EL METODO DE REGENERACION PARA LA ESTIMACION DE LA ESTABILIDAD DE PROCESOS CONTROLABLES CON COSTO DESCONTADO

TESIS QUE PRESENTA:

TORRES RIOS NORBERTO

PARA OBTENER EL GRADO DE MAESTRO EN MATEMATICAS

DIRECTOR: Dr. EVGENI I. GORDIENKO

UNIVERSIDAD AUTONOMA MATROPOLITANA-IZTAPALAPA
DIVICION DE CIENCIAS BASICAS E INGENIERIA
DEPARTAMENTO DE MATEMATICAS

MARZO DE 1998

EL METODO DE REGENERACION PARA LA ESTIMACION DE LA ESTABILIDAD DE PROCESOS CONTROLABLES CON COSTO DESCONTADO

TESIS QUE PRESENTA:

TORRES RIOS NORBERTO

PARA OBTENER EL GRADO DE MAESTRO EN MATEMATICAS

Deseo expresar mi reconocimiento y agradecimiento al Dr. Evgeni I. Gordienko por dirigir esta tesis y al Dr. Raúl Montes de Oca M. por sus valiosas sugerencias y comentarios.

Deseo también manifestar mi agradecimiento a CONACYT por su apoyo económico y al departamento de Matemáticas de la UAM-I por las facilidaes que me brindó.

INDICE

INTRODUCCION

En los problemas aplicados de control estocástico se encuentran, como regla, incertidumbres en la información de entrada, la cual se origina por la falta de información precisa, por ejemplo, sobre la ecuación que describe la dinámica del "sistema real" de control (o sobre la distribución de los vectores aleatorios, involucrados en la descripción de la evolución de este sistema), el cual denotaremos por SR. Consideraremos otro sistema de control que aproxima al sistema SR, que denotaremos por $\widetilde{SA}$, suponiendo que este sistema es conocido completamente (es decir, se tiene toda la información sobre este proceso) y que un controlador tiene en mente, el fin de encontrar la política óptima para el sistema SR, contando con el sistema aproximante $\widetilde{SA}$.

Se encuentra un incremento en el costo, en la situación cuando una política de control escogida por medio de datos aproximados, es decir, por medio del sistema $\widetilde{SA}$, se aplica para controlar al sistema original SR. Esta situación surge frecuentemente en las aplicaciones cuando se usan aproximaciones de las distribuciones de probabilidad en SR, mediante estimaciones estadísticas.

La estimación del aumento del costo total descontado, en términos de la precisión de la aproximación, es la tarea del estudio de la estabilidad que se realiza en esta tesis.

En este trabajo consideramos sistemas dados por procesos de control de Markov (PCM's), a tiempo discreto, definidos de la siguiente manera:

$$(\mathrm{SR}) \qquad \mathrm{x}_{t+1} = \Phi(\mathrm{x}_t, a_t, \xi_t), \qquad t = 0, 1, ... \tag{0.1}$$

$$(\widetilde{\mathrm{SA}}) \qquad \tilde{\mathrm{x}}_{t+1} = \Phi(\tilde{\mathrm{x}}_t, \tilde{a}_t, \tilde{\xi}_t), \qquad t = 0, 1, ... \tag{0.2}$$

Por lo tanto SA y $\widetilde{\mathrm{SR}}$, representan **procesos controlables** y, en (0.1) y (0.2), denotaremos:

x_t, $\tilde{\mathrm{x}}_t \in$ X son estados del proceso al tiempo t.

$a_t, \tilde{a}_t \in$ A son controles (acciones) aplicadas en el tiempo t.

$\{\xi_t\}, \{\tilde{\xi}_t\}$, llamadas a la vez "sucesiones gobernantes", son sucesiones de vectores (en R^n) aleatorios independientes e idénticamente distribuidos (v.a'.s i.i.d.), sus distribuciones comúnes las denotaremos por F y $\widetilde{\mathrm{F}}$, respectivamente.

Los procesos (0.1) y (0.2) están definidos sobre los mismos espacios generales de Borel. También, la función c(x,a), x $\in$ X, a $\in$A, la cual supondremos acotada, está dada para describir el costo por etapa para ambos procesos (0.1) y (0.2). La única diferencia entre estos procesos, son las funciones de distribución F y $\widetilde{\mathrm{F}}$, de manera que la **distribución conocida** $\widetilde{\mathrm{F}}$, se utiliza como la aproximación de la distribución F, **conocida parcialmente** (o **desconocida** como en los problemas de control adaptado, véase [8, 13, 14]).

Como criterio para optimizar políticas de control $\pi = \{a_t, t = 0, 1, ...\}$ con horizonte infinito, se considerará, en este trabajo, el índice de **costo total descontado,** el cual denotaremos por $Q\beta(\mathrm{x}, \pi)$ y $\tilde{Q}\beta(\mathrm{x}, \tilde{\pi})$ para los procesos (0.1) y (0.2), respectivamente, (véase las definiciones de estos índices en el siguiente capítulo). En las notaciones de los costos descontados, introducidos anteriormente, se subraya la dependencia del costo con el estado inicial, $\mathrm{x}_0 = \mathrm{x}$, $\tilde{\mathrm{x}}_0 = \mathrm{x} \in \mathrm{X}$, de los procesos, de la política π aplicada y del factor de descuento $\beta \in (0, 1)$, dado.

Las políticas óptimas π_* y $\tilde{\pi}_*$ para los procesos (0.1) y (0.2), respectivamente, están definidas como sigue (las cuestiones de la existencia de éstas se posponen hasta el capítulo 2):

$$
\begin{aligned}
Q\beta(\mathrm{x}, \pi_*) &= \inf \pi Q\beta(\mathrm{x}, \pi) \quad \text{para cada} \quad \mathrm{x} \in \mathrm{X} \\
\tilde{Q}\beta(\mathrm{x}, \tilde{\pi}_*) &= \inf \pi \tilde{Q}\beta(x, \pi) \quad \text{para cada} \quad \mathrm{x} \in \mathrm{X}
\end{aligned}
$$

A pesar de que el problema de optimización original, es la búsqueda de la política π_*, la cual en general, no puede ser encontrada si no se tiene la información suficiente acerca de F, podemos encontrar la política $\tilde{\pi}_*$ (ya que se conoce la información suficiente acerca de $\tilde{\mathrm{F}}$) e intentar usarla, en lugar de π_*, para controlar al proceso original (SR). En otras palabras, usar la política $\tilde{\pi}_*$ como una aproximación accesible, de la política π_*.

Siguiendo el enfoque de [5, 7], estimaremos la pérdida adicional (debido al uso de $\tilde{\pi}_*$ en vez de π_*) mediante el siguiente **índice de estabilidad descontado**:

$$\triangle\beta(\mathrm{x}) := Q\beta(\mathrm{x}, \tilde{\pi}_*) - Q\beta(\mathrm{x}, \pi_*) \geq 0 \tag{0.3}$$

O bien, ya que a veces es más importante estimar la pérdida relativa, es mejor usar el siguiente **índice de estabilidad relativo**:

$$\overline{\triangle}\beta(\mathrm{x}) := \frac{\triangle\beta(\mathrm{x})}{|Q\beta(\mathrm{x},\pi_*)|} \tag{0.4}$$

(admitiendo el valor $+\infty$ si $Q(\mathrm{x},\pi_*) = 0$)

El problema de estabilidad de los sistemas de control (SR) y $(\widetilde{\mathrm{SA}})$, en (0.1) y (0.2), puede plantearse, como el problema de obtener condiciones para garantizar que $\triangle\beta(\mathrm{x}) \to 0$ ó $\overline{\triangle}\beta(\mathrm{x}) \to 0$, cuando, $\mu(\mathrm{F},\tilde{\mathrm{F}}) \to 0$, donde μ es una métrica en el espacio de funciones de distribución. De manera más apropiada, desde el punto de vista de las aplicaciones, consiste en obtener cotas de la tasa de convergencia $\triangle\beta(\mathrm{x}) \to 0$, es decir, buscar una desigualdad del tipo:

$$\triangle\beta(\mathrm{x}) \leq g_{\mathrm{x}}(\mu(\mathrm{F}, \tilde{\mathrm{F}})) \tag{0.5}$$

donde $g_{\mathrm{x}}(\mathrm{s}) \to 0$ cuando $\mathrm{s} \to 0$.

La búsqueda de las cotas superiores de (0.5), o bien de:

$$\overline{\triangle}\beta(\mathrm{x}) \leq g_{\mathrm{x}}^{(1)}(\mu(\mathrm{F}, \tilde{\mathrm{F}})) \tag{0.6}$$

en el caso cuando $\mu = \sigma$, donde σ es la métrica de variación total, es el primer objetivo de esta tesis. El segundo objetivo, es el problema de estudiar el comportamiento asíntotico del lado derecho de (0.6), cuando el factor de descuento β tiende a 1. Lo que se quiere tener, son cotas uniformes sobre $\beta \in (0,1)$, en (0.6), es decir:

$$\sup_{\beta\in(0,1)}\overline{\triangle}\beta(\mathrm{x}) \leq g_{\mathrm{x}}^{(2)}(\mu(\mathrm{F},\tilde{\mathrm{F}})) \tag{0.7}$$

Obtener (0.7) es un problema más difícil, que proponer algunas condiciones para obtener (0.5). En efecto, en palabras generales, la estabilidad del proceso (0.1) cuando $\beta \to 1$, está relacionada con la estabilidad, con respecto al costo promedio. En su momento, para probar esta última estabilidad, necesitaremos usar algunas condiciones de ergodicidad de los procesos inducidos por (0.1) y (0.2), al aplicar las políticas de control estacionarias (véase [2, 5, 6, 7]).

Obsérvese, que existen ejemplos de sistemas del tipo (0.1) y (0.2), **no estables** en costo descontado, es decir, ejemplos para los cuales $\sup_{\beta\in(0,1)}\overline{\triangle}\beta(\mathrm{x}) = \infty$, para algunas distribuciones F y $\tilde{\mathrm{F}}$, tales que, $\mu(\mathrm{F},\tilde{\mathrm{F}}) \to 0$, donde μ es la métrica de Prokhorov (véase [10]).

