{t-1}, a_t, \xi_t), \qquad t = 1, 2, \ldots \tag{1,1}$$

$$\widetilde{x}_t = F(\widetilde{x}_{t-1}, \widetilde{a}_t, \widetilde{\xi}_t), \qquad t = 1, 2, \ldots \tag{1,2}$$

Aquí, los $x_t, \widetilde{x}_t \in X$ de nuevo denotan los estados de los procesos al tiempo t; $a_t \in A(x_{t-1}) \subset A$, $\widetilde{a}_t \in A(\widetilde{x}_{t-1}) \subset A$ son las acciones de control, y $\{\xi_t, t \geq 1\}$, $\{\widetilde{\xi}_t, t \geq 1\}$ son dos sucesiones de elementos aleatorios independientes e idénticamente distribuidos (i.i.d.) que toman valores en el espacio de Bore S. El espacio de estados X y el espacio de acciones A son también espacios de Borel, y a cada estado $x \in X$ su correspondiente conjunto de acciones admisibles $A(x)$ es compacto, el conjunto $\mathbb{K} := \{(x, a) : x \in X, a \in A(x)\}$ es medible en $X \times A$ por hipótesis. En lo que sigue, las métricas en los correspondientes espacios X, A y S se denotan respectivamente por ρ, d y r, mientras que $\mathbb{K}$ posee una métrica $\mu := \text{máx}\{\rho, d\}$. El costo por etapa $c(x, a), (x, a) \in \mathbb{K}$ será no-negativo, medible y, en general, no acotado.

Sean $x \in X$ y π be, respectivamente un estado inicial dado y una *política de control* dada (véase e.g. [1,7] para las definiciones). Para los procesos (1.1) y (1.2) definimos los correspondientes *costos promedio* por unidad de tiempo $J, \widetilde{J}$ de la siguiente manera:

$$J(x, \pi) := \limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E_x^\pi c(x_{t-1}, a_t), x \in X, \tag{1,3}$$

$$\widetilde{J}(x, \pi) := \limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E_x^\pi c(\widetilde{x}_{t-1}, \widetilde{a}_t), x \in X. \tag{1,4}$$

En este capítulo nos interesa analizar el problema clásico de minimizar J y $\widetilde{J}$ en el contexto de las desigualdades de estabilidad que hemos venido estudiando a lo largo de la tesis.

Supondremos que el problema primario es la búsqueda de una *política óptima* π_* para el proceso (1.1) tal que

$$J(x, \pi_*) = \inf_{\pi \in \sqcap} J(x, \pi), \quad x \in X, \qquad (1{,}5).$$

Pero en nuestro contexto este problema no se puede resolver directamente por las razones antes expuestas. (En (1.5) $\sqcap$ denota la clase de todas las políticas.)

Tal como ha sido el tema recurrente en esta tesis, en nuestro contexto, la estimación de estabilidad consiste en dada la distribución real pero desconocida D_ξ del elemento aleatorio ξ_1 in (1.1) es aproximarla por una distribución disponible $D_{\widetilde{\xi}}$ de la perturbación aleatoria $\widetilde{\xi}_1$ que define el proceso de control (1.2). En particular, en la sección 2 retornamos al muy importante caso en el que la aproximación es la distribución empírica $D_{\widetilde{\xi}}$ que se obtiene de observaciones independientes del elemento aleatorio con distribución desconocida D_ξ.

El objetivo inmediato de este capítulo consiste en introducir condiciones bajos las cuales existen políticas óptimas π_* y $\widetilde{\pi}_*$ para el proceso (1.2).

La última política satisface la ecuación:

$$\widetilde{J}(x, \widetilde{\pi}_*) = \inf_{\pi \in \sqcap} \widetilde{J}(x, \pi), \quad x \in X.$$

También se establecerá que ambas "funciones de valor"$J(x, \pi_*)$ y $\widetilde{J}(x, \widetilde{\pi}_*)$ no dependen del estado inicial $x \in X$.

Supongamos que la política $\widetilde{\pi}_*$ se puede determinar (al menos desde el punto de vista teórica), y es la que se utiliza para controlar el proceso "original"(1.1). El correspondiente costo promedio en exceso que se paga es determinado por el índice de estabilidad:

$$\Delta := J(\widetilde{\pi}_*) - J(\pi_*) \geq 0. \qquad (1{,}6)$$

Siguiendo el enfoque de Gordienko (see [2,4,5,6,8]) se plantea el problema de estimación de estabilidad como el de encontrar "desigualdades de estabididad"de la forma

$$\Delta \leq \psi\big[\beta(D_\xi, D_{\widetilde{\xi}})\big], \qquad (1{,}7)$$

donde $\psi(t) \to 0$ cuando $t \to 0+$, y β alguna métrica de probabilidad adecuada al contexto del problema. (Véase por ejemplo [9] para la definición de métrica de probabilidad

El caso en el que β la métrica de variación total ha sido considerado en [2,6,8] y en [5] donde ha sido posible establecer desigualdades de estabilidad como (1.7) para el problema de optimización de costo total descontado, (También son relevantes los trabajoes [10, 11]). Sin embargo, bajo la métrica de variación total y otras métricas “fuertes"no es posible estimar el índice de estabilidad Δ cuando la aproximación se realiza por medio de médidas empíricas, por ejemplo, tal como se ilustró en el primer capítulo.

El problema básico que se propuso para esta tesis doctoral fue el de analizar las desigualdades de tipo (1.7) tomando métricas de probabilidadi “débiles"β, i.e. métricas compatibles con la convergencia débil. En este capítulo se muestra cómo se puede llegar a una desigualdad de este tipo utilizando la métrica de Kantorovich.

En el capítulo anterior se describe como es posible establecer una desigualdad tipo (1.7) for para el problema de optimizacion de costo total descontado precisamente cuando β es la métrica de Kantorovich:

$$\ell(D_\xi, D_{\widetilde{\xi}}) := \sup\left\{|E\varphi(\xi) - E\varphi(\widetilde{\xi})| : \varphi : |\varphi(s) - \varphi(s')| \leq r(s,s'); s,s' \in S\right\}. \tag{1,8}$$

Recordemos que r es la métrica sobre el espacio S.

Tal como se menciona en el primer capítulo, se sabe (véase, e.g. [9]) que, por ejemplo, para $S = \mathbb{R}^k$ la ℓ-convergencia es equivalente a la convergencia débil más convergencia de losprimeros momentos absolutos. En este capílo se ilustra cómo utilizando los resultados del capítulo anterior y los interesantes avances contenidos en [12] respecto a las propiedades contractivas del operador funcional correspondiente a la ecuación de optimalidad del caso de costo promedio, se puede deducir la siguiente cota del índice de estabilidad Δ:

$$\Delta \leq B\ell(D_\xi, D_{\widetilde{\xi}}), \tag{1,9}$$

donde B es una constante que se puede calcular explícitamente.

Para establecer la desigualdad (1.9) requeriremos condiciones tipo Lyapunov (como las utilizadas, por ejemplo, en [3, 12]) y condiciones de “suavidad” impuestas sobre las probabilidades de transición de los procesos (1.1) y (1.2).

Como ejemplos de aplicación de la desigualdad (1.9) se considerará la aproximación por medio de medidas empíricas y la estimación de estabilidad de un sistema autoregresivo simple.

Hay que señalar que las hipótesis relativas a las condiciones tipo Lyapunov por un lado, y a la "suavidad" de las probabilidades de transición, por el otro lado, son muy importantes, en particular, en su ausencia se puede presentar el fenómeno de la inestabilidad de los procesos de control. Se analizarán dos ejemplos de procesos inestables. En ellos, las distancias $\ell(D_\xi, D_{\widetilde{\xi}_n})$ se van a cero, cuando $n \to \infty$, pero, apesar de ello, existe una constante positiva γ (independiente de n) tal que

$$\Delta_n := J(\widetilde{\pi}_*^{(n)}) - J(\pi_*) \geq \gamma \qquad \text{para } n = 1, 2, 3 \ldots.$$

4.2. Condiciones y resultado principal

Primero enunciaremos las condiciones que involucran la función estocástica de Lyapunov W. El uso de este tipo de condiciones no es raro, en particular se puede consultar [3,12].

Bloque 1 de Condiciones. Existe una función medible $W : X \to [1, \infty)$, una constante $\lambda \in [0, 1)$, una medida de probabilidad ν sobre $(X, \mathcal{B}_X)$ y una función continua $\phi : \mathbb{K} \to [0, \infty)$ tales que: para cada $(x, a) \in \mathbb{K}$ se cumple:

(a) $c(x, a) \leq W(x)$;

(b) $P\big(F(x, a, \xi) \in B\big) \geq \phi(x, a)\nu(B)$

$P\big(F(x, a, \widetilde{\xi}) \in B\big) \geq \phi(x, a)\nu(B)$, para toda $B \in \mathcal{B}_X$;

(c) $EW\big[F(x, a, \xi)\big] \leq \lambda W(x) + \phi(x, a) \displaystyle\int_X W d\nu$,

$EW\big[F(x, a, \widetilde{\xi})\big] \leq \lambda W(x) + \phi(x, a) \displaystyle\int_X W d\nu$;

(d) $b := \displaystyle\sup_{(x,a)\in\mathbb{R}} \phi(x, a) \int_X W d\nu < \infty$;

(e) $\displaystyle\inf_{f\in\Delta_0} \int_X \phi\big(x, f(x)\big)\nu(dx) > 0$, donde Δ_0 denota la clase de todos los selectores medibles $f : X \to A$ tales que $f(x) \in A(x)$, $x \in X$.