Por otro lado, hasta donde conocemos, los únicos trabajos donde se obtuvieron las cotas para el índice de estabilidad descontado, como en (0.5), son: [1, 2, 9, 10, 12]. En estos artículos, las desigualdades de estabilidad se demuestran por métodos distintos de los que se utilizan en el presente trabajo (en particular por la técnica de operadores de contracción), pero bajo algunas condiciones generales (sin hipótesis de ergodicidad). Estas desigualdades tienen la siguiente estructura:

$$\triangle\beta(\mathrm{x}) \leq \mathrm{k}(\beta) g_{\mathrm{x}}^{(3)}(d(\mathrm{SR},\widetilde{\mathrm{SA}})) \tag{0.8}$$

donde d es una medida de la diferencia entre los sistemas: original y aproximante (en [9, 10], d es una métrica ponderada). La función $g_{\mathrm{x}}^{(3)}$ no depende del parámetro β, mientras que:

$$\mathrm{k}(\beta) \sim \frac{\mathrm{k}_0}{(1-\beta)^2} \quad \text{cuando} \quad \beta \to 1 \tag{0.9}$$

Se conoce que bajo las condiciones generales (véase [15]), el valor óptimo de costo descontado $Q\beta(\mathrm{x},\pi_*)$, en (0.4), es del orden:

$$\frac{\text{cte.}}{1-\beta} \qquad \text{si} \qquad \beta \to 1. \tag{0.10}$$

Teniendo en mente obtener desigualdades como en (0.6) y utilizando (0.8) y (0.10), podemos escribir

$$\overline{\triangle}\beta(\mathrm{x}) \leq \mathrm{k}_1(\beta) g_{\mathrm{x}}^{(3)}(d(\mathrm{SR}, \widetilde{\mathrm{SA}})) \tag{0.11}$$

Ahora, de (0.9) y (0.10) se tiene:

$$\mathrm{k}_1(\beta) \sim \frac{\text{cte.}}{(1-\beta)} \qquad \text{si} \quad \beta \to 1.$$

Por lo tanto, desigualdades como en (0.8), impiden obtener la desigualdad (0.7) y, además, la cota en (0.11) no es de utilidad cuando β está cerca de 1.

Para obtener la desigualdad (0.7) elegiremos un método de análisis alternativo, imponiendo condiciones suficientemente restrictivas acerca de los sistemas (0.1) y

(0.2), reduciendo por esto, el problema de estabilidad al estudio de la estabilidad de procesos regenerativos, con "buenas" propiedades del tiempo de regeneración. Para procesos regenerativos, los métodos de estimación han sido desarrollados en la última década (véase [17, 18, 19]). Tal enfoque reduce el área de aplicaciones posibles, pero ayudará a obtener la siguiente cota uniforme para $\overline{\triangle}\beta(\mathrm{x})$ (obtenida en el presente trabajo):

$$\sup_{\beta\in(0,1)}\overline{\triangle}\beta(\mathrm{x}) \leq \tilde{\mathrm{k}}\sigma(\mathrm{F},\tilde{\mathrm{F}})\left[1-\ln[\alpha\sigma(\mathrm{F},\tilde{\mathrm{F}})]\right] \tag{0.12}$$

donde σ es la métrica de variación total y las constantes $\tilde{\mathrm{k}}$ y α pueden calcularse mediante las cantidades dadas en las suposiciones del capítulo 1 y 3.

La desigualdad (0.12) se cumple bajo condiciones de ergodicidad del tipo de Liapunov, que se escribirán en el capítulo 3, y que, en efecto, garantizan la convergencia geométrica de la distribución de (0.1) a su distribución estacionaria correspondiente, cuando se aplican políticas estacionarias. De hecho, en problemas de control aplicado, la clase de procesos que demuestran la convergencia geométrica es suficientemente amplia, por ejemplo: procesos en control de inventarios, colas controlables, etc. (véase por ejemplo[16]).

El material de este trabajo está organizado de la siguiente manera:

Capítulo 1. Se dará la descripción: de los modelos de control Markoviano, de las políticas de control y del criterio de optimización con el cual se trabaja. También en este capítulo, se presenta el índice de costo total descontado, definimos las políticas óptimas y además las condiciones generales acerca de los modelos, bajo las cuales se verifica la existencia de las políticas óptimas estacionarias.

Capítulo 2. Se plantearán los problemas a estudiar y los objetivos de este trabajo.

Capítulo 3. Introduciremos las suposiciones del tipo de Liapunov para el tiempo de regreso a un estado fijo.

Capítulo 4. Daremos los resultados obtenidos en este trabajo, es decir, las cotas superiores para los índices $\triangle\beta(\mathrm{x})$ y $\overline{\triangle}\beta(\mathrm{x})$.

Capítulo 5. Se introducen resultados auxiliares y se demuestran las afirmaciones del capítulo 4.

Capítulo 6. Se considera un ejemplo de control de inventarios, donde se satisfacen todas las suposiciones dadas en el capítulo 3 y , por consiguiente, se aplicaran las desigualdades de estabilidad, obtenidas en el capítulo 4.

Capítulo 7. Se da la conclusión de este trabajo y se discuten, brevemente, algunos problemas abiertos relacionados con este trabajo.

CAPITULO 1

PROCESOS DE CONTROL MARKOVIANO CON COSTO DESCONTADO

1.1 ELEMENTOS DEL MODELO DE CONTROL MARKOVIANO

Recuérdese que **un espacio de Borel** es un subconjunto medible de un espacio métrico, separable y completo. Si Y es un espacio de Borel, entonces, denotaremos por B(Y) a su $\sigma-$ álgebra de Borel.

Un modelo controlable de Markov general (véase [3, 13, 16]) se define, de una manera estándar, por los siguientes objetos: $\{X, A, \{A(x), x \in X\}, p, c\}$, donde:

X, llamado **espacio de estados**, es un espacio de Borel. Los elementos x ∈ X son llamados **estados.**

A, el **conjunto de acciones o controles**, es un espacio de Borel.

{A(x), x ∈ X} es la familia de **conjuntos de controles admisibles**. Es decir, a cada estado x ∈ X se le asocia un subconjunto A(x) ⊂ A no vacío y medible, cuyos elementos son acciones admisibles cuando el proceso, asociado con el modelo, está en el estado x ∈ X.

El conjunto

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

es supuesto un subconjunto medible del espacio producto XA. Los elementos $(x,a) \in \mathbb{K}$ son denotados por k.

p ó p(B/k), donde B ∈ B(X), k ∈ $\mathbb{K}$ es la **ley de transición**, la cual es un kérnel estocástico de X dado $\mathbb{K}$, es decir, p(./k) es una medida de probabilidad sobre el espacio (X, B(X)), para cada k ∈ $\mathbb{K}$, y p(B/.) es una función medible sobre $\mathbb{K}$, para cada B ∈ B(X).

c: $\mathbb{K}\to$R es una función medible, llamada **función de costo en un paso**, la cual para nuestros propósitos la supondremos acotada.

En el caso particular, pero suficientemente general, con el cual vamos a trabajar, la probabilidad **de transición** esta dada por:

$$p(B/k) = P(x_{t+1} \in B/x_t = x, a_t = a), \qquad B \in B(X),\ (x,a) \in \mathbb{K}.$$

donde, x_{t+1} está dado como en (0.1) por:

$$x_{t+1} = \Phi(x_t, a_t, \xi_t), \qquad t = 0, 1, ...; \quad x_0 = x. \tag{1.1}$$

con Φ: IKR^n $\rightarrow$X , una función medible dada, $x_t \in X$, $a_t \in A(x_t)$ son estados y controles en la t-ésima etapa de la evolución del proceso controlable, que se asocia con el modelo descrito de Markov y $\xi_0, \xi_1, ...$son vectores aleatorios, independientes e idénticamente distribuidos (v.a's.i.i.d), con valores en algún espacio Euclidiano R^n. Es fácil ver en este caso que:

$$p(B/k) = \int_{R^n} 1_B \,[\Phi(k,s)]\, F(ds) \tag{1.2}$$

donde F es la distribución común de las v.a's. $\xi_0, \xi_1, ...$

La dinámica del proceso de control (1.1) se describe de la siguiente manera: si en el tiempo t se observa al proceso en el estado $x \in X$, es decir, $x_t = x \in X$, se selecciona la acción $a_t = a \in A(x_t) \subset A$ y la variable aleatoria (v.a.) ξ_t toma el valor s (es decir, ξ_t = s), entonces, se genera un costo c(x,a) y el proceso, en el tiempo t+1, pasa al siguiente estado x_{t+1}= $\Phi(x, a, s)$ y así sucesivamente trazando una **trayectoria** $x_0, a_0, x_1, a_1, ...$ si se proporciona una regla para seleccionar las acciones $a_0, a_1, a_2, ...$. Tal regla se llama una **política de control**, definida en la siguiente sección.

1.2 CLASES DE POLITICAS DE CONTROL

Para cada $t \geq 0$, definimos al espacio $\mathbb{H}_t$ de **historias** hasta el tiempo t como sigue:

$$\mathbb{H}_0 = X, \qquad \mathbb{H}_t = \mathbb{K}^t X, \qquad t = 0,1,....$$

Un elemento genérico $h_t \in \mathbb{H}_t$ es la parte de la trayectoria del proceso hasta el tiempo t y éste representa un vector de la forma:

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

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

DEFINICION 1.1 (véase [3,16]) Una **política de control** es una sucesión $\pi = \{\pi_t, t = 0, 1, ...\}$, tal que, para cada t, π_t es un kérnel en A dado $\mathbb{H}_t$, que satisface la siguiente propiedad: $\pi_t(A(x_t)/h_t) = 1$, para cada $h_t \in \mathbb{H}_t$.

En otras palabras, la política π es una regla para elegir un control en cada momento t, tal que si el proceso está en el estado x_t y se realiza la historia $h_t = (x_0, a_0, ..., x_{t-1}, a_{t-1}, x_t)$, entonces, según π, la acción escogida es $a_t \in D \subset A(x_t)$ con probabilidad $\pi_t(D/h_t)$.

Una política $\pi = \{\pi_t\}$ se llama **determinística** si las distribuciones π_t están concentradas en subconjuntos que consisten de un único punto, es decir, existen funciones medibles f_t, t = 0,1,..., tales que, para cuando se da la historia h_t, la regla π_t obliga a elegir el control $a_t = f_t(h_t) \in A(x_t)$.

DEFINICION 1.2 (véase [3, 16]) Diremos que una política es **estacionaria** (y determinística) si dichas funciones f_t, t = 0,1,.... no dependen de t, es decir, $f_t = f$, t = 0,1,2,.... y la dependencia entre f y h_t se tiene sólo por x_t, el último elemento de la historia, es decir, $f(h_t) \equiv f(x_0, a_0, ..., x_{t-1}, a_{t-1}, x_t) = f(x_t)$ para toda historia con el mismo último estado x_t.

Así, hemos visto que una política **estacionaria** $\pi = \{f, f, ...\}$ está dada por la función medible:

$$f : X \to A, \qquad \text{con} \qquad f(x) \in A(x), \quad x \in X. \tag{1.3}$$

De modo que en cada momento t, se elige el control $a_t = f(x_t)$ independientemente de la "prehistoria" $x_0, a_0, ..., x_{t-1}, a_{t-1}$. Por conveniencia, será útil denotar por S, como al conjunto de funciones medibles f : X→A, las cuales satisfacen (1.3), por consiguiente de aquí en adelante identificaremos a f como una política estacionaria $\pi = \{f, f, ...\}$.

La clase de todas las políticas de control que satisfagan la definición 1.1 las denotaremos por Π.