Observación 1. Cada selector $f \in \Delta_0$ define una *política estacionaria de control* $\pi = \{f, f, \dots\}$. En particular, cuando sea conveniente, interpretaremos a Δ_0 como la clase de todas las políticas estacionarias.

El segundo bloque de condiciones es bastante restrictivo: impone "condiciones tipo Lipschitz" sobre la función de costo por etapa y las probabilidades de transición de los procesos (1.1) y (1.2).

No está de más recordar que la distancia de Hausdorff h entre conjuntos compactos B y C de A es:

$$h(B, C) := \max\{\sup_{x \in B} d(x, C), \sup_{y \in C} d(y, B)\}.$$

De ahora en adelante, para las parejas $(x, a) \in \mathbb{K}$ utilizaremos la notación: $\mathbf{k} = (x, a)$, y denotaremos por ξ y $\widetilde{\xi}$ a los elementos genéricos, respectivamente para $\xi_1, \xi_2, \dots$ y $\widetilde{\xi}_1, \widetilde{\xi}_2, \dots$ in (1.1), (1.2).

Bloque 2 de Condiciones. Existen constantes finitas L_0, L, L_1 y L_* tales que:

(a) $h\big[A(x), A(x')\big] \le L_0 \rho(x, x'); x, x' \in X$;

(b) $|c(\mathbf{k}) - c(\mathbf{k}')| \le L_1 \mu(\mathbf{k}, \mathbf{k}'), \mathbf{k}, \mathbf{k}' \in \mathbb{K}$;

(c) $d_W\big(F(\mathbf{k}, \xi), F(\mathbf{k}', \xi)\big) \le L \mu(\mathbf{k}, \mathbf{k}'),\ \mathbf{k}, \mathbf{k}' \in \mathbb{K}$,
donde se define

$$d_W(\zeta, \widetilde{\zeta}) := \sup_{u : |u| \le W} |Eu(\zeta) - Eu(\widetilde{\zeta})| \qquad (2{,}1)$$

para los elementos aleatorios $\zeta, \widetilde{\zeta}$ en $(X, \mathcal{B}_X)$.

(d) $\sup_{(x,a) \in \mathbb{K}} \rho\big(F(x, a, s), F(x, a, s')\big) \le L_* r(s, s'), s, s' \in S$;

(e) para cada $x \in X$ el mapeo $a \to Eu\Big[F(x, a, \widetilde{\xi})\Big]$ es continuo sobre $A(x)$, cuando $u = W$ y cuando u es una función arbitraria medible y acotada.

Remark 2. In Assumption 3 in [4] the above condition (d) was omitted due to our oversight. Correspondingly, the right-hand side of inequality (2.9) in [4] must be corrected multiplying it by L_*.***Corregir esto en el capítulo anterior***

Se sabe (véanse por ejemplo [3,7]) que los Bloques de Condiciones 1 y 2 proveen condiciones suficientes para la existencia de políticas estacionarias óptimas $\pi_* = \{\varphi, \varphi, \dots\}$, $\widetilde{\pi}_* = \{\widetilde{\varphi}, \widetilde{\varphi}, \dots\}$, respectivamente, para los procesos de control (1.1) y (1.2). Además, los costos promedio mínimos ("las funciones de valor") $J(x, \pi_*)$, $\widetilde{J}(x, \widetilde{\pi}_*)$ no dependen del estado inicial $x \in X$, y

$$J_* := J(\pi_*) = \inf_{\pi \in \Pi} J(\pi) < \infty, \tag{2,2}$$

$$\widetilde{J}_* := \widetilde{J}(\widetilde{\pi}_*) = \inf_{\pi \in \Pi} \widetilde{J}(\pi) < \infty. \tag{2,3}$$

Consecuentemetne, el índice de estabilidad Δ está bien definido por (1.6). Ya está todo listo para formular el resultado principal.

Theorem. Bajos los Bloques de Condiciones 1 y 2

$$\Delta \leq B\ell(D_\xi, D_{\widetilde{\xi}}), \tag{2,4}$$

donde ℓ es la métrica de Kantorovich definida en (1.8), y

$$B = 6(L_0 + 1)L_* \left[1 + \frac{b}{(1-\lambda)^2}\right] \left[L_1 + \frac{L}{1-\lambda} \max\left\{1,\ \frac{b}{1-\lambda}\right\}\right]. \tag{2,5}$$

Observación 3. Si $S = \mathbb{R}$ entonces (ver [9]) $\ell(D_\xi, D_{\widetilde{\xi}}) = \int_{-\infty}^{\infty} |F_\xi(x) - F_{\widetilde{\xi}}(x)| dx$, donde F_ζ es la función de distribución de la variable aleatoria ζ. No hemos logrado localizar ninguna referencia sobre la fórmula correspondiente para vectores aleatorios: quizás éste constituye un problema abierto aún, y de ser así, constituiría un problema muy interesante.

Consideremos el caso en el que la distribución desconocida D_ξ es aproximada por la distribución empírica:

$$D_{\widetilde{\xi}} \equiv \widehat{D}(\xi_1, \xi_2, \dots, \xi_n) := \frac{1}{n} \sum_{i=1}^{n} \delta_{\xi_i},$$

obtenido de realizaciones independientes $\xi_1, \xi_2, \dots, \xi_n$ del elemento aleatorio ξ (donde δ_{ξ_i} es la medida de Dirac). El corolorio abajo sigue directamente de

(2.4) y el resultado en [4] se refiere a la tasa de convergencia de las medidas empíricas en la métrica de Kantoróvich.

Sea $n \geq 2$ un elemento aleatorio arbitrario fijo $\widetilde{\xi}$ in (1.2) con distribución $\widehat{D}(\xi_1, \xi_2, \ldots, \xi_n)$, y la política $\widetilde{\pi}_*$ en (1.6) se obtiene respecto a esta distribución.

Corollary. Supongamos que $S = \mathbb{R}^k$, y para alguna $\gamma > 0$, $Ee^{\gamma|\xi|} < \infty$. Entonces, bajo el Bloque de Condiciones 1 y 2 a cada $\epsilon \in (0, 1/2)$ le corresponde una constante $K = K(\epsilon)$ tal que

$$E\Delta \leq Kn^{\kappa(\epsilon,k)} \ln(n), \qquad n = 2, 3, \ldots, \tag{2,6}$$

donde

$$\kappa(\epsilon, k) := \begin{cases} 1/2 & \text{if} \quad k = 1, \\ 1/2 - \epsilon & \text{if} \quad k = 2, \\ 1/k & \text{if} \quad k \geq 3. \end{cases}$$

4.3. Contraejemplos

En esta sección ilustramos la importancia de las condiciones aparentemente restrictivas que se han impuesto por medio de dos contra ejemplos. En la siguiente sección aplicamos la desigualdad (2.4) al estudio de la estabilidad de un proceso controlado autoregresivo.

Ejemplo 1. *(Verificando la importancia del Bloque 1 de Condiciones en el Teorema.)* Sea $X = \{0, 1, 2, \ldots\}$, $A(x) \equiv A = \{0, 1\}$, $x \in X$, $S = \{0, 1, 2, \ldots\}$,

$$x_t = a_t x_{t-1} + \xi_t, \tag{3,1}$$

$$\widetilde{x}_t = \widetilde{a}_t \widetilde{x}_{t-1} + \widetilde{\xi}_t, \quad t = 1, 2, \ldots, \tag{3,2}$$

donde $\widetilde{\xi} \equiv 0$, y para cualquier $\epsilon \in (0, 1)$ la variable aleatoria $\xi \equiv \xi_\epsilon$ posee distribución de Poisson distribution con parámetro ϵ. Sea también $c(x, a) \equiv c(x)$, $x \in X$ la siguiente función: $c(0) = 1$, $c(1) = 0$ y $c(x) = 3$ para $x \geq 2$.

Eligiendo $x_0 = \widetilde{x}_0 = 1$ vemos quela política estacionaria $\widetilde{\pi}_* = \{1, 1, \ldots\}$ es óptima en promedio para el proceso (3.2) con $\widetilde{J}_*(1) = \widetilde{J}(1, \widetilde{\pi}_*) = 0$. Aplicando esta política al proceso (3.1) encontramos que $x_t = 1 + \xi_1 + \cdots + \xi_t \xrightarrow{\text{a.s.}} \infty$, y por lo tanto $J(1, \widetilde{\pi}_*) = 3$. Por otro lado, para la política estacionaria $\pi = \{0, 0, \ldots\}$ por medio de un cálculo directo podemos verificar que para una ϵ suficientemente pequeña, $J(1, \pi) < 2$.