OBSERVACION 1.3. Aplicando una política estacionaria $f \in S$ al proceso (1.1), éste se puede reescribir de la siguiente manera :

$$x_{t+1} = \Phi(x_t, f(x_t), \xi_t), \qquad t = 0, 1, ...$$

Con esto se sigue, que $\{x_t\}$ es un proceso de Markov con espacio de estados X y probabilidad de transición:

$$
\begin{aligned}
p(B/x) &= p(B/x, f(x)) = P(x_{t+1} \in B/x_t = x) = P(\Phi(x, f(x), \xi_t) \in B) \\
&= \int_{R^n} 1_B\,[\Phi(x,f(x),s)]\,F(ds) \qquad (1.4)
\end{aligned}
$$

1.3 INDICE DE COSTO TOTAL DESCONTADO

El objetivo del control óptimo es escoger una política de control que sea óptima para el índice de funcionamiento, para el sistema de control dado por (1.1). Así, necesitaremos introducir un índice como criterio de optimización. En esta tesis se trabajará con el índice de costo total descontado que en particular, es apropiado para problemas de control con contexto económico (véase [3]).

Primero notemos, que para cada estado inicial $x_0 = x \in X$ del proceso (1.1) y para cada política de control $\pi \in \Pi$ existe una medida de probabilidad $P_x\pi$, con el espacio de las trayectorias $\{x_0, a_0, ..., x_t, a_t, ...\}$ (véase [3]). La esperanza con respecto a esta medida, la denotaremos por $E_x\pi$.

Sea $\beta \in (0,1)$ un número, llamado factor de descuento. Para cualquier política $\pi \in \Pi$, dado el estado inicial $x \in X$, se define **el costo total decontado** $Q\beta$ por (véase [3, 13]):

$$Q\beta(x,\pi) := \sum_{t=0}^{\infty} \beta^t E_x^\pi c(x_t, a_t) \tag{1.5}$$

donde c es la función de costo en un paso y $\{(x_t, a_t),\ t = 0,1,2,...\}$ es la trayectoria del proceso (1.1), inducida por la aplicación de la política π y el estado inicial x. Como $\beta < 1$ y c es acotada, es fácil ver, que la serie en (1.5) converge, así, el índice $Q\beta(x,\pi)$ está bién definido. Por comodidad, llamaremos a $Q\beta(x,\pi)$ simplemente costo descontado.

DEFINICION 1.4. Una política $\pi_* \in \Pi$ es **óptima en costo descontado** (o simplemente óptima) si para todo estado inicial $x \in X$ se cumple lo siguiente:

$$Q\beta(x,\pi_*) = Q\beta^*(x) := \inf_{\pi\in\Pi} Q\beta(x,\pi) \tag{1.6}$$

La función $Q\beta^*(x)$, $x \in X$, representa el **costo óptimo** ($\beta-$descontado).

1.4 EXISTENCIA DE POLITICAS OPTIMAS ESTACIONARIAS

Para definir el índice de estabilidad descontado, como en (0.3), necesitaremos asegurarnos de que existen políticas óptimas para los procesos (0.1) y (0.2). Además, nuestro método de análisis de estabilidad requiere de la existencia de políticas óptimas **estacionarias**. Así para garantizar la existencia de estas últimas políticas, introduciremos algunas suposiciones de continuidad de los procesos (0.1) y (0.2) (o bien, del proceso (1.1)).

La ley de transición del proceso (0.1), que es la misma para (1.1), está dada en (1.2). De igual manera, la ley de transición, denotada por $\tilde{p}$, para el sistema de control (0.2), esta dadá por:

$$\tilde{p}(B/k) = \int_{R^n} 1_B\, [\Phi(k,s)]\, \tilde{F}(ds)$$

donde,

$B \in B(X)$,

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

$\tilde{F}$ es la disribución común de los v.a's.i.i.d. $\tilde{\xi}_t$ en (0.2).

SUPOSICION S.1.5.

(a) Para cada $x \in X$, el conjunto $A(x)$ es **compacto.**

(b) Existe una constante $r > 0$, tal que, $|c(x,a)| \leq r$, para todo par $(x,a) \in \mathbb{K}$; además, para cada $x \in X$ la función $c(x,.)$: $A(x) \rightarrow R$ es continua.

(c) Para cada $x \in X$ y cada función u: $X \rightarrow R$ medible y acotada, la aplicación

$\int_X u(y)p(\, dy/x,.)$: $A(x) \rightarrow R$ es continua.

(d) La condición (c) se satisface para $\tilde{p}$.

OBSERVACION 1.6. Pueden darse algunas condiciones sencillas en términos de Φ, F, $\tilde{F}$, en (0.1) y (0.2), suficientes para que se cumpla S.1.5,(c)-(d), (véase [11, 16] o el ejemplo en el capítulo 6 de este trabajo).

Una política óptima para "el proceso original" (0.1) está especificada en la definición (1.6). De la misma manera, decimos que $\tilde{\pi}_*$ es una política óptima para el proceso de control "aproximante" (0.2), si $\tilde{\pi}_*$ minimiza el costo descontado para el proceso (0.2), definido como sigue:

$$\tilde{Q}\beta(x, \tilde{\pi}_*) = \tilde{Q}\beta^*(x) := \inf_{\pi\in\Pi}\tilde{Q}\beta(x, \pi) \qquad x \in X; \tag{1.7}$$

donde, análogamente a (1.5), se define el índice de costo descontado para este proceso como sigue:

$$\tilde{Q}\beta(x, \pi) := \sum_{t=0}^{\infty} \beta^t E_x \pi c(\tilde{x}_t, a_t) \tag{1.8}$$

donde, $\{(\tilde{x}_t, a_t),\ t = 0,1,2,...\}$ es la trayectoria del proceso (0.2), inducida por la aplicación de la política π y el estado inicial x. De nuevo, la función $\tilde{Q}\beta(x,\tilde{\pi}_*)$, $x \in X$, es llamada la función de costo óptimo para el proceso (0.2).

PROPOSICION 1.7. (véase [13, Cap.2]) Supóngase que se satisfacen las condiciones de S.1.5. Entonces, para $\beta \in (0, 1)$, existen políticas óptimas estacionarias $f_*, \tilde{f}_* \in S$, bajo el índice de costo descontado, respectivamente, para los procesos (0.1) y (0.2).

CAPITULO 2

PLANTEAMIENTO DEL PROBLEMA DE ESTABILIDAD EN COSTO DESCONTADO

En esta sección se siguen los razonamientos de la introducción y nos permitiremos repetir algunas definiciones y conceptos de la introducción, usando terminologías (dadas en el capítulo anterior) más adecuadas para este trabajo .

Sean $\{\xi_t\}$ y $\left\{\widetilde{\xi}_t\right\}$ dos sucesiones de v.a's.i.i.d. (en R^n) con funciones de distribución comúnes, denotadas, respectivamente, por F y $\widetilde{\mathrm{F}}$. Se considerarán dos procesos de control Markoviano dados (como en (0.1) y (0.2)) por las siguientes relaciones de recurrencia:

$$\mathrm{x}_{t+1} = \Phi(\mathrm{x}_t, a_t, \xi_t), \qquad t = 0, 1, ...; \quad \mathrm{x}_0 = \mathrm{x} \tag{2.1}$$

$$\widetilde{\mathrm{x}}_{t+1} = \Phi(\widetilde{\mathrm{x}}_t, \widetilde{a}_t, \widetilde{\xi}_t), \qquad t = 0, 1, ...; \quad \widetilde{\mathrm{x}}_0 = \mathrm{x} \tag{2.2}$$

donde,

$x_t, \tilde{x}_t \in X$,

$a_t \in A(x_t) \subset A, \quad t = 0,1,.....$

$\tilde{a}_t \in A(\tilde{x}_t) \subset A, \quad t = 0,1,.....$

y los conjuntos X, A, $\{A(x), x \in X\}$ fueron definidos en el capítulo anterior.

Sea $x \in X$, un estado arbitrario fijo, utilizando a éste, como estado inicial para ambos procesos $\{x_t\}$ y $\{\tilde{x}_t\}$, es decir, $x_0 = x$, $\tilde{x}_0 = x$. Como se dijo anteriormente la función de costo en un paso $c(x,a)$ es la misma para ambos procesos (2.1) y (2.2). Supondremos también que las leyes de transición p y $\tilde{p}$, definidas en (1.2) y (1.4), para los procesos (2.1) y (2.2), satisfacen la suposición S.1.5, por consiguiente, de la proposición 1.7, existen políticas estacionarias f_*, $\tilde{f}_* \in S$ óptimas para los procesos (2.1) y (2.2), respectivamente, bajo el índice de costo descontado, dado en (1.5) y (1.8). Por lo tanto, podemos reescribir los índices de estabilidad, introducidos en (0.3) y (0.4) como sigue:

Para el índice de Estabilidad Descontado:

$$\triangle\beta(x) := Q\beta(x, \tilde{f}_*) - Q\beta(x, f_*) \tag{2.3}$$

Para el índice de estabilidad relativo:

$$\overline{\triangle}\beta(x) := \frac{Q\beta(x, \tilde{f}_*) - Q\beta(x, f_*)}{|Q\beta(x, f_*)|} \tag{2.4}$$

Observemos que en (2.4), $Q\beta(x,f_*) = Q\beta^*(x)$ es el costo descontado óptimo para el sistema de control (2.1), con estado inicial $x \in X$.

Ahora, para plantear nuestros objetivos necesitaremos de la siguiente (véase, por ejemplo, [23]):

DEFINICION 2.1. Sean ξ, η, dos elementos aleatorios con valores en un espacio de Borel Y, con distribuciones $F\xi$, $F\eta$, respectivamente. Entonces, la métrica de variación total σ está dada por:

$$\sigma(\xi,\eta) \equiv \sigma(F\xi, F\eta) := 2 \sup_{B \in B(Y)} |P(\xi \in B) - P(\eta \in B)| \tag{2.5}$$

El objetivo de este trabajo, es proporcionar algunas condiciones acerca de la clase de sistemas de control (o procesos de control) considerados, tales que, estas condiciones nos permitan obtener las siguientes desigualdades de estabilidad (o en otras palabras, las cotas superiores de estabilidad):

$$\begin{aligned} \triangle\beta(x) &\leq g(\sigma(F,\tilde{F})) \\ \sup_{\beta\in(0,1)}\overline{\triangle}\beta(x) &\leq g_2(\sigma(F\xi, F\eta)) \end{aligned}$$

donde, las funciones explícitas $g, g_2 : [0,\infty) \to [0,\infty)$ son tales que, si s$\to$ 0, entonces, las funciones $g(s)$ y $g_2(s)$ se aproximan a cero, suficientemente rápido. En efecto obtendremos las desigualdades de estabilidad con las funciones que poseen la siguiente asíntota:

$$(cte)(s) \ln(\frac{1}{s}), \quad \text{cuando} \quad s \to 0.$$

CAPITULO 3

SUPOSICIONES EN LOS SISTEMAS DE CONTROL.