De esa manera, para toda tal ϵ,

$$\Delta(1) = J(1, \widetilde{\pi}_*) - J(1, \pi_*) \geq J(1, \widetilde{\pi}_*) - J(1, \pi) \geq 1.$$

Es sencillo verificar que $\ell(D_\xi, D_{\xi_\epsilon}) \to 0$ cuando $\epsilon \to 0$. Notemos que en este ejemplo simple el Bloque de Condiciones 2 se cumple, pero que al menos el inciso (c) del Bloque de Condiciones 1 no se llega a cumplir.

Ejemplo 2. (*Ilustrando que el cumplimiento de los Bloques de Condiciones 1 y 2 no son suficientes para la estabilidad con respecto a la métrica de variación total*).

Sea $X = [0, \infty)$, $A(x) \equiv A = \{0, 1\}$, $x \in X$, $S = [0, \infty)$ y

$$x_t = a_t \xi_t; \quad \widetilde{x}_t = \widetilde{a}_t \widetilde{\xi}_t, \quad t = 1, 2, \ldots.$$

For $\epsilon \in (0, 1/2)$ hagamos

$$D_{\widetilde{\xi}} = \frac{1}{2}\,\delta_0 + \frac{1}{2}\,U(0, 1),$$

$$D_\xi = \frac{1}{2}\,\delta_0 + \left(\frac{1}{2} - \epsilon\right) U(0, 1) + \epsilon \delta_{b(\epsilon)},$$

donde δ_x es la distribución de Dirac concentrada en el punto x, $U(0, 1)$ es la distribución uniforme sobre el intervalo $(0, 1)$, y $b(\epsilon) > 1$ es un número que especificaremos en su momento. SeaLet $c(x, a) \equiv c(x)$ una función no-negativa de costo que satisface la condición de Lipschitz y tal que:

- $c(0) = 1, \quad c(x) \leq 1, \quad 0 < x < 1, \quad c(1) = 0;$
- c es decreciente sobre $[0, 1]$ y $\displaystyle\int_0^1 c(x)dx = 8/9;$
- $c(x) = x - 1, \quad x > 1;$

  Es inmediato verificar el Bloque de Condiciones 2 en este ejemplo. Para que se cumpla el Bloque de Condiciones 1 basta con elegir:

- $w(x) = \begin{cases} 1, & 0 \leq x \leq 1, \\ x, & x > 1; \end{cases}$
- $\phi(x, a) \equiv 0{,}4;$

- $\nu = \delta_0$; $\lambda = 0{,}9$; $b(\epsilon) = 0{,}3\epsilon^{-1}$.

Dado que c es decreciente sobre $[0, 1]$, la política estacionaria $\widetilde{\pi}_* = \{1, 1, \dots\}$ es óptima respecto al proceso $\{\widetilde{x}_t\}$. Por otro lado,

$$J(\widetilde{\pi}_*) = \frac{1}{2} + \left(\frac{1}{2} - \epsilon\right)\frac{8}{9} + \epsilon b(\epsilon) \geq 1{,}2$$

para cualquier $\epsilon \in (0, 0{,}05)$, y para la política estacionaria $\pi = \{0, 0, \dots\}$ tenemos: $J_* \leq J(\pi) = 1$. De esa manera, para las constantes antes mencionadas se cumple ϵ, $\Delta \geq J(\widetilde{\pi}_*) - 1 \geq 0{,}2$. Al mismo tiempo, la distancia de Variación Total $V(D_\xi, D_{\widetilde{\xi}}) \to 0$ cuando $\epsilon \to 0$.

Pero hay que notar que este ejemplo no contradice (2.4) porque $\ell(D_\xi, D_{\widetilde{\xi}})$ no se va a cero cuando $\epsilon \to 0$ (dado que $E\widetilde{\xi}$ no converge a $E\xi$).

Se puede construir un contraejemplo que demuestra que el Teorema el inciso (d) del Bloque 2, también es esencial.

4.4. Estabilidad en Control Autoregresivo

Analicemos ahora un caso de mucho interés el cual es c-estable.

Ejemplo (*Aplicación de la desigualdad de estabilidad a un proceso controlado autoregresivo*).

Sea $X = S = \mathbb{R}$, $A = [-\gamma, \gamma]$ with a given $0 < \gamma < 1$, $A(x) \equiv A$, $x \in X$, y

$$\begin{aligned} x_t &= a_t x_{t-1} + \xi_t, \\ \widetilde{x}_t &= \widetilde{a}_t \widetilde{x}_{t-1} + \widetilde{\xi}_t, \qquad t = 1, 2, \dots. \end{aligned}$$

Supongamos que las variables aleatorias ξ y $\widetilde{\xi}$ (genérica para ξ_t y $\widetilde{\xi}_t$) tienen densidades continuas y estrictamente positivas sobre $\mathbb{R}$. Denotaremos dichas por f_ξ y $f_{\widetilde{\xi}}$, respectivamente. Adicionalmente, supondremos que existen funciones medibles $\psi : \mathbb{R} \to \mathbb{R}$, $\varphi : \mathbb{R} \times [-\gamma, \gamma] \times \mathbb{R} \to \mathbb{R}$ tales que:

- $f_\xi(t - ax) = \psi(t)\varphi(t, a, x)$; $t, x \in \mathbb{R}$, $a \in [-\gamma, \gamma]$; (3.3)

- $\displaystyle\int_{-\infty}^{\infty} |\psi(t)| dt = k_0 < \infty, \quad \int_{-\infty}^{\infty} |t||\psi(t)| dt = k_1 < \infty;$ (3.4)

- ϕ es diferenciable en a,x, y

$$\sup_{\substack{t,x\in\mathbb{R},\\ a\in[-\gamma,\gamma]}} \left|\frac{\partial\varphi}{\partial a}\,(t,a,x)\right| \leq k_2 < \infty; \tag{3,5}$$

$$\sup_{\substack{t,x\in\mathbb{R},\\ a\in[-\gamma,\gamma]}} \left|\frac{\partial\varphi}{\partial x}\,(t,a,x)\right| \leq k_3 < \infty. \tag{3,6}$$

Además, supondremos que la densidad $f_{\tilde{\xi}}$ satisface las condiciones (3.3)-(3.6).

Observación 4. Las condicioness (3.3)-(3.6) se satisfacen, por ejemplo, cuando f_ξ es la densidad normal con parámetros μ,σ.

En este caso se puede elegir:

$$\psi(t) = \mathbf{e}^{-\frac{t^2}{4\sigma^2}}.$$

Sea ahora $c : \mathbb{R}\times[-\gamma,\gamma] \to [0,\infty)$ una función de costo por etapa que satisface la condición de Lipschitz del Bloque 2, (b) y tal que para alguna constante $b_0 \geq 1$

$$c(x,a) \leq |x| + b_0, \qquad (x,a)\in\mathbb{K}.$$

Es claro que los incisos (a) y (d) del Bloque 2 se satisfacen con las constantes $L_0 = 0, L_* = 1$.

Eligamos $W(x) := |x| + \bar{b}$, $x\in\mathbb{R}$, donde $\bar{b} \geq b_0$ será especificada posteriormente. Para revisar el cumplimiento del inciso (c) del Bloque 2 hacemos la observación de que es directo verificarlo para $\mathbf{k} = (x,a)$, $\mathbf{k}' = (x',a')$,

$$\begin{aligned} d_W(ax+\xi,\ a'x'+\xi) &= \int_{-\infty}^{\infty} W(t)\,|f_{ax+\xi}(t) - f_{a'x'+\xi}(t)|\,dt \\ &= \int_{-\infty}^{\infty} (|t|+\bar{b})|f_\xi(t-ax) - f_\xi(t-a'x')|dt. \end{aligned}$$

Utilizando (3.3)-(3.6) obtenemos:

$$|t||f_\xi(t-ax) - f_\xi(t-a'x')| \leq |t|\psi(t)(k_2|a-a'| + k_3|x-x'|),$$

y by (3.4)

$$\int_{-\infty}^{\infty} |t|\big|f_\xi(t-ax) - f_\xi(t-a'x')\big|dt \leq k_1(k_2+k_3)\mu\big((x,a),(x',a')\big).$$

De manera similar,

$$\int_{-\infty}^{\infty} \bigl|f_\xi(t-ax) - f_\xi(t-a'x')\bigr|dt \le k_0(k_2+k_3)\mu\bigl((x,a),(x',a')\bigr).$$

Por lo tanto, la desigualdad en el Bloque 2, (c) se cumple con

$$L = \bigl(\bar{b}k_0 + k_1)(k_2+k_3).$$

Sean $\kappa, d > 0$ números que se determinarán más tarde, y $\phi : \mathbb{R} \to \mathbb{R}$ una función continua tal que:

$$\phi(x,a) \equiv \phi(x) := \begin{cases} \kappa, \ x \in [-d,d], \\ 0, \ x \in \mathbb{R}\backslash I_d, \\ \text{lineal en } (-(d-1),-d) \\ \text{y en } (d,(d+1)), \end{cases} \tag{3,7}$$

donde $I_d := [-(d+1),(d+1)]$.

Sean

$$H := \text{máx}\{E|\xi|, E|\widetilde{\xi}|\},$$

y la probabilidad ν en el Bloque1 la distribución uniforme sobre $[-d,d]$.

Para garantizar el cumplimiento del Bloque 1, (c), definimos:

$$\lambda = \lambda(\bar{b}) := \text{máx}\left\{\frac{1+\gamma}{2}, 1-\frac{1}{\bar{b}}\right\} \in (\gamma, 1). \tag{3,8}$$

Entonces

$$\begin{aligned} & EW[F(x,a,\xi)] = E\bigl(|ax+\xi| + \bar{b}\bigr) \le \\ & \gamma|x| + H + \bar{b} = \\ & \lambda W(x) - (\lambda-\gamma)|x| + H + \bar{b}(1-\lambda) \le \\ & \lambda W(x) - (\lambda-\gamma)|x| + H + 1. \end{aligned} \tag{3,9}$$

En el lado derecho de la primer desigualdad del Bloque 1, (c) obtenemos

$$\phi(x,a)\int_X W(x)\nu(x) = \kappa\int_{-d}^{d}|x|\frac{dx}{2d} + \kappa b = \kappa\left[\frac{d}{2}+\bar{b}\right], \tag{3,10}$$

para cada $x \in [-d,d]$.

Elegimos $\bar{b} = \bar{b}(\kappa)$ de la siguiente manera:

$$\bar{b} := \text{máx}\left\{\frac{H+1}{\kappa}, b_0\right\}. \tag{3,11}$$

De esa manera, se cumple $x \in [-d, d]$,

$$EW\big[F(x,a,\xi)\big] \leq \lambda W(x) + \phi(x,a)\int_X W d\nu, \tag{3,12}$$

debido a (3.9)-(3.11).

Para garantizar el cumplimiento de (3.12) para x con $|x| \geq d$, definimos

$$d := \frac{2(H+1)}{1-\gamma}.$$

Entonces por (3.8) $\lambda - \gamma \geq \frac{1-\gamma}{2}$ y

$$(\lambda - \gamma)|x| \geq H + 1$$

para cada $x \in \mathbb{R}\backslash[-d, d]$. Por lo tanto (3.9) implica la desigualdad

$$EW[F(x,a,\xi)] \leq \lambda W(x).$$

Dado que $\phi(x) = 0$ sobre $\mathbb{R}\backslash I_d$, para verificar el Bloque 1, (b) basta considerar $x \in I_d$.

Para una $B \in \mathcal{B}_X$ arbitraria tenemos:

$$P(F(x,a,\xi) \in B) = P(\xi \in B - ax)$$

$$\begin{aligned} \geq \inf_{y\in I_d} P(\xi \in B - y) \quad &= \quad \inf_{y\in I_d}\int_{B-y} f_\xi(t)dt \\ \geq \inf_{y\in I_d}\int_{B\cap I_d} f_\xi(t+y)dt \quad &\geq \quad \lambda(B\cap I_d)h \geq 2dh\nu(B), \end{aligned}$$

donde elegimos una constante positiva conocida$h$ de tal manera que

$$h \leq \text{mín}\left\{\inf_{z\in 2I_d} f_\xi(z), \inf_{z\in 2I_d} f_{\tilde{\xi}}(z)\right\}, \tag{3,13}$$

(λ es la medida de Lebesgue).

Para cumplir el Bloque 1, (b) basta fijar en (3.7)

$$\kappa := 2dh = \frac{4(H+1)}{1-\gamma}\, h. \tag{3,14}$$

Observación 5. De hecho podemos relajar la hipótesis respecto a la positividad de f_ξ y $f_{\widetilde{\xi}}$ asumiendo sólo que son estríctamente positivas en el intervalo

$$\left[-\frac{4(H+1)}{1-\gamma}, \frac{4(H+1)}{1-\gamma}\right].$$

Las restantes desigualdades en el Bloque 1, (b), (c) se verifican de manera similar.

La condición (e) en el Bloque 1 se satisface trivialmente dado que la function ϕ no depende de a. Además, por (3.10), (3.11) y (3.14) tenemos que en el Bloque 1, (d),

$$b = \frac{4h(H+1)}{1-\gamma}\left[\frac{H+1}{1-\gamma} + \text{máx}\left\{b_0, \frac{1-\gamma}{4h}\right\}\right]. \tag{3,15}$$

Todo esto constituye el cumplimiento de las hipótesis del Teorema de la sección 2. Consecuentemente, (2.4) nos brinda la siguiente desigualdad de estabilidad para el ejemplo bajo consideración:

$$\Delta \leq \overline{B}\ell(D_\xi, D_{\widetilde{\xi}}), \tag{3,16}$$

donde, según (2.5) y el cálculo anterior tenemos:

$$\overline{B} = 6\left[1 + \frac{b}{(1-\lambda)^2}\right]\left[L_1 + \frac{L}{1-\lambda}\,\text{máx}\left\{1,\ \frac{b}{1-\lambda}\right\}\right],$$

b está dado por (3.15),

$$\lambda = \text{máx}\left\{\frac{1+\gamma}{2}, 1 - \frac{1}{\text{máx}\left\{\frac{1-\gamma}{4h}, b_0\right\}}\right\},$$

$$L = \left[k_0\,\text{máx}\left\{\frac{1-\gamma}{4h}, b_0\right\} + k_1\right](k_2 + k_3),$$

y la constante $h = h(H)$ se define en (3.13).

La distancia de Kantoróvich en (3.16) se puede calcular de la siguiente manera:

$$\ell(D_\xi, D_{\widetilde{\xi}}) = \int_{-\infty}^{\infty}\left|\int_{-\infty}^{x} f_\xi(t)dt - \int_{-\infty}^{x} f_{\widetilde{\xi}}(t)dt\right| dx.$$

4.5. Demostración del Teorema

Utilizando la función W del Bloque 1 se introduce el espacio de Banach $\mathbb{B}_W$ de las funciones medibles $u : X \to \mathbb{R}$ para las cuales la norma:

$$\|u\|_W := \sup_{x\in X} \frac{|u(x)|}{W(x)} \text{ es finita.}$$

En el artículo [12] (ver también [3]) se prueba que bajo los Bloques 1 y 2 (y de hecho, bajo condiciones aún más débiles) existen políticas estacionarias óptimasπ_* y $\widetilde{\pi}_*$ con correspondientes costos promedio óptimos J_*, $\widetilde{J}_*$ (see (2.2), (2.3)), así como que los selectores $\varphi, \widetilde{\varphi}$ que definen las políticas π_* y $\widetilde{\pi}_*$ satisfacen las siguientes ecuaciones de optimalidad:

$$\begin{aligned} h(x) &= \inf_{a\in A(x)} \big\{c(x,a) - J_* + Eh\big[F(x,a,\xi)\big]\big\} \\ &= c\big(x,\varphi(x)\big) - J_* + Eh\big[F(x,\varphi(x),\xi)\big], \quad x \in X, \end{aligned} \tag{4,1}$$

$$\begin{aligned} \widetilde{h}(x) &= \inf_{a\in A(x)} \Big\{c(x,a) - \widetilde{J}_* + E\widetilde{h}\big[F(x,a,\widetilde{\xi})\big]\Big\} \\ &= c\big(x,\widetilde{\varphi}(x)\big) - \widetilde{J}_* + E\widetilde{h}\big[F(x,\widetilde{\varphi}(x),\widetilde{\xi})\big], \quad x \in X, \end{aligned} \tag{4,2}$$

donde $h, \widetilde{h}$ son ciertas funciones que pertenecen a $\mathbb{B}_W$.

Para cualquier $(x,a) \in \mathbb{K}$ definimos las siguientes cantidades:

$$H(x,a) := c(x,a) + Eh\big[F(x,a,\xi)\big] - J_*, \tag{4,3}$$

$$\widetilde{H}(x,a) := c(x,a) + E\widetilde{h}\big[F(x,a,\widetilde{\xi})\big] - \widetilde{J}_*. \tag{4,4}$$

Denotaremos por medio de $\{X_t,\ t \geq 0\}$ al proceso de Markov inducido por el uso de la política de control $\widetilde{\pi}_* = \{\widetilde{\varphi}, \widetilde{\varphi}, \dots\}$ en el proceso (1.1). Fijando su "histor" hasta el momento t:

$$\Gamma_t := \{x, a_1, x_1, a_2, \dots, x_{t-1}, a_t\},$$

utilizando (4.1)-(4.4) y la propiedad Markov podemos escribir (cambiando de notación: $\pi :\equiv \widetilde{\pi}_*$ por simplicidad):

$$\zeta_t := E^\pi\big[h(X_t)|\Gamma_t\big] = H(x_{t-1}, a_t) - c(x_{t-1}, a_t) + J_* =$$

$$H(x_{t-1}, a_t) - \inf_{a\in A(x_{t-1})} H(x_{t-1}, a) + h(x_{t-1}) - c(x_{t-1}, a_t) + J_* =$$

$$\Lambda_t - c(x_{t-1}, a_t) + h(x_{t-1}) + J_*,$$

donde

$$\Lambda_t := H(x_{t-1}, a_t) - \inf_{a \in A(x_{t-1})} H(x_{t-1}, a). \tag{4,5}$$

Thus,

$$E_x^\pi h(X_t) = E_x^\pi \zeta_t = E_x^\pi h(X_{t-1}) - E_x^\pi c(X_{t-1}, a_t) + J_* + E_x^\pi \Lambda_t.$$

Sumando las últimas igualdades sobre $t = 1, 2, \ldots, n$ obtenemos:

$$\frac{1}{n} \sum_{t=1}^{n} E_x^\pi c(X_{t-1}, a_t) = J_* + \frac{1}{n} \big[h(x) - E_x^\pi h(X_n) \big] + \frac{1}{n} \sum_{t=1}^{n} E_x^\pi \Lambda_t. \tag{4,6}$$

En [3] se demostró que para cualquier politica estacionaria π el límite del lado izquierdo de (4.6) existe y es finito independiente del estado inicial $x \in X$. Adicionalmente, de [3] se sigue que existe un elemento aleatorio $X(\pi)$ tal que $d_W(X_n, X(\pi)) \to 0$ cuando $n \to \infty$. A partir de esto , y del hecho de que $h \in \mathbb{B}_W$ se obtiene $\frac{1}{n}\ E_x^\pi h(X_n) \to 0$ as $n \to 0$.