Cuando se aplica cualquier política estacionaria $f \in S$ a los procesos (2.1) y (2.2), obtendremos dos procesos de Markov (véase la observación 1.3) dados por las siguientes relaciónes de recurrencia:

$$
\begin{aligned}
x_{t+1} &= \Phi(x_t, f(x_t), \xi_t), \quad t = 0, 1, ...; \quad x_0 = x && (3.1)\\
\tilde{x}_{t+1} &= \Phi(\tilde{x}_t, f(\tilde{x}_t), \tilde{\xi}_t), \quad t = 0, 1, ...; \quad \tilde{x}_0 = x && (3.2)
\end{aligned}
$$

Las probabilidades de trancisión p y $\tilde{p}$, respectivamente, de (3.1) y (3.2), están dadas por (1.4) y por la siguiente relación:

$$
\begin{aligned}
\tilde{p}(B/x) &= \tilde{p}(B/x, f(x)) = P(\tilde{x}_{t+1} \in B/\tilde{x}_t = x) = P(\Phi(x, f(x), \tilde{\xi}_t) \in B)\\
&= \int_{R^n} 1_B(\Phi(x, f(x), s))\tilde{F}(ds), \qquad B \in B(X), \quad x \in X \qquad (3.3)
\end{aligned}
$$

Para hacer notar la dependencia de los procesos (3.1) y (3.2) con respecto a la política $f \in$ S, vamos a denotar estos procesos por $\left\{x_t^{(f)}\right\}$ y $\left\{\tilde{x}_t^{(f)}\right\}$, respectivamente, y sus probabilidades de transición en (1.4) y (3.3) las escribiremos usando las notaciones:

$$p(B/x, f(x)) \qquad y \qquad \tilde{p}(B/x, f(x)), \qquad B \in B(X), \qquad x \in X$$

Ahora, introduciremos las siguientes condiciones, bajo las cuales se trabaja, para los modelos (3.1) y (3.2).

SUPOSICION S.3.1. Existe un estado $x_* \in X$ y un número positivo δ, tales que, se cumple lo siguiente:

(a) $\inf_{f \in S} p(\{x_*\}/x_*, f(x_*)) \geq \delta$.

(b) $\inf_{f \in S} \tilde{p}(\{x_*\}/x_*, f(x_*)) \geq \delta$.

SUPOSICION S.3.2 Existe una función medible V: X→ $[1, \infty)$ y una constante $\mu > 0$, tales que, para cada política estacionaria $f \in$ S, se cumple para el proceso (3.1) lo siguiente:

(a) $\int_X V(z)p(z/x,f(x)) \leq e^{-\mu}V(x)$. para toda $x \neq x_*$. además

(b) $d := \sup_{f \in S} \int_X V(z)p(dz/x_*,f(x_*)) < \infty$.

SUPOSICION S.3.3 Lo mismo que en la suposición S.3.2, con alguna función $\tilde{V}$ y constantes $\tilde{\mu}$, $\tilde{d}$, se cumple para el proceso (3.2).

OBSERVACION 3.4 Hay sistemas de control aplicados (por ejemplo, algunos modelos de inventarios) donde el requisito de que se cumplan S.3.1 - S.3.3, para **todas** las políticas estacionarias, es demasiado restrictivo. De hecho, como se demostró en

[18], las suposiciones S.3.1-S.3.3 son suficientes para la convergencia geométrica, con la métrica σ, de las distribuciones de $\left\{x_t^{(f)}\right\}$ y $\left\{\tilde{x}_t^{(f)}\right\}$ a las correspondientes distribuciones invariantes. Esta convergencia puede no darse en algunos sistemas, cuando se aplican políticas estacionarias "no estables" (es decir, que puede suceder que los estados $x_t^{(f)} \to \infty$, cuando, $t \to \infty$), sin embargo, para los sistemas mencionados, hay como regla, una subclase $S_0 \subset S$ de políticas estacionarias, tales que, para cada $f \in S_0$, los procesos $\left\{x_t^{(f)}\right\}$ y $\left\{\tilde{x}_t^{(f)}\right\}$ tienen "buenas" propiedades de ergodicidad. Por otro lado, de las demostraciones del capítulo 5, se puede ver, que podriamos relajar, esencialmente, nuestras condiciones, suponiendo que existe una subclase $S_0 \subset S$ de políticas estacionarias, tales que, las políticas óptimas f_*, $\tilde{f}_*$ estén en S_0, y las suposiciones S.3.1, S.3.2 y S.3.3 se cumplen, para cada $f \in S_0$.

En el capítulo 6 discutiremos un ejemplo de control de inventario, donde se satisfacen todas las condiciones S.3.1-S.3.3 y también se satisface la condición S.1.5.

CAPITULO 4

RESULTADOS PRINCIPALES.

Los resultados de esta tesis se resumen en los dos teoremas siguientes:

TEOREMA 4.1 Supongamos que se satisfacen las suposiciones S.1.5, S.3.1 - S.3.3 y que $x_0 = \tilde{x}_0 = x_*$, entonces, se cumple que:

$$\triangle\beta(x_*) \leq \frac{k_1}{1-\beta}\sigma(F,\tilde{F})\left[1 - \ln[\alpha\sigma(F,\tilde{F})]\right] \tag{4.1}$$

donde las constantes k_1, $\alpha > 0$, se pueden estimar, en los términos de las cantidades r, $V(x_*)$, μ, d, $\tilde{V}(x_*)$, $\tilde{\mu}$, $\tilde{d}$, involucradas en las suposiciones del capítulo 1 y del capítulo 3.

TEOREMA 4.2 Supongamos que se satisfacen las suposiciones S.1.5, S.3.1, S.3.2, S.3.3 y, además, $\inf_{(x,a)\in \mathbb{K}} c(x,a) = \kappa > 0$, $x_0 = \tilde{x}_0 = x_*$. Entonces se cumple que:

$$\sup_{\beta\in(0,1)} \overline{\triangle}\beta(x_*) \leq \frac{k_1}{\kappa}\sigma(F,\tilde{F})\left[1 - \ln[\alpha\sigma(F,\tilde{F})]\right] \tag{4.2}$$

Recordemos que la función de distribución $\widetilde{\mathrm{F}}$, de las v.a's $\widetilde{\xi}_t$ en (2.2), se utiliza como la aproximación de la función de distribución desconocida F, de las v.a's ξ_t en (2.1). También se utiliza la métrica de variación total σ para estimar esta aproximación y para el caso de que se tenga que $\sigma(\mathrm{F},\widetilde{\mathrm{F}}) \to 0$, entonces, los lados derechos de (4.1) y (4.2) se anulan con rapidez "casi lineal". En particular, esto significa, que bajo las condiciones dadas anteriormente, el sistema de control (2.1) es estable. No es difícil de checar (por ejemplos) que, en general, no existen cotas de estabilidad del tipo (4.1):

$$\Delta\beta(\mathrm{x}) \leq \mathrm{cte.}g(\sigma(\mathrm{F},\widetilde{\mathrm{F}}))$$

con la función $g(\mathrm{s})$ que se anula cuando $\mathrm{s} \to 0$, más rápido que una función lineal.

CAPITULO 5

RESULTADOS AUXILIARES Y DEMOSTRACIONES

5.1 RESULTADOS PRELIMINARES

Primero reducimos el problema de estimación de $\triangle\beta(x)$ al problema de la búsqueda de cotas uniformes con respecto de $f \in S$, para la diferencia $\left|Q\beta(x,f) - \tilde{Q}\beta(x,f)\right|$, de los costos descontados para los procesos (2.1) y (2.2).

LEMA 5.1 Bajo la suposición S.1.5 se cumple que:

$$\triangle\beta(x) \leq 2\sup_{f\in S}\left|Q\beta(x,f) - \tilde{Q}\beta(x,f)\right| \tag{5.1}$$

donde $Q\beta$ y $\tilde{Q}\beta$ están definidos, respectivamente, en (1.5) y (1.8).

DEMOSTRACION. De la definición de $\triangle\beta(x)$ en (2.3), sumando y restando el término $\tilde{Q}\beta(x,\tilde{f}_*)$, se tiene lo siguiente (véase la definición de las políticas óptimas en (1.6) y (1.7)):

$$
\begin{aligned}
\triangle\beta(\mathrm{x}) &= \mathrm{Q}\beta(\mathrm{x},\tilde{f}_*) - \tilde{\mathrm{Q}}\beta(\mathrm{x},\tilde{f}_*) + \tilde{\mathrm{Q}}\beta(\mathrm{x},\tilde{f}_*) - \mathrm{Q}\beta(\mathrm{x},f_*) \\
&= \left[\inf_{f\in \mathrm{S}} \tilde{\mathrm{Q}}\beta(\mathrm{x},f) - \inf_{f\in \mathrm{S}} \mathrm{Q}\beta(\mathrm{x},f)\right] + \left[\mathrm{Q}\beta(\mathrm{x},\tilde{f}_*) - \tilde{\mathrm{Q}}\beta(\mathrm{x},\tilde{f}_*)\right] \\
&\leq \sup_{f\in \mathrm{S}} \left|\tilde{\mathrm{Q}}\beta(\mathrm{x},f) - \mathrm{Q}\beta(\mathrm{x},f)\right| + \left[\mathrm{Q}\beta(\mathrm{x},\tilde{f}_*) - \tilde{\mathrm{Q}}\beta(\mathrm{x},\tilde{f}_*)\right] \\
&\leq 2\sup_{f\in \mathrm{S}} \left|\tilde{\mathrm{Q}}\beta(\mathrm{x},f) - \mathrm{Q}\beta(\mathrm{x},f)\right| \qquad (5.2)
\end{aligned}
$$

ya que

$\left|\inf_{f\in \mathrm{S}} \widetilde{H}(f) - \inf_{f\in \mathrm{S}} H(f)\right| \leq \sup_{f\in \mathrm{S}} \left|\widetilde{H}(f) - H(f)\right|$

para cualesquiera dos funciones finitas H y $\widetilde{H}$ en S. Así, (5.2) es la desigualdad que se quería. ■

LEMA 5.2 Bajo la suposición S.1.5 se cumple que:

$$
\triangle\beta(\mathrm{x}) \leq \frac{2\mathrm{r}}{1-\beta} \sup_{f\in \mathrm{S}} \sup_{t\geq 0} \sigma(\mathrm{x}_t^{(f)}, \tilde{\mathrm{x}}_t^{(f)}) \qquad (5.3)
$$

donde r es la constante involucrada en S.1.5, (b) y los procesos $\left\{\mathrm{x}_t^{(f)}\right\}$ y $\left\{\tilde{\mathrm{x}}_t^{(f)}\right\}$ fueron definidos en el capítulo 3.

DEMOSTRACION. Fijemos cualquier política $f \in \mathrm{S}$, por (1.5) y (1.8) se tiene:

$$
\left|\tilde{\mathrm{Q}}\beta(\mathrm{x},f) - \mathrm{Q}\beta(\mathrm{x},f)\right| = \left|\sum_{t=0}^{\infty} \beta^t \mathrm{E}_{\mathrm{x}}^f \mathrm{c}(\mathrm{x}_t, f(\mathrm{x}_t)) - \sum_{t=0}^{\infty} \beta^t \mathrm{E}_{\mathrm{x}}^f \mathrm{c}(\tilde{\mathrm{x}}_t, f(\tilde{\mathrm{x}}_t))\right|
$$

$$\left|\tilde{Q}\beta(x,f) - Q\beta(x,f)\right| \leq r\sum_{t=0}^{\infty}\beta^t \left|\frac{E_x^f c(x_t, f(x_t))}{r} - \frac{E_x^f c(\tilde{x}_t, f(\tilde{x}_t))}{r}\right|$$

$$\leq r\sum_{t=0}^{\infty}\beta^t \sup_{\rho:||\rho||_\infty \leq 1}\left|E_x^f \rho(x_t) - E_x^f \rho(\tilde{x}_t)\right| \tag{5.4}$$

ésto es consecuencia de: S.1.5,(b) y de la función: $\dfrac{c(x_t, f(x_t))}{r}$ que es acotada por 1.