Recordando la definición del índice de estabilidad en (1.6), utilizando el hecho de que $J_* = J(\pi_*)$ se sigue de (4.6) que

$$\Delta = \lim_{n \to \infty} \frac{1}{n} \ \sum_{t=1}^{n} E_x^\pi \Lambda_t. \tag{4,7}$$

A partir de (4.2)-(4.5) y de la definición de la política $\pi = \{\widetilde{\varphi}, \widetilde{\varphi}, \ldots\}$ como una óptima para el proceso (1.2) se observa que

$$\Lambda_t = H(x_{t-1}, a_t) - \widetilde{H}(x_{t-1}, a_t) + \inf_{a \in A(x_{t-1})} \widetilde{H}(x_{t-1}, a) - \inf_{a \in A(x_{t-1})} H(x_{t-1}, a).$$

Así (véase (4.3), (4.4)),

$$|\Lambda_t| \le 2 \sup_{a \in A(x_{t-1})} |H(x_{t-1}, a) - \widetilde{H}(x_{t-1}, a)| \le 2|J_* - \widetilde{J}_*| + 2I_t, \tag{4,8}$$

donde

$$I_t := \sup_{a \in A(x_{t-1})} \Big| Eh\big[F(x_{t-1}, a, \xi)\big] - E\widetilde{h}\big[F(x_{t-1}, a, \widetilde{\xi})\big] \Big|. \tag{4,9}$$

Para encontrar una cota superior para I_t observamos que:

$$\begin{aligned} I_t \le & \sup_{a \in A(x_{t-1})} \Big| Eh\big[F(x_{t-1}, a, \xi)\big] - Eh\big[F(x_{t-1}, a, \widetilde{\xi})\big] \Big| + \\ & \sup_{a \in A(x_{t-1})} \Big| Eh\big[F(x_{t-1}, a, \widetilde{\xi})\big] - E\widetilde{h}\big[F(x_{t-1}, a, \widetilde{\xi})\big] \Big|. \end{aligned} \tag{4,10}$$

Para evaluar el segundo sumando del lado derecho de (4.10) se introduce la siguiente pareja de operadores $T, \widetilde{T} : \mathbb{B}_W \to \mathbb{B}_W$:

$$Tu(x) := \inf_{a \in A(x)} \left\{ c(x,a) - J_* + Eu\big[F(x,a,\xi)\big] - \phi(x,a) \int_X u d\nu \right\}, \ x \in X, \tag{4,11}$$

$$\widetilde{T}u(x) := \inf_{a \in A(x)} \left\{ c(x,a) - \widetilde{J}_* + Eu\big[F(x,a,\widetilde{\xi})\big] - \phi(x,a) \int_X u d\nu \right\}, \ x \in X.$$

En [12] bajo el Bloque 1 se demostró que T y $\widetilde{T}$ efectivamente mapean $\mathbb{B}_W$ sobre $\mathbb{B}_W$, y adicionalmente, que dichos operadores son contractivos con módulo λ respecto a la norma $\| \cdots \|_W$. Por otro lado, en el artículo antes mencionado se demuestra también que

$$Th = h; \qquad \widetilde{T}\widetilde{h} = \widetilde{h}, \tag{4,12}$$

donde h y $\widetilde{h}$ son respectivamente soluciones de las ecuaciones (4.1) y (4.2).

Regresando a (4.10) observamos que

$$\sup_{a \in A(x_{t-1})} \left| Eh\big[F(x_{t-1},a,\widetilde{\xi})\big] - E\widetilde{h}\big[F(x_{t-1},a,\widetilde{\xi})\big] \right| \frac{W\left[F(x_{t-1},a,\widetilde{\xi})\right]}{W\left[F(x_{t-1},a,\widetilde{\xi})\right]} \le$$

$$\|h - \widetilde{h}\|_W \sup_{a \in A(x_{t-1})} EW\big[F(x_{t-1},a,\widetilde{\xi})\big]. \tag{4,13}$$

Iterando la segunda desigualdad en el Bloque 1, (c) se obtiene que para cada $t \ge 1, x \in X$,

$$\sup_{a \in A(x_{t-1})} EW\big[F(x_{t-1},a,\widetilde{\xi})\big] \le \lambda^t W(x) + \frac{b}{1-\lambda}, \tag{4,14}$$

donde b es la constante del Bloque 1, (d).

Respecto al primer factor del lado derecho de (4.13) tenemos:

$$\|h - \widetilde{h}\|_W \le \|Th - \widetilde{T}h\|_W + \|\widetilde{T}h - \widetilde{T}\widetilde{h}\|_W \le \|Th - \widetilde{T}h\|_W + \lambda \|h - \widetilde{h}\|_W,$$

o

$$\|h-\widetilde{h}\|_W \leq \frac{1}{1-\lambda}\sup_{x\in X}\frac{1}{W(x)}\sup_{a\in A(x)}\Big|-J_*+Eh\big[F(x,a,\xi)\big]+\widetilde{J}_*-Eh\big[F(x,a,\widetilde{\xi})\big]\Big| \leq$$

$$\frac{1}{1-\lambda}\left[\big|J_*-\widetilde{J}_*\big|+\sup_{(x,a)\in\mathbb{K}}\Big|Eh\big[F(x,a,\xi)\big]-Eh\big[F(x,a,\widetilde{\xi})\big]\Big|\right]. \tag{4,15}$$

Combinando las desigualdades (4.8), (4.10), (4.13), (4.14) y (4.15) obtenemos la siguiente desigualdad:

$$\begin{aligned}|\Lambda_t| \leq 2\Bigg\{&\left(1+\frac{b}{(1-\lambda)^2}\right)|J_*-\widetilde{J}_*|+\left(1+\frac{b}{(1-\lambda)^2}\right)Q+\\ &\lambda^t W(x)\left[\frac{1}{1-\lambda}|J_*-\widetilde{J}_*|+\frac{1}{1-\lambda}Q\right]\Bigg\},\end{aligned} \tag{4,16}$$

donde

$$Q := \sup_{(x,a)\in\mathbb{K}}|Eh\big[F(x,a,\xi)-Eh\big[F(x,a,\widetilde{\xi})\big]|. \tag{4,17}$$

Ahora se demostrará que para cada $(x,a)\in\mathbb{K}$ la función $h[F(x,a,\bullet)]$ in (4.17) es Lipschitz sobre S con la constante

$$\bar{L} := (L_0+1)L_*\left[L_1+\frac{L}{1-\lambda}\ \text{máx}\left\{1,\ \frac{b}{1-\lambda}\right\}\right]. \tag{4,18}$$

UUtilizando (4.12) y la propiedad contractiva de T se obtiene:

$$\|h\|_W = \|Th-T0+T0\|_W \leq \lambda\|h\|_W+\|T0\|_W.$$

De esa manera, en vista de (4.11),

$$\|h\|_W \leq \frac{1}{1-\lambda}\sup_{x\in X}\frac{1}{W(x)}\left|\inf_{a\in A(x)}c(x,a)-J_*\right|. \tag{4,19}$$

Utilizando el Bloque 1, (a) observamos que para la política estacionaria óptima $\pi_*=(\varphi,\varphi,\dots)$ se sigue:

$$\begin{aligned}J_* &= \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}E_x^{\pi_*}c\big(X_{t-1},\varphi(X_{t-1})\big) \leq\\ &\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}E_x^{\pi_*}W(X_{t-1}) \leq \frac{b}{1-\lambda},\end{aligned} \tag{4,20}$$

dado que, de manera análoga a(4.14),

$$EW(X_{t-1}) \leq \lambda^{t-1} W(x) + \frac{b}{1-\lambda}.$$

A partir de (4.19) y (4.20) obtenemos:

$$\|h\|_W \leq \frac{1}{1-\lambda} \ \text{máx}\left\{1, \ \frac{b}{1-\lambda}\right\} =: r. \tag{4,21}$$

Determinemos ahora la constante $K < \infty$ tal que la siguiente función involucrada en la ecuación (4.1):

$$g(\mathbf{k}) := c(\mathbf{k}) + Eh\big[F(\mathbf{k}, \xi)\big]$$

es de Lipschitz con constante K. Para arbitrarias pero fijas $\mathbf{k} = (x, a), \mathbf{k}' = (x', a') \in \mathbb{K}$ podemos verificar que (ver Bloque 2, (b), (c)) que dado (4.21) y la definición de la métrica d_W en (2.1),

$$\begin{aligned} |g(\mathbf{k}) - g(\mathbf{k})| \leq L_1 \mu(\mathbf{k}, \mathbf{k}') + |Eh[F(\mathbf{k}, \xi)] - Eh[F(\mathbf{k}', \xi)]| \leq \\ L_1 \mu(\mathbf{k}, \mathbf{k}') + rL\mu(\mathbf{k}, \mathbf{k}') \leq K\mu(\mathbf{k}, \mathbf{k}') \end{aligned}$$

where $K := L_1 + rL$.

Usando la ecuación (4.1) y el Bloque 2, (a), (d) observamos que para probar que h es de Lipschitz con la constante $(L_0+1)L_* K$ it basta fundamentar la siguiente afirmación:

Para una función $g : \mathbb{K} \to \mathbb{R}$ tal que

$$|g(\mathbf{k}) - g(\mathbf{k}')| \leq K\mu(\mathbf{k}, \mathbf{k}'), \ \ \mathbf{k}, \mathbf{k}' \in \mathbb{K}$$

se sigue

$$g(x, y) := \left| \inf_{a \in A(x)} g(x, a) - \inf_{a \in A(y)} g(y, a) \right| \leq (L_0 + 1) K \rho(x, y)$$

para cada $x, y \in X$.

Efectivamente,

$$\begin{gathered} g(x, y) \leq \left| \inf_{a \in A(x)} g(x, a) - \inf_{a \in A(x)} g(y, a) \right| + \\ \left| \inf_{a \in A(x)} g(y, a) - \inf_{a \in A(y)} g(y, a) \right| \leq K\rho(x, y) + I. \end{gathered}$$

Se demostrará que

$$I := \left| \inf_{a \in A(x)} g(y,a) - \inf_{a \in A(y)} g(y,a) \right| \leq K L_0 \rho(x,y).$$

Supongamos que éste no es el caso, entonces existe $\epsilon > 0$ tal que $I > KL_0\rho(x,y) + \epsilon$.

Sea, por ejemplo,

$$I = \inf_{a \in A(x)} g(y,a) - \inf_{a \in A(y)} g(y,a).$$

Dada la compacidad de $A(x), A(y)$ y la continuidad de g, existen $a_x \in A$, $a_y \in A_y$ para los cuales los ínfimos anteriores se alcanzan. De esa manera

$$g(y, a_x) > g(y, a_y) + KL_0\rho(x,y) + \epsilon.$$

Since $h\left[A(x), A(y)\right] \leq L_0\rho(x,y)$, there exists $\bar{a}_x \in A(x)$ tal que $d(a_y, \bar{a}_x) \leq L_0\rho(x,y)$, y consecuentemente

$$|g(y, \bar{a}_x) - g(y, a_y)| \leq KL_0\rho(x,y).$$

Comparando ambas desigualdades tenemos que $g(y, a_x) > g(y, \bar{a}_x) + \epsilon$, contradiciendo el hecho de que a_x es el elemento para el cual el ínfimo de $g(y,a)$ sobre $a \in A(x)$ se alcanza.

Aplicando la definición de la métrica de Kantorovich en (1.8), vemos que(4.17)

$$Q \leq \bar{L}\ell(D_\xi, D_{\widetilde{\xi}}), \tag{4,22}$$

donde $\bar{L}$ es la constante de (4.18).

Comparando (4.7), (4.16) y (4.22) se deduce que

$$\Delta \leq 2\left(1 + \frac{b}{(1-\lambda)^2}\right)\left[\bar{L}\ell(D_\xi, D_{\widetilde{\xi}}) + |J_* - \widetilde{J}_*|\right]. \tag{4,23}$$

La siguiente tarea consiste en acotar la magnitud$|J_* - \widetilde{J}_*|$ en términos de la distancia de Kantorovich entre ξ y $\widetilde{\xi}$. Para lograrlo utilizaremos el **vanishing discount approach** y algunos resultados apropiados de [4].

Fijemos $\alpha \in (0,1)$, y para un estado inicial$x \in X$ y una política $\pi \in \sqcap$ definamos el costo total descontado para los procesos(1.1) y (1.2), como respectivamente:

$$V_\alpha(x,\pi) := E_x^\pi \sum_{t=1}^{\infty} \alpha^{t-1} c(x_{t-1}, a_t),$$

$$\widetilde{V}_\alpha(x,\pi) := E_x^\pi \sum_{t=1}^{\infty} \alpha^{t-1} c(\widetilde{x}_{t-1}, \widetilde{a}_t).$$

Las correspondientes funciones de valor se definen de la siguiente manera:

$$V_\alpha^*(x) := \inf_{\pi\in\Pi} V_\alpha(x,\pi), \qquad x \in X;$$

$$\widetilde{V}_\alpha^*(x) := \inf_{\pi\in\Pi} \widetilde{V}_\alpha(x,\pi), \qquad x \in X.$$

Para el Bloque 1, (c), (d) y los resultados en [4] se sigue que

$$\|V_\alpha^* - \widetilde{V}_\alpha^*\|_{W_\alpha} \leq (1-\beta)^{-1} \|T_\alpha V_\alpha^* - \widetilde{T}_\alpha V_\alpha^*\|_{W_\alpha} \tag{4,24}$$

where $W_\alpha(x) := W(x) + b\,\dfrac{2\alpha}{1-\alpha}$,

$$T_\alpha u(x) := \inf_{a\in A(x)} \big\{ c(x,a) + \alpha E u\big[F(x,a,\xi)\big] \big\}, \qquad x \in X$$

(y $\widetilde{T}_\alpha$ being defined similarly).

Utilizando (4.24) es inmediato demostrar que

$$|V_\alpha^*(z) - \widetilde{V}_\alpha^*(z)| \leq \frac{2}{1-\alpha}\, \psi(\alpha)\, Q_\alpha(\xi, \widetilde{\xi}), \tag{4,25}$$

donde

$$\psi(\alpha) \to 1 \quad \text{as} \quad \alpha \to 1, \tag{4,26}$$

y

$$Q_\alpha(\xi,\widetilde{\xi}) := \sup_{(x,a)\in\mathbb{K}} \Big| E V_\alpha^*\big[F(x,a,\xi)\big] - E V_\alpha^*\big[F(x,a,\widetilde{\xi})\big] \Big| . \tag{4,27}$$

Se sigue de los resultados en[3] que bajo los Bloques 1 y 2, una sucesión $\{\alpha_n\} \subset (0,1)$ se puede elegir de tal manera que cuando $n \to \infty$:

- $\alpha_n \to 1$;
- $(1-\alpha_n) V_{\alpha_n}^*(z) \to J_*$; (4.28)
- $(1-\alpha_n) \widetilde{V}_{\alpha_n}^*(z) \to \widetilde{J}_*$; (4.29)
- $V_{\alpha_n}^*(x) - V_{\alpha_n}^*(z) \to h'(x), \quad x \in X,$ (4.30)

para cierta función $h' \in \mathbb{B}_W$ que satisface la ecuación:

$$h'(x) = \inf_{a\in A(x)} \big\{ c(x,a) - J_* + E h'\big[F(x,a,\xi)\big] \big\} . \tag{4,31}$$

Dado que $h' \in \mathbb{B}_W$ y $\displaystyle\int_X W d\nu < \infty$, tenemos que $\gamma := \displaystyle\int_X h'(x)\nu(dx)$ es un número finito.

Using (4.25) - (4.29) we establish that for every integer N,

$$
\begin{aligned}
&|J_* - \widetilde{J}_*| \leq 2 \limsup_{n\to\infty} Q_{\alpha_n}(\xi, \widetilde{\xi}) \leq \\
&2 \sup_{n\geq N} \sup_{(x,a)\in\mathbb{K}} \left| Eh_{\alpha_n}\big[F(x,a,\xi)\big] - Eh_{\alpha_n}\big[F(x,a,\widetilde{\xi})\big] \right|,
\end{aligned}
\tag{4,32}
$$

donde $h_{\alpha_n}(x) := V^*_{\alpha_n}(x) - V^*_{\alpha_n}(z) - \gamma$, $x \in X$.

Por (4.30) $h_{\alpha_n}(x) \to h'(x) - \gamma$, $x \in X$, cuando $n \to \infty$. Por lo tanto la función $h := h' - \gamma$ es también solución de la ecuación (4.31), y dado que $\int_X h(x)\nu(dx) = 0$, por [12] se deduce que$Th = h$, para el operador contractivo T definido en (4.11).