En (5.4) se denota $||\rho||_\infty := \sup_{x\in X}|\rho(x)|$. Luego utilizando la siguiente definición, de la métrica de variación total:

$$\sigma(\xi,\eta) := \sup_{\rho:||\rho||_\infty \leq 1}|E\rho(\xi) - E\rho(\eta)| \tag{5.5}$$

que es equivalente a la definición dada en (2.5) (véase [23]) y de (5.4) se sigue que:

$$\left|\tilde{Q}\beta(x,f) - Q\beta(x,f)\right| \leq r\sum_{t=0}^{\infty}\beta^t\sigma(x_t,\tilde{x}_t) \leq \frac{r}{1-\beta}\sup_{t\geq 0}\sigma(x_t,\tilde{x}_t)$$

y de (5.1) se sigue la desigualdad (5.3). ■

5.2 REDUCCION DEL PROBLEMA A COMPARACION DE PROCESOS REGENERATIVOS

En virtud del lema 5.2, para demostrar el teorema 4.1 es suficiente encontrar una cota superior adecuada para $\sup_{f \in S} \sup_{t \geq 0} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)})$. En otras palabras, hay que buscar una cota superior para $\sup_{t \geq 0} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)})$ que no dependa de cualquier política $f \in S$. La idea general de como vamos a hacerlo, es usando S.3.1-S.3.3 para verificar que para la política estacionaria f, los procesos $\left\{x_t^{(f)}\right\}$ y $\left\{\tilde{x}_t^{(f)}\right\}$ son procesos regenerativos, con períodos de regeneración que poseen algunas propiedades "buenas" de "mezcla" y de existencia de los momentos exponenciales. Luego aplicaremos los métodos desarrollados para la estimación de las distancia, $\sigma(x_t^{(f)}, \tilde{x}_t^{(f)})$, uniforme sobre el tiempo t para los procesos regenerativos $\left\{x_t^{(f)}\right\}$ y $\left\{\tilde{x}_t^{(f)}\right\}$ (véase [17, 18, 19]). La clave aquí, es la de obtener "estimaciones de cercanía" de los procesos regenerativos $\left\{x_t^{(f)}\right\}$ y $\left\{\tilde{x}_t^{(f)}\right\}$, que no dependan de $f \in S$.

En el resto de este capítulo, sea $f \in S$ una política estacionaria arbitraria fija. Siempre supondremos que $x_0^{(f)} = \tilde{x}_0^{(f)} = x_*$, donde, x_* es el estado involucrado en S.3.1.

Definimos los momentos de regreso al estado x_* como sigue:

$$\tau \ : \ = \inf\left\{t \geq 1 \text{ tal que } x_t^{(f)} = x_*\right\} \tag{5.6}$$

$$\widetilde{\tau} \ : \ = \inf\left\{t \geq 1 \text{ tal que } \widetilde{x}_t^{(f)} = x_*\right\} \tag{5.7}$$

Si no existe tal t en (5.6) ó (5.7), entonces, hacemos que $\tau = \infty$ ó $\widetilde{\tau} = \infty$.

PROPOSICION 5.3. Bajo las suposiciones S.3.1, S.3.2 y S.3.3 se cumple lo siguiente:

$$\text{(a)}\mathrm{E}\exp(\mu\tau) \ \leq \ \mathrm{d}e\mu \tag{5.8}$$

$$\text{(b)}\mathrm{E}\exp(\widetilde{\mu}\widetilde{\tau}) \ \leq \ \widetilde{\mathrm{d}}e^{\widetilde{\mu}} \tag{5.9}$$

DEMOSTRACION. De hecho, las desigualdades (5.8) y (5.9) son consecuencias del corolario 2, del párrafo 5.2.1 en [17] , de donde, como un caso particular, se tiene que S.3.2 implica que $\mathrm{E}\exp(\mu\tau) \leq e\mu \int_X V(z)p(dz/x_*, f(x_*))$, de la misma manera se tiene la cota para $\mathrm{E}\exp(\widetilde{\mu}\widetilde{\tau})$. ■

OBSERVACION 5.4

(a) Como los estados iniciales $x_0 = \widetilde{x}_0 = x_*$ y la política $f \in S$ son fijos, en esta sección, entonces en la proposición 5.3 denotamos la esperanza por E en lugar de E_x^f.

(b) Nótese que de S.3.2, (b) los lados derechos de las desigualdades (5.8) y (5.9) no depende de la política $f \in S$. Por lo tanto, hemos obtenido las cotas uniformes (con respecto de f) para los momentos exponenciales de τ y $\widetilde{\tau}$.

PROPOSICION 5.5 Bajo las suposiciones S.3.1-S.3.3 y tomando a $x_0 = \widetilde{x}_0 = x_*$, los procesos $\left\{x_t^{(f)}\right\}$ y $\left\{\widetilde{x}_t^{(f)}\right\}$, son procesos regenerativos con periódo de regeneración, repectivamente, τ y $\widetilde{\tau}$, definidos en (5.6) y (5.7).

DEMOSTRACION. Es claro, que es suficiente mostrar la proposición para el proceso $\left\{x_t^{(f)}\right\}$. De (5.8) la esperanza de la función, estrictamente positiva, exp(μs) es finita y de ésto se sigue que $P(\tau < \infty) = 1$. La v.a. τ es tiempo de paro para el proceso $\left\{x_t^{(f)}\right\}$. Así, como cualquier proceso de Markov a tiempo discreto tiene la propiedad de Markov fuerte (véase, por ejemplo, [17, Cap.5]), se sigue que dado $x\tau^{(f)} = x_*$, las distribuciones condicionales de $x_{\tau+m}^{(f)}, m > 1$, no dependen de lo que ocurre con $x_t^{(f)}$, para $t < \tau$. Si definimos por inducción "los momentos de regeneración" como sigue:

$$
\begin{aligned}
\tau_1 \; &: = \tau; \\
\tau_2 \; &: = inf \left\{ t > \tau_1 \quad \text{tal que } x_t^{(f)} = x_* \right\} \\
\tau_3 \; &: = inf \left\{ t > \tau_2 \quad \text{tal que } x_t^{(f)} = x_* \right\} \\
&\;\vdots
\end{aligned}
$$

Entonces, para cada $i \geq 1$, la parte de la trayectoria denotada por Tr(i), esta dada por:

$$
\mathrm{Tr(i)} := \left\{ x_{\tau_i}^{(f)},\ x_{\tau_{i+1}}^{(f)},\ x_{\tau_{i+2}}^{(f)}, \ldots\ldots \right\},
$$

y esta es independiente de $x_t^{(f)}$, con $t < \tau_i$, y además, la distribución de la Tr(i) coincide con la distribución de la trayectoria $\left\{x_0^{(f)},\ x_1^{(f)},\ x_2^{(f)}, \ldots\ldots\right\}$. Estos dos hechos significan que $\left\{\ x_t^{(f)}\right\}$ es un proceso regenerativo. (véase [17, Cap.7]). ■

5.3 DEMOSTRACION DE LOS TEOREMAS 4.1 Y 4.2

Usaremos la siguiente definición (véase [17, Cap.7]).

DEFINICION 5.6 Sean N $\geq$ 1 un número entero y q > 0. Se dice que la distribución de la variable aleatoria (v.a) τ, con valores en $\{1, 2, 3, ...\}$, pertenece a la clase UNP(N,q) si el máximo común divisor de los enteros j $\leq$ N, tal que, $P(\tau = j) \geq q$, es igual a 1.

DEMOSTRACION DEL TEOREMA 4.1

En virtud de la desigualdad (5.3) y la proposición 5.5, solo falta verificar que para los procesos regenerativos $\left\{ x_t^{(f)} \right\}$ y $\left\{ \tilde{x}_t^{(f)} \right\}$, la siguiente desigualdad se cumple para algunas constantes k_2,$\alpha > 0$, independientes de la política $f \in S$.

$$\sup_{t\geq 0} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}) \leq k_2 \sigma(F, \tilde{F}) \left[1 - \ln[\alpha\sigma(F, \tilde{F})]\right] \tag{5.10}$$

Para hacerlo, utilizamos los resultados sobre las estimaciones cuantitativas de estabilidad de procesos regenerativos en [17], de éstos se tiene, que la desigualdad (5.10) es válida para algunas constantes k_2,α, calculadas mediante las constantes, δ, μ, V(x), d, $\tilde{\delta}$, $\tilde{\mu}$, $\tilde{V}(x)$, $\tilde{d}$, involucradas en las suposiciones S.3.1-S.3.3, si se cumplen las siguientes condiciones:

H1: $\max_{0\leq t\leq n} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}) \leq n\sigma(F, \tilde{F})$, n= 1, 2, ... (5.11)

H2: Las distribuciones de las variables aleatorias τ y $\tilde{\tau}$ éstan en la clase UNP(N,q) para algún $N \geq 1$ y $q > 0$.

H3: $\max\{E\exp(\mu\tau), E\exp(\mu\tilde{\tau})\} < \infty$.

Según el corolario 1 de [17], si verificamos las condiciones H1-H3, entonces, (5.10) se cumple con algunas constantes k_2 y α, que son determinadas en función de las constantes mencionadas anteriormente, y por tanto k_2 y α no dependen de cualquier política estacionaria $f \in S$.

La condición H3 se sigue de la proposición 5.3. Debido a la suposición S.3.1, pues $P(\tau = 1) = P(x_1^{(f)} = x_* \,/\, x_0^{(f)} = x_*) \geq \delta > 0$.

Por último, el máximo común divisor, de la definición 5.6, es 1 y si tomamos $q = \delta$, entonces obtenemos que la distribución de la v.a. τ está en la clase UNP(1,δ), por ésto se satisface la condición H2.

Para concluir la demostración, falta verificar las desigualdades (5.11) en la condición H1. Primero mostraremos que:

$$\max_{0 \leq t \leq n} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}) \leq n \sup_{x \in X} \sup_{a \in A(x)} \sigma(p(./x,a), \tilde{p}(./x,a)), \qquad n = 1, 2, ... \tag{5.12}$$

Teniendo en cuenta que $\sigma(x_0^{(f)}, \tilde{x}_0^{(f)}) = \sigma(x_*, x_*) = 0$, la desigualdad (5.12) se sigue por inducción de la siguiente desigualdad:

$$\sigma(x_{t+1}^{(f)}, \tilde{x}_{t+1}^{(f)}) \leq \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}) + \sup_{x \in X} \sup_{a \in A(x)} \sigma(p(./x,a), \tilde{p}(./x,a)) \tag{5.13}$$

$$\text{para} \quad t = 0, 1, ...$$

Demostración de (5.13). De la definición (5.5) se tiene:

$$
\begin{aligned}
&\sigma(\mathrm{x}_{t+1}^{(f)}, \tilde{\mathrm{x}}_{t+1}^{(f)}) \\
&= \sup_{\rho:\|\rho\|_\infty \leq 1} \left| \mathrm{E}\rho(\mathrm{x}_{t+1}^{(f)}) - \mathrm{E}\rho(\tilde{\mathrm{x}}_{t+1}^{(f)}) \right| \\
&= \sup_{\rho:\|\rho\|_\infty \leq 1} \left| \int_{\mathrm{X}} \rho(\mathrm{z}) \mathrm{p}^{t+1}(\mathrm{dz}/\mathrm{x}_*, f(\mathrm{x}_*)) - \int_{\mathrm{X}} \rho(\mathrm{z}) \tilde{\mathrm{p}}^{t+1}(\mathrm{dz}/\mathrm{x}_*, f(\mathrm{x}_*)) \right|, \quad (5.14)
\end{aligned}
$$

donde, p^{t+1} y $\tilde{\mathrm{p}}^{t+1}$ son las probabilidades de transición en t+1 etapas, respectivamente, para los procesos $\left\{ \mathrm{x}_t^{(f)} \right\}$ y $\left\{ \tilde{\mathrm{x}}_t^{(f)} \right\}$. Aplicando la ecuación de Chapman-Kolmogorov y el teorema de Fubini a (5.14) se tiene:

$$
\begin{aligned}
\sigma(\mathrm{x}_{t+1}^{(f)}, \tilde{\mathrm{x}}_{t+1}^{(f)}) &= \sup_{\rho:\|\rho\|_\infty \leq 1} \Big| \int_{\mathrm{X}} \rho(\mathrm{z}) \int_{\mathrm{X}} \mathrm{p}(\mathrm{dz}/\mathrm{y}, f(\mathrm{y})) \mathrm{p}^t(\mathrm{dy}/\mathrm{x}_*, f(\mathrm{x}_*)) - \\
&\quad \int_{\mathrm{X}} \rho(\mathrm{z}) \int_{\mathrm{X}} \tilde{\mathrm{p}}(\mathrm{dz}/\mathrm{y}, f(\mathrm{y})) \tilde{\mathrm{p}}^t(\mathrm{dy}/\mathrm{x}_*, f(\mathrm{x}_*)) \Big|
\end{aligned}
$$

$$
\begin{aligned}
\sigma(\mathrm{x}_{t+1}^{(f)}, \tilde{\mathrm{x}}_{t+1}^{(f)}) &= \sup_{\rho:\|\rho\|_\infty \leq 1} \Big| \int_{\mathrm{X}} \mathrm{p}^t(\mathrm{dy}/\mathrm{x}_*, f(\mathrm{x}_*)) \int_{\mathrm{X}} \rho(\mathrm{z}) \mathrm{p}(\mathrm{dz}/\mathrm{y}, f(\mathrm{y})) - \\
&\quad \int_{\mathrm{X}} \tilde{\mathrm{p}}^t(\mathrm{dy}/\mathrm{x}_*, f(\mathrm{x}_*)) \int_{\mathrm{X}} \rho(\mathrm{z}) \tilde{\mathrm{p}}(\mathrm{dz}/\mathrm{y}, f(\mathrm{y})) \Big| \qquad (5.15)
\end{aligned}
$$

Al sumar y restar el término:

$$\int_X \tilde{p}^t(dy/x_*, f(x_*)) \int_X \rho(z)p(dz/y, f(y))$$

en (5.15), obtenemos:

$$\sigma(x_{t+1}^{(f)}, \tilde{x}_{t+1}^{(f)}) \leq \sup_{\rho:\|\rho\|_\infty \leq 1} \left| \int_X p^t(dy/x_*, f(x_*))h(y) - \int_X \tilde{p}^t(dy/x_*, f(x_*))h(y) \right| +$$

$$\sup_{\rho:\|\rho\|_\infty \leq 1} \left| \int_X \tilde{p}^t(dy/x_*, f(x_*)) \left[\int_X \rho(z)p(dz/y, f(y)) - \int_X \rho(z)\tilde{p}(dz/y, f(y)) \right] \right|, \quad (5.16)$$

dónde $h(y) := \int_X \rho(z)p(dz/y, f(y))$, es una función medible y además

$$\begin{aligned} \|h\|_\infty &= \sup_{y\in X} |h(y)| \leq \sup_{y\in X} \int_X |\rho(z)|\, p(dz\,/\,y, f(y)) \\ &\leq \sup_{y\in X} \int_X p(dz\,/\,y, f(y)) = 1 \end{aligned}$$

ya que, $\sup_{z\in X} |h(z)| \leq 1$, se sigue de (5.16) que

$$\sigma(x_{t+1}^{(f)}, \tilde{x}_{t+1}^{(f)}) \leq \sup_{h:\|h\|_\infty \leq 1} \left| \int_X p^t(dy/x_*, f(x_*))h(y) - \int_X \tilde{p}^t(dy/x_*, f(x_*))h(y) \right| +$$

$$\sup_{\rho:\|\rho\|_\infty \leq 1} \sup_{y\in X} \left| \int_X \rho(z)p(dz/y, f(y)) - \int_X \rho(z)\tilde{p}(dz/y, f(y)) \right| \qquad (5.17)$$

Como $p^t(./x_*, f(x_*))$ y $\tilde{p}^t(./x_*, f(x_*))$ son las distribuciones de $x_t^{(f)}$ y $\tilde{x}_t^{(f)}$, obtenemos (véase (5.5)) que el primer término en el lado derecho de (5.17) es exactamente $\sigma(x_t^{(f)}, \tilde{x}_t^{(f)})$. Cambiando el orden de los supremos sobre ρ y sobre "y", en el segundo término del lado derecho de (5.17), se obtiene, que este sumando es:

$$\sup_{y\in X} \sigma(p(./y, f(y)), \tilde{p}(./y, f(y))$$

y como $f(y) \in A(y)$, entonces, se tiene:

$$\sup_{y\in X} \sigma(p(./y, f(y)), \tilde{p}(./y, f(y)) \le \sup_{x\in X} \sup_{a\in A(x)} \sigma(p(./x, a), \tilde{p}(./x, a))$$

Por lo tanto, llegamos a la desigualdad (5.13), por lo que se prueba (combinando el razonamiento por inducción) la desigualdad (5.12). ■

Ahora (5.12) implicara la condición H1 (véase (5.11)) si se tiene la siguiente desigualdad:

$$\sup_{x\in X} \sup_{a\in A(x)} \sigma(p(./x,a),\tilde{p}(./x, a)) \le \sigma(F, \tilde{F}) \tag{5.18}$$

Usando las relaciones análogas a (1.4) y (3.3) para $p(./x,a)$ y $\tilde{p}(./x,a)$ obtenemos (véase también (2.5));

$$\sigma(p(./x, a), \tilde{p}(./x, a)) = 2 \sup_{B\in B(X)} |p(B/x, a) - \tilde{p}(B/x, a)|$$

$$= 2 \sup_{B\in B(X)} \mid P(x_{t+1} \in B/x_t = x, a_t = a) - P(\tilde{x}_{t+1} \in B/\tilde{x}_t = x, \tilde{a}_t = a) \mid$$

$$= 2 \sup_{B\in B(X)} \mid P(\Phi(x, a, \xi_t) \in B) - P(\Phi(x, a, \tilde{\xi}_t) \in B) \mid$$

$$\sigma(\mathrm{p}(./\mathrm{x},a),\tilde{\mathrm{p}}(./\mathrm{x},a)) = 2 \sup_{\mathrm{B}\in \mathrm{B}(\mathrm{X})} \mid \mathrm{P}(\xi_t \in \mathrm{G}_{\mathrm{x},a}^{-1}(\mathrm{B})) - \mathrm{P}(\tilde{\xi}_t \in \mathrm{G}_{\mathrm{x},a}^{-1}(\mathrm{B})) \mid \tag{5.19}$$

donde $\mathrm{G}_{\mathrm{x},a}(.) := \Phi(\mathrm{x},a,.)$. Pero, para cada $\mathrm{B} \in \mathrm{B}(\mathrm{X})$, $(\mathrm{x},a) \in \mathrm{IK}$, el conjunto $\mathrm{G}_{\mathrm{x},a}^{-1}(\mathrm{B})$ es de Borel en R^n. Por esto, el último término en (5.19) es menor o igual al termino:

$$2 \sup_{\mathrm{D}\in \mathrm{B}(\mathrm{R}^n)} \left|\mathrm{P}(\xi_t \in \mathrm{D}) - \mathrm{P}(\tilde{\xi}_t \in \mathrm{D})\right| = \sigma(\mathrm{F},\tilde{\mathrm{F}}).$$

Resumiendo lo dicho anteriormente, hemos verificado las condiciones H1-H3 y, por lo tanto, las desigualdades (5.10) y (5.3) implican la desigualdad (4.1) con

$\mathrm{k}_1 := 2\mathrm{rk}_2$ ■

DEMOSTRACION DEL TEOREMA 4.2

De la definición (2.4) y de la desigualdad (4.1) se sigue que

$$\overline{\triangle}\beta(\mathrm{x}_*) = \frac{\triangle\beta(\mathrm{x}_*)}{|\mathrm{Q}\beta(\mathrm{x}_*,f_*)|} \leq \frac{\mathrm{k}_1}{1-\beta}\frac{1}{|\mathrm{Q}\beta(\mathrm{x}_*,f_*)|}\sigma(\mathrm{F},\tilde{\mathrm{F}})\left[1-\ln[\alpha\sigma(\mathrm{F},\tilde{\mathrm{F}})]\right] \tag{5.20}$$

Por otro lado (véase (1.5), (1.6)) como $\mathrm{c}(\mathrm{x},a) \geq \kappa$, para toda $(\mathrm{x},a) \in \mathrm{IK}$, se obtiene

$$\mathrm{Q}\beta(\mathrm{x}_*,f_*) = \sum_{t=0}^{\infty}\beta^t \mathrm{E}_{\mathrm{x}_*}^{f_*}\mathrm{c}(\mathrm{x}_t,f_*(\mathrm{x}_t)) \geq \kappa\sum_{t=0}^{\infty}\beta^t = \kappa(1-\beta)^{-1} \tag{5.21}$$

Combinando (5.20) y (5.21) obtenemos la cota:

$$\frac{\mathrm{k}_1}{\kappa}\sigma(\mathrm{F},\tilde{\mathrm{F}})\left[1-\ln[\alpha\sigma(\mathrm{F},\tilde{\mathrm{F}})]\right]$$

que no depende de β, y, por lo tanto, se sigue la desigualdad (4.2). ■

OBSERVACION 5.6 Como ya se observó en la formulación del teorema 4.1, para las constantes $k_1 = 2rk_2$ y α, en (4.1) se puede acotar por arriba en los términos de las constantes involucradas en las suposiciones S.3.1-S.3.3. Los procedimientos de como hacerlo, están indicados en los capítulos 6 y 7 en [17]. Esta técnica se basa en la aplicación de la proposición 5.3 y de la condición H2, para estimar los momentos exponenciales de una v.a. que se llama "el tiempo de cruce" de los procesos regenerativos $\left\{ x_t^{(f)} \right\}$ y $\left\{ \tilde{x}_t^{(f)} \right\}$. El algoritmo de tal estimación es suficientemente complicado, por esta causa no podríamos dar las fórmulas explícitas para las constantes k_1 y α.