Por otro lado, dado que para cada α, $V^*_\alpha = T_\alpha V^*_\alpha$ ([3,7]), se sigue que para cada n, $x \in X$:

$$
\begin{aligned}
h_{\alpha_n}(x) \;\; = \;\; & \inf_{a\in A(x)} \Big\{ c(x,a) - (1-\alpha_n)V^*_{\alpha_n}(z) \\
& + \alpha_n E h_{\alpha_n}\big[F(x,a,\xi)\big] - (1-\alpha_n)\gamma \Big\},
\end{aligned}
\tag{4,33}
$$

or

$$
h_{\alpha_n} = T'_{\alpha_n} h_{\alpha_n},
$$

donde el operador $T'_{\alpha_n} : \mathbb{B}_W \to \mathbb{B}_W$ está definido por el lado derecho de

(4.33). Así

$$
\begin{aligned}
&\|h - h_{\alpha_n}\|_W = \|Th - T'_{\alpha_n} h_{\alpha_n}\|_W \leq \\
&\|Th - Th_{\alpha_n}\|_W + \|Th_{\alpha_n} - T'_{\alpha_n} h_{\alpha_n}\|_W \leq \\
&\lambda \|h - h_{\alpha_n}\|_W + \|Th_{\alpha_n} - T'_{\alpha_n} h_{\alpha_n}\|_W, \quad \text{or} \\
&\|h - h_{\alpha_n}\|_W \leq \frac{1}{1-\lambda} \|Th_{\alpha_n} - T'_{\alpha_n} h_{\alpha_n}\|_W \leq \\
&\frac{1}{1-\lambda} \sup_{x \in X} \frac{1}{W(x)} \Bigg\{ \quad \big| J_* - (1-\alpha_n) V^*_{\alpha_n}(z) \big| \\
&+ \sup_{a \in A(x)} (1-\alpha_n) | E h_{\alpha_n} \big[F(x,a,\xi) \big] | + \\
&(1-\alpha_n)\gamma + \sup_{a \in A(x)} \phi(x,a) \int_X h_{\alpha_n}(x) \nu(dx) \Bigg\}.
\end{aligned}
\tag{4,34}
$$

El primer sumando del lado derecho de (4.34) tiende a cero cuando $n \to \infty$ debido a (4.28). En [3] se demuestra que existe una constante $\bar{K}$ tal que para cada $n \in X$,

$$
\big| V^*_{\alpha_n}(x) - V^*_{\alpha_n}(z) \big| \leq \bar{K} W(x). \tag{4,35}
$$

Tomando en cuenta esta desigualdad y el Bloque 1, (c), (d) concluimos que los segundo (y también tercer) sumandos del lado derecho de (4.34) tienden a cero cuando $n \to \infty$. Además, por la continuidad de ϕ, (4.35), Bloque 1, (d), el Teorema de Convergencia Dominada de Lebesgue y el hecho de que $\int_X h d\nu = 0$, se obtiene que el último sumando en (4.34) converge a cero cuando $n \to \infty$. Por lo tanto, para cada $\epsilon > 0$ podemos elegir una N tal que para $n \geq N$

$$
\|h_{\alpha_n}\|_W \leq \|h\|_W + \|h - h_{\alpha_n}\|_W \leq r + \epsilon, \tag{4,36}
$$

donde la constante r se definió en (4.21).

Utilizando (4.36), la ecuación (4.33), el Bloque 2, (a), (b), (c) y repitiendo el argumento utilizado para establecer la propiede de Lipschitz de h, encontramos que para $n \geq N$ la función h_{α_n} es Lipschitz con la constante $(L_0 + 1)\big(L_1 + L(r+\epsilon)\big)$, y en consecuencia (ver Bloque 2, (d)) para cada$n \geq N$, $(x,a) \in \mathbb{K}$ la función $h_{\alpha_n}\big[F(x,a,\bullet)\big]$ es Lipschitz sobre S con la constante $L_*(L_0+1)\big(L_1+L(r+\epsilon)\big)$. De esa manera, a partir de (1.8) y (4.32) se sigue

$$|J_* - \widetilde{J}_*| \leq 2L_*(L_0+1)\big(L_1+L(r+\epsilon)\big)\ell(D_\xi, D_{\tilde{\xi}}). \tag{4,37}$$

Dado que ϵ es artibrario, podemos tomar $\epsilon = 0$ en la última desigualdad.

Combinando las desigualdades (4.23) y (4.37)se logra finalmente demostrar la desigualdad de estabilidad deseada (2.4).

∎

Bibliografía

[1] Dynkin EB, Yushkevich AA (1979) Controlled Markov processes. Springer, New York.

[2] Gordienko E (1992). An estimate of the stability of optimal control of certain stochastic y deterministic systems. J. Soviet Math. 59: 891-899.

[3] Gordienko E, Hernández-Lerma O (1995) Average cost Markov control processes with weighted norms: existence of canonical policies. Appl. Math. 23: 199-218.

[4] Gordienko E, Lemus-Rodríguez E, Montes-de-Oca R (2008) Discounted cost optimality problem: stability with respect to weak metrics. Math. Methods Oper. Res. In press.

[5] Gordienko E, Salem F (2000) Estimates of Stability of Markov control processes with unbounded cost. Kybernetika 36: 195-210.

[6] Gordienko E, Yushkevich A (2003) Stability estimates in the problem of average optimal switching of a Markov chain. Math. Methods Oper. Res. 57:345-365.

[7] Hernández-Lerma O, Lasserre JB (1996) Discrete-Time Markov Control Processes. Basic Optimality Criteria. Springer, New York.

[8] Montes-de-Oca R, Salem-Silva F (2005) Estimates for perturbations of average Markov decision processes with a minimal state and upper bounded by stochastically ordered Markov chains. Kybernetika 41:757-772.

[9] Rachev ST, Rüschendorf L (1998) Mass transportation problem, Vol II: Applications. Springer, New York.

[10] Van Dijk NM (1988) Perturbation theory for unbounded Markov reward processes with applications to queueing. Adv. In Appl. Probab. 20:91-111.

[11] Van Dijk NM, Sladky K (1999) Error bounds for nonnegative dynamic models. J. Optim Theory Appl. 101: 449-474.

[12] Vega-Amaya O (2003) The average cost optimality equation: a fixed point approach. Bol. Soc. Mat. Mexicana 9:185-195.

Capítulo 5

Conclusiones

A lo largo de estos años dedicados a la investigación doctoral, se puede, afortunadamente, hablar de suficientes problemas como para tener una razonable esperanza de estar ante una fecunda, robusta y longeva línea de investigación. Describiremos escuetamente los problemas esenciales.

5.1. Estabilidad: Hadamard, problemas bien planteados y la no-unicidad.

Un problema interesante que plantea la reflexión sobre la estabilidad en problemas de control es su contraste brusco con teorias clásicas de estabilidad como la de Hadamard en el contexto de las ecuaciones en derivadas parciales. Recordemos que dicho concepto está implícito en la concepción de Hadamard de cuándo un problema está bien planteado (es el tercer punto):

1. La solución existe.
2. Es única.
3. Depende de manera continua (en alguna topologia adecuada) de los parámetros.

De hecho, es claro que en el caso de modelos estocásticos encontrar la topologia adecuada es un reto tanto matemático como filosófico y de modelación.

Pero nos llama mucho la atención el punto dos: que la solución sea única. Si nuestra ecuacion diferencial aspira a servir como instrumento para modelar un aspecto de la realidad, es claro que la falta de unicidad planteaaria un gran problema: ¿cuál de las soluciones realmente deberiamos tomar como modelo? Sin embargo, cuando, en un segundo paso, pasamos de modelar a tomar una decisión es posible, y de hecho, ocurre en muchos casos, que la solución no es única. Por ejemplo, en una esfera existen infinitas trayectorias de un punto a su antipoda que tienen longitud mínima. En una cultura matemática que durante siglos ha cultivado la unicidad como una virtud, incluso al punto de forzarla via clases de equivalencia, el alto grado de no-unicidad que presentan diversos problemas de optimización plantea una anomalía conceptual. Pero desde el punto de vista del decisor esa no-unicidad puede ser buena, tener más alternativas puede, como se hace en algunas ocasiones, permitir optimizar lexicográficamente, es decir, primero optimizar el criterio básico y luego seleccionar a partir de un segundo criterio, etc. El problema que creemos se plantea en este contexto es analizar la estabilidad de la no-unicidad: ante pequeñas variaciones, cuándo se conserva la no-unicidad y en qué casos, por el contrario, la no-unicidad es inestable, es decir, toda pequeña perturbación genera un nuevo problema que sí tiene óptimo único. La respuesta a estas cuestiones permitiria en particular extender la noción de problema bien planteado a los problemas de optimización.

5.2. Coherencia de problemas, subproblemas y superproblemas.

Existe otro contexto en el que la estabilidad puede tener sentido, que podriamos denominar meta-optimización al menos en esta sección. Nos referimos a los cambios que puede sufrir el óptimo se el problema original se reduce o se aumenta (es decir, qué tanto difieren los óptimos del problema, de subproblemas y superproblemas).