CAPITULO 6

UN EJEMPLO: ESTABILIDAD EN UN SISTEMA DE CONTROL DE INVENTARIOS

En este capítulo, estudiaremos la estabilidad de un sistema de inventario con capacidad de producción y almacenamiento finitos, en el cual, el "exceso" de la demanda se considera como una deuda que puede saldarse en el futuro. (Acerca de modelos de inventario estocásticos véase, por ejemplo, [22]). Nuestro fin aquí, es checar, que para este sistema de control se cumplen las suposiciones S.1.5, S.3.1-S.3.3 y también daremos las cotas de estabilidad, obtenidas en el capítulo 4, en situaciones particulares.

Considerando el funcionamiento del sistema a tiempo discreto t = 0,1,... (que corresponde a etapas o períodos de producción-almacenamiento) describiremos la dinámica del sistema original ("real") como sigue:

$$\begin{aligned} x_{t+1} &= \min\left\{U, x_t + a_t^{(1)}\xi_t - a_t^{(2)}\eta_t\right\}, \qquad t = 0, 1, \ldots \qquad (6.1) \\ x_0 &= x \in [0, U] \end{aligned}$$

donde, para la t-ésima etapa se denota:

U - es la capacidad máxima del almacén;

x_t - es el nivel de producto en el inventario;

$a_t^{(1)}$ - es la cantidad de producto solicitado a la unidad de producción, es decir, la cantidad planeada a producir (o a comprar) en la t-ésima etapa;

$a_t^{(2)}$ - es cantidad planeada de producto para la venta en la t-ésima etapa;

$a_t^{(1)}\xi_t$ - es la cantidad real producida (tomando en cuenta posibles desviaciones aleatórias) en la t-ésima etapa;

$a_t^{(2)}\eta_t$ - la venta real (tomando en cuenta la demanda real) realizada en la t-ésima etapa;

ξ_t - es la v.a. que describe las desviaciones de producción, en comparación a las cantidades planeadas de producción en la t-ésima etapa;

η_t - es la v.a. que describe la influencia de la demanda en la venta real del producto en la t-ésima etapa.

Nótese, que según (6.1), el estado del proceso x_t puede ser un número negativo (cuando hay exceso en la demanda). En este caso $x_t < 0$ se interpreta como la deuda del tamaño $|x_t|$ en la t-ésima etapa, que deberá pagarse por producción en las siguientes etapas (t+1), (t+2),....

SUPOSICION S.6.1.

(a) Las v.a's. $\xi_0, \xi_1, ...$ son no negativas e i.i.d. con densidad en común $\rho\xi$.

(b) Las v.a's. $\eta_0, \eta_1, ...$ son no negativas e i.i.d. con densidad en común $\rho\eta$ que es acotada y continua en $[0, \infty)$, además, para cada s $\in [0, \infty]$, la siguiente función es integrable:

$$\Psi(z) := \sup_{a^{(1)} \in [b, U], a^{(2)} \in [\theta_1, \theta_2]} \rho\eta\left(\frac{a^{(1)}s\text{-}z}{a^{(2)}}\right)$$

(c) Para cada t, las v.a's. ξ_t, η_t, son independientes.

Bajo esta suposición la relación de recurrencia (6.1) define un proceso de control de Markov, como en (2.1), con espacio de estados $X = (-\infty, U]$, conjunto de controles $A = [0, \infty)$ x $[0, \infty)$ y los conjuntos de controles admisibles como sigue:

$A(x) = \left\{(a^{(1)}, a^{(2)})\right\} = A_1(x) x A_2(x)$

con:

$$A_1(x) = \begin{cases} [b, U - x] & si \quad U - x \geq b \\ \{b\} & \text{si} \quad U - x < b \end{cases} \tag{6.2}$$

$$A_2(x) = \left\{a^{(2)}\right\} = [\theta_1, \theta_2] \tag{6.3}$$

donde, $0 < \theta_1 \leq \theta_2$, $b > 0$ son números dados.

La restricción (6.2) significa que hay un nivel minimo de la producción planeada $b > 0$ (que no depende del presente nivel del inventario y refleja la esperanza del producto a vender, por lo menos esta cantidad de producto). Por otro lado, si $x > b$, según (6.2), está prohibido producir más de U-x, para no sobrellenar el almacén. En todo lo que se sigue, denotaremos al par de controles $(a^{(1)}, a^{(2)}) \in \mathrm{A}(\mathrm{x})$ por a.

Supóngase que el costo por etapa se tiene de la siguiente forma:

$$\mathrm{c}(\mathrm{x}, a) = \mathrm{Ec}_1(\mathrm{x}, a^{(1)}, a^{(2)}, \xi_0, \eta_0), \tag{6.4}$$

donde c_1 es una función dada, acotada y continua en $a^{(1)}$ y $a^{(2)}$.

También se considerará el siguiente proceso que aproxima al proceso (6.1), dado de la siguiente manera:

$$\begin{aligned} \tilde{\mathrm{x}}_{t+1} &= \min\left\{U, \tilde{\mathrm{x}}_t + \tilde{a}_t^{(1)}\tilde{\xi}_t - \tilde{a}_t^{(2)}\tilde{\eta}_t\right\} \qquad t = 0, 1, ... \\ \tilde{\mathrm{x}}_0 &= \mathrm{x} \in [0, U] \end{aligned} \tag{6.5}$$

Con espacios de estados $[-\infty, U]$, conjunto de controles admisibles $(\tilde{a}_t^{(1)}, \tilde{a}_t^{(2)}) = \tilde{a} \in \mathrm{A}(\tilde{\mathrm{x}}_t)$ y función de costo por etapa (como en (6.4)) son los mismos que se tienen para el proceso (6.1).

SUPOSICION S.6.2.

(a) Las v.a's. $\tilde{\xi}_0, \tilde{\xi}_1, ...$ son no negativas e i.i.d. con densidad en común $\tilde{\rho}\xi$.

(b) Las v.a's. $\tilde{\eta}_0, \tilde{\eta}_1, ...$ son no negativas e i.i.d. con densidad en común $\tilde{\rho}\eta$ que es acotada y continua en $[0, \infty)$, además, para cada s $\in [0, \infty]$, la siguiente función es integrable:

$$\tilde{\Psi}(z) := \sup_{a^{(1)} \in [b, U], a^{(2)} \in [\theta_1, \theta_2]} \tilde{\rho}\eta\left(\frac{a^{(1)}\mathrm{s\text{-}z}}{a^{(2)}}\right)$$

(c) Para cada t, las v.a's. $\tilde{\xi}_t$ y $\tilde{\eta}_t$ son independientes.

Ahora verificaremos que la suposición de la secciones 1.4 y las suposiciones de capítulo 3 se satisfacen para el sistema de control descrito. Comenzaremos con S.1.5.

De (6.2)-(6.4), se ve facilmente, que las condiciones S.1.5,(a)-(b) se cumplen, Chequemos S.1.5,(c). Fijemos $x \in (-\infty, U]$ y sea $u : (-\infty, U] \to R$ una función medible y acotada, arbitraria. Por (1.2), S.6.1 y aplicando el teorema de Fubini, obtenemos:

$$\begin{aligned} q(a) \quad : \quad &= \int_X u(y)p(dy/x, a) = Eu\left[\min\left\{U, x + a^{(1)}\xi_0 - a^{(2)}\eta_0\right\}\right] \\ &= \int_0^\infty \int_0^\infty u\left[\min\left\{U, x + a^{(1)}s_1 - a^{(2)}s_2\right\}\right] \rho\xi(s_1)\rho\eta(s_2)ds_1ds_2 \\ &= \int_0^\infty \rho\xi(s_1)ds_1 \int_0^\infty u\left[\min\left\{U, x + a^{(1)}s_1 - a^{(2)}s_2\right\}\right] \rho\eta(s_2)ds_2 \ , \end{aligned}$$

y tomando el cambio de variable $z = a^{(1)}s_1 - a^{(2)}s_2$ en la segunda integral se tiene:

$$q(a) = \int_0^\infty \rho\xi(s_1)ds_1 \frac{1}{a^{(2)}} \int_{-\infty}^{a^{(1)}s_1} u\left[\min\left\{U, x + z\right\}\right] \rho\eta(\frac{a^{(1)}s_1 - z}{a^{(2)}})dz \qquad (6.6)$$

Observemos que como $a^{(2)} \in [\theta_1, \theta_2]$, $\theta_1 > 0$ y la función u es acotada, se sigue que la función:

$$\rho_{a^{(1)},a^{(2)}}(s_1) := \frac{1}{a^{(2)}} \int_{-\infty}^{a^{(1)}s_1} u\left[\min\left\{U, x + z\right\}\right] \rho\eta(\frac{a^{(1)}s_1 - z}{a^{(2)}})dz$$

es acotada uniformemente con respecto de $a^{(1)}, a^{(2)}$. Así, para mostrar que la función $q(a)$, en (6.6), es continua, bastará probar, utilizando el teorema de Lebesgue sobre convergencia dominada, que para cada $s_1 \in [0,\infty)$, fijo, la continuidad de la función $g(a)$, definida como sigue:

$$g(a) = \int_{-\infty}^{a^{(1)}s_1} u\left[\min\{U, x+z\}\right]\rho\eta(\frac{a^{(1)}s_1 - z}{a^{(2)}})dz.$$

Consideramos cualquier sucesión $\{a_n\} = \left\{a_n^{(1)}, a_n^{(2)}\right\} \in A(x)$, tal que, $a_n \to a$, entonces

$$|g(a) - g(a_n)|$$

$$\leq \left|\int_{a_n^{(1)}s_1}^{a^{(1)}s_1} u\left[\min\{U, x+z\}\right]\rho\eta(\frac{a^{(1)}s_1 - z}{a^{(2)}})dz\right| +$$

$$\int_{-\infty}^{a^{(1)}s_1} |u\left[\min\{U, x+z\}\right]| \left|\rho\eta(\frac{a_n^{(1)}s_1 - z}{a_n^{(2)}})dz - \rho\eta(\frac{a^{(1)}s_1 - z}{a^{(2)}})\right| dz \qquad (6.7)$$

El primer sumando en (6.7) se anula cuando $n \to \infty$, ya que, u y $\rho\eta$ son acotadas, mientras que el segundo converge a cero, por el teorema de Lebesgue, que es aplicable aquí debido a la condición S.6.1.(b).

La condición S.1.5,(d) para la probabilidad de transición $\tilde{p}$ se verifica de igual manera. Pasemos a verificar las suposiciones S.3.1 - S.3.3, para esto se necesitan más condiciones sobre los procesos (6.1) y (6.5), que a continuación se dan.

SUPOSICION S.6.3. Existe una constante $\delta > 0$, tal que

.(a)$P(b\xi_0 \geq \theta_2\eta_0) \geq \delta$

(b) $P(b\tilde{\xi}_0 \geq \theta_2\tilde{\eta}_0) \geq \delta$.