Pensemos por ejemplo en el horizonte de control: si aumenta o disminuye el intervalo de tiempo en el cual el sistema esta bajo control, ¿qué ocurre con el óptimo? ¿en qué sentidos podría ser o no estable? ¿bajo qué condiciones?

Situaciones similares surgen si aumenta o disminuye el conjunto de acciones admisibles en cada estado.

5.3. La Métrica de Kantorovich

Tal como se detalla en el capítulo 3, el trabajo previo sobre el tema de estabilidad se había realizado en el caso de la métrica de Variación Total y la Métrica de Variación Total Ponderada. Estas métricas excluyen el estudio de la estabilidad a distribuciones de probabilidad que sean mutuamente singulares, y por ende, no permiten analizar la aproximación de una medida dada por una densidad utilizando una medida concentrada en un conjunto finito. Sin embargo, existen otras métricas que sí permiten estudiar este caso, como la de Kantoróvich. Pero la métrica de Kantoróvich no es la única, y surge la pregunta natural de porqué elegirla. Podemos aducir las siguientes razones:

- Esta métrica metriza la convergencia débil de medidas.
- Tiene una definición matemática elegante.
- Su definición involucra la familia de funciones de Lipschitz.
- Para medidas en $\mathbb{R}$ admite una formulación muy cómoda en términos de las funciones de distribución.
- Se conocen diversas desigualdades que la relacionan con otras métricas.
- Está íntimamente relacionada con el problema de Transporte Óptimo de Masa de Monge-Kantoróvich.
- Existen una literatura relativamente accesible y clara sobre el tema.

Por supuesto, ésto no implica que no sea de interés analizar el problema de estabilidad respecto a otras métricas. En ese sentido, este sería un primer problema abierto. Mucho de la dificultad estriba en la gran variedad de métricas existentes y el escaso orden de la literatura sobre el tema.

Para concluir esta sección, es relevante plantear el siguiente problema: ¿es posible analizar el problema correspondiente de estabilidad cuándo el sistema aproximante es una discretización del sistema original? El poder utilizar una métrica como la de Kantoróvich tiene la ventaja de que permite analizar un sistema continuo aproximándolo por uno finito. Pareciera entonces que este problema ya está básicamente resuelto por los resultados presentados en la tesis. Lamentablemente no es así: un control óptimo del sistema finito en principio no se puede aplicar directamente en el sistema original. Es decir, al problema hay que añadirle la información que permite interpolar un control del sistema original a partir del control del sistema discretizado. Es plausible que por ejemplo, se puede intentar resolver este problema en esquemas de discretización como el de Kushner et al., que discretizan una difusión controlada al aproximarla por una Cadena Controlada de Márkov.

En discusiones con el Dr. Andrés Christen, surge la pregunta de si es posible replantear el problema en un enfoque adaptivo y un contexto Bayesiano al estudiar la relación entre la justificación de la teoría Bayesiana por medio de la teoría de la información, y la métrica de Kantoróvich. Por ejemplo, si μ_0 es una medida a priori que aproxima a μ y por medio de una muestra X obtenemos una medida a posteriori $\mu(X)$, ¿estará más $\mu(X)$ de μ en la métrica de Kantoróvich que μ_0 de μ? Si se analiza el paso de la medida a priori a la a posteriori como un operador, ¿es este operador una contracción respecto a la métrica de Kantoróvich?

5.4. Consideraciones Finales

Tal como se planteó al inicio del texto, el objetivo de la presente tesis no es meramente comunicar por escrito los resultados obtenidos sobre el problema doctoral. También queremos lograr motivar el interés por el tema y facilitar el acercamiento a cualquiera que le llame la atención. Por esa razón algunas consideraciones son relevantes.

El tema de la estabilidad es un tema que no ha recibido la importancia que merece, al encontrarse en la frontera entre las matemáticas y los campos en los que se aplica el control. Precisamente por esa razón presenta un terreno fértil para la investigación. Aunque en México aún existe el prejuicio de diversos matemáticos que se mantienen alejados de sus colegas de otras disciplinas, ahora es más urgente que nunca ese acercamiento. El control estocástico presente un muy buen contexto para lograrlo.

Por otro lado, los problemas matemáticos planteados gracias a dicha interacción son interesantes per se y requieren matemáticas como análisis funcional y teoría de la medida, pero en un contexto donde su utilidad se renueva. Es sorprendente cómo una técnica tan clásica como el Teorema del Punto Fijo de Banach aún guarda mucha vitalidad. Surge la pregunta obligada ¿tendrán aplicación similar algunos de los otros teoremas de Punto

Fijo?

En el plano propio del Control Estocástico, se utilizó de manera crucial el método de aproximaciones sucesivas que en este contexto se conoce como Iteración de Valores. Pero existe otro método utilizado con cierta frecuencia en la práctica pero que no admite un análisis matemático tan directo, conocido como Iteración de Políticas. Sin embargo, se sabe que en cierto sentido Iteración de Políticas en un Método tipo Newton. Un problema interesante consistiría en revisar los resultados de los dos capítulos anteriores en el contexto de Iteración de Políticas: ¿aún se puede utilizar la técnica de los operadores contractivos?

No está de más comentar que esta tesis se realizó en el marco de una comunidad de nuevos y entusiastas investigadores que están consolidando el estudio de diversos problemas de modelación y optimización probabilísticas, entre los cuales se pueden mencionar muchos nombres como Hugo Adán Cruz, Heliodoro Cruz, Francisco Salem, Oscar Vega, Juan González y muchos otros. De entre los nombres ya consolidados como los doctores Onésimo Hernández-Lerma, Rolando Cavazos y mi asesor E. Gordienko, es muy grato también mencionar nuevos líderes como el Dr. Raúl Montes de Oca.

Aprovecho finalmente la oportunidad para volver a agradecer al Prof. E. Gordienko por su gran generosidad intelectual y humana, al compartir tanto durante la elaboración de la tesis, en particular, al ligarnos a la tradición matemática rusa en probabilidad que se remonta de Chebyshev, Lyapunov, Markov a Jinchin llegando a Kolmogórov y su discípulo Prójorov, de quien el Prof. Gordienko fue alumno. Spasibo!

Bibliografía

[1] R. Montes-de-Oca, A. Sakhanenko y F. Salem-Silva *Estimates for Perturbations of General Discounted Markov Control Chains* Applicationes Mathematicae **30**, 3 (2003) 287–304

[2] E. Gordienko *An Estimate of Certain Stability of Optimal Control of Certain Stochastic and Deterministic Systems* J. Soviet Math. **59** (1992) 891–899

[3] O. Hernández-Lerma y J.B. Lasserre *Further Topics on Discrete-Time Markov Control Processes* Springer Verlag 1999

[4] R.M. Dudley, E. Giné y J. Zinn *Uniform and Universal Glivenko-Cantelli Classes* Journal of Theoretical Probability **4** (1991) 485–510

[5] A. L. Gibbs y F.E. Su *On choosing and bounding probability metrics.* Manuscrito (2002)

[6] V. M. Zolotarev *Probability Metrics* Theory Probab. Appl. **28** (1983) 278–302

[7] E. Gordienko A. Yushkevich *Stability estimates in the problem of average optimal switching of a Markov chain.* Math. Meth. Oper. Res. **57** (2003) 345–365

[8] O. Vega-Amaya *The average cost optimality equation: a fixed point approach.* Bol. Soc. Mat. Mexicana **9** (2003) 185–195

[9] S.P. Meyn y R. L. Tweedie *Markov Chains and Stochastic Stability* Springer Verlag 1993

[10] E. Gordienko y F. Salem *Estimates of stability of Markov control processes with unbounded costs.* Kybernetika **36** (2000) 195–210

[11] E. Gordienko y E. Lemus-Rodríguez *Estimation of robustness for controlled diffusion processes.* Stochastic Anal. Appl. **17** (1999) 421–441

[12] E. Gordienko *Uniform exponential convergence of Markov processes in metrics corresponding to a weak topology.* Theory Probab. Appl. **28** (1993) 570–571 (in russian)

[13] S. T. Rachev y L. Rüschendorf *Mass transportation problems vol. II : Applications* Springer Verlag 1998

[14] R. M. Dudley *The speed of mean Glivenko-Cantelly convergence.* Annals of Math. Stat. **40** (1969) 40-50

[15] J. Horowitz R.L. Karandikar *Mean rates of convergence of empirical measures in the Wasserstein metric.* J. Comput. Appl. Math. **55** (1994) 261–273

[16] S.T. Rachev L. Rüschendorf *Probability metrics and recursive algorithms.* J. Appl. Probab. **27** (1995) 770–799

[17] O. Hernández-Lerma y J.B. Lasserre *Markov Chains and Invariant Probabilities.* Birkhäuser 2003

[18] T.R. Bielecki y J.A. Filar *Singularly perturbed Markov control problems: limiting average cost.* Ann. Oper. Research **28** (1991) 153–168

[19] M.L. Puterman *Markov Decision Processes* Wiley 1994

[20] S.T. Rachev *Probabilty Metrics and the Stability of Stochastic Models* Wiley 1991