También introduciremos las siguientes notaciones:

$\xi := \theta_2\eta_0 - \mathrm{b}\xi_0, \qquad \widetilde{\xi} := \theta_2\widetilde{\eta}_0 - \mathrm{b}\widetilde{\xi}_0$

Y requeriremos de lo siguiente:

SUPOSICION S.6.4. Existen dos constantes, $\mu_0 > 0$ y $\widetilde{\mu}_0 > 0$, tales que:

(a) $\mathrm{E}\exp(\mu_0\xi) < \infty$.

(b) $\mathrm{E}\exp(\widetilde{\mu}_0\widetilde{\xi}) < \infty$.

Por la suposición S.6.3, es fácil ver, que se cumple la suposición S.3.1, tomando $\mathrm{x}_* = U$. De hecho, con respecto de S.3.1,(a) se tiene, que para cualquier política $f = (f_1, f_2) \in \mathrm{S}$ lo siguiente:

$$\mathrm{p}(\{\mathrm{x}_*\}/\mathrm{x}_*, f(\mathrm{x}_*)) = \mathrm{P}\left(\min\{U, U + f_1(U)\xi_0 - f_2(U)\eta_0\} = U\right)$$

$$= \mathrm{P}\left(f_1(U)\xi_0 - f_2(U)\eta_0 \geq 0\right) \geq \mathrm{P}\left(\mathrm{b}\xi_0 - \theta_2\eta_0 \geq 0\right) \geq \delta > 0$$

Debido a que: $f_1(U) = \mathrm{b}$, $f_2(U) \leq \theta_2$, (véase (6.2), (6.3)) y de S.6.3,(a). Del mismo modo se verifica S.3.1,(b).

La verificación de las suposiciónes S.3.2 y S.3.3 se realizan de la misma manera, entonces es suficiente checar S.3.2. Ahora para mostrar S.3.2 se hace un cambio de variable, por comodidad, en (6.1) tomando $\mathrm{y}_t := U - \mathrm{x}_t$. Entonces, de (6.1) se tiene:

$$U - \mathrm{y}_{t+1} = \min\left\{U, U - \mathrm{y}_t + a_t^{(1)}\xi_t - a_t^{(2)}\eta_t\right\}$$

o bién:

$$-\mathrm{y}_{t+1} = \min\left\{0, -\mathrm{y}_t + a_t^{(1)}\xi_t - a_t^{(2)}\eta_t\right\}$$

y como - $\min\{b, d\} = \max\{-b, -d\}$, entonces, se tiene:

$$y_{t+1} = \max\left\{0, y_t + a_t^{(2)}\eta_t - a_t^{(1)}\xi_t\right\}, \qquad t = 0, 1, 2, ... \tag{6.8}$$

La relación (6.8) define un proceso de control de Markov (con espacio de estados $Y = [0, \infty))$ que es equivalente al proceso original (6.1), si adicionalmente hacemos el cambio de variable $x = U - y$, en la definición del costo (6.4). Observemos que él estado $x_* = U$ para el proceso (6.1) pasa al estado $y_* = 0$ para (6.8).

Para ver que se cumple S.3.2. elegimos la "función de prueba" V como sigue:

$$V(y) = \begin{cases} 1 & \text{si } \; y = 0 \\ \tilde{b}e^{\gamma y} & \text{si } \; y > 0 \end{cases} \tag{6.9}$$

donde $\tilde{b} \geq 1$ y $\gamma > 0$, son algunas constantes.

Para cada política estacionaria $f = (f_1, f_2)$ se tiene:

$$\begin{aligned} \int_X V(z)p(dz/y, f(y)) &= EV\left[\max\{0, y + f_2(y)\eta_0 - f_1(y)\xi_0\}\right] \\ &\leq EV\left[\max\{0, y + \theta_2\eta_0 - b\xi_0\}\right] = \\ &= EV\left[\max\{0, y \ + \xi\}\right] \end{aligned} \tag{6.10}$$

La desigualdad en (6.10) es válida, ya que la función V es creciente, $f_2(y) \in [\theta_1, \theta_2]$ y $f_1(y) \geq b$, entonces:

$$y + f_2(y)\eta_0 - f_1(y)\xi_0 \leq y + \theta_2\eta_0 - b\xi_0$$

En el libro [17, Cap.8] , bajo la suposición S.6.4,(a), se demuestra que existe una constante $\mu > 0$, tal que para toda $y \neq 0$, la última esperanza en (6.10) es menor que $e^{-\mu}V(y)$, por esto se obtiene S.3.2,(a). De nuevo aplicando la desigualdad (6.10) con $y = y_* = 0$, tenemos, que para cada $f \in S$:

$$\int_X V(z)p(dz/y_*, f(y_*)) \leq EV[\max\{0,\xi\}] < \infty$$

El lado derecho de esta desigualdad es independiente de la política $f \in S$. Por consiguiente se cumple S.3.2,(b).

Resumiremos lo antes dicho en la siguiente afirmación:

PROPOSICION 6.5 Si se satisfacen las condiciones S.6.1 - S.6.4, entonces, se cumplen las suposiciones S.1.5, S.3.1 - S.3.3 y suponiendo que $x_0 = \tilde{x}_0 = U$, para los sistemas de control (6.1) y (6.5) se aplican las desigualdades de estabilidad (4.1) y (4.2) (esto último, a condición de que el costo en (6.4) es estrictamente positivo.). Además en este ejemplo se podra utilizar en lugar de $\sigma(F,\tilde{F})$ en (4.1) y (4.2), la siguiente relación más sencilla:

$$\varphi(F,\tilde{F}) := \int_0^\infty |\rho\xi(s) - \tilde{\rho}\xi(s)|\, ds + \int_0^\infty |\rho\eta(s) - \tilde{\rho}\eta(s)|\, ds \tag{6.11}$$

DEMOSTRACION. Hay que mostrar sólo la última afirmación, primero aseguremos que se cumple:

$$\sigma(F,\tilde{F}) \leq \varphi(F,\tilde{F}) \tag{6.12}$$

En efecto, se sabe (véase, por ejemplo,[4]) que si existen densidades $\rho_{\eta,\xi}$ y $\tilde{\rho}_{\eta,\xi}$, respectivamente, de las distribuciones F y $\widetilde{F}$, entonces, la distancia de variación total σ esta representada por la siguiente fórmula:

$$\sigma(F,\widetilde{F}) = \int_{R^2} \left|\rho_{\xi,\eta}(s,u) - \widetilde{\rho}_{\xi,\eta}(s,u)\right| dsdu$$

que a su vez por independencia de las v.a. ξ y η se tiene:

$$\sigma(F,\widetilde{F}) = \int_{R^2} |\rho\xi(s)\rho\eta(u) - \widetilde{\rho}\xi(s)\widetilde{\rho}\eta(u)|\, dsdu \tag{6.13}$$

Restando y sumando en (6.13) el término $\rho\xi(s)\widetilde{\rho}\eta(u)$ y aplicando el teorema de Fubini, se tiene:

$$\begin{aligned}
\sigma(F,\widetilde{F}) &\leq \int_0^\infty \int_0^\infty |\rho\xi(s)\rho\eta(u) - \rho\xi(s)\widetilde{\rho}\eta(u)|\, dsdu + \int_0^\infty \int_0^\infty |\rho\xi(s)\widetilde{\rho}\eta(u) - \widetilde{\rho}\xi(s)\widetilde{\rho}\eta(u)|\, dsdu \\
&= \int_0^\infty \rho\xi(s)ds \int_0^\infty |\rho\eta(u) - \widetilde{\rho}\eta(u)|\, du + \int_0^\infty \widetilde{\rho}\eta(u)du \int_0^\infty |\rho\xi(s) - \widetilde{\rho}\xi(s)|\, ds \\
&= \varphi(F,\widetilde{F})
\end{aligned}$$

Lo cual prueba (6.12). ■

Finalmente, es muy fácil ver que la función $s[1 - \ln(\alpha s)]$, en los lados derechos de (4.1) y (4.2), es no decreciente. Entonces, de (6.12) podemos utilizar en las desigualdades (4.1) y (4.2) la cantidad $\varphi(F,\widetilde{F})$ en vez de $\sigma(F,\widetilde{F})$.

Para concluir este capítulo, consideremos las aplicaciones de las cotas (4.1) y (4.2) a los procesos (6.1) y (6.5), en el caso particular cuando las v.a's. ξ, $\widetilde{\xi}$; η, $\widetilde{\eta}$ tienen las distribuciones exponenciales con parámetros, λ, $\lambda + \epsilon$, ω, $\omega + \epsilon$, respectivamente.

Primero estimaremos la siguiente integral:

$$\mathrm{I}_1(\epsilon) := \int_0^{\infty} \left|\lambda e^{-\lambda s} - (\lambda+\epsilon)e^{-(\lambda+\epsilon)s}\right| ds \tag{6.14}$$

que es un sumando en (6.11), y tomando $s_* = -\ln(\frac{\lambda}{\lambda+\epsilon})^{\frac{1}{\epsilon}}$ se tiene:

$$\begin{array}{lcl} \lambda e^{-\lambda s} - (\lambda+\epsilon)e^{-(\lambda+\epsilon)s} & < & 0 \qquad \text{si} \quad s < s_* \\ \lambda e^{-\lambda s} - (\lambda+\epsilon)e^{-(\lambda+\epsilon)s} & \geq & 0 \qquad \text{si} \quad s \geq s_* \end{array}$$

entonces, se tiene lo siguiente:

$$\mathrm{I}_1(\epsilon) := \int_0^{s_*} \left[(\lambda+\epsilon)e^{-(\lambda+\epsilon)s} - \lambda e^{-\lambda s}\right] ds + \int_{s_*}^{\infty} \left[\lambda e^{-\lambda s} - (\lambda+\epsilon)e^{-(\lambda+\epsilon)s}\right] ds$$

$$\begin{aligned} \mathrm{I}_1(\epsilon) &= 2(e^{-\lambda s_*} - e^{-(\lambda+\epsilon)s_*}) \\ &= 2\left\{(\frac{\lambda}{\lambda+\epsilon})^{\frac{\lambda}{\epsilon}} - (\frac{\lambda}{\lambda+\epsilon})^{\frac{\lambda+\epsilon}{\epsilon}}\right\} \\ &= 2(1-\frac{\lambda}{\lambda+\epsilon})(\frac{\lambda}{\lambda+\epsilon})^{\frac{\lambda}{\epsilon}} \\ &= 2(\frac{\epsilon}{\lambda+\epsilon})(\frac{\lambda}{\lambda+\epsilon})^{\frac{\lambda}{\epsilon}} \\ &= \frac{2\epsilon}{\lambda}(\frac{\lambda}{\lambda+\epsilon})^{\frac{\lambda+\epsilon}{\epsilon}} \end{aligned} \tag{6.15}$$

Pero:

$$\frac{\lambda+\epsilon}{\epsilon} \geq 1, \frac{\lambda}{\lambda+\epsilon} \leq 1$$

ésto implica:

$$(\frac{\lambda}{\lambda+\epsilon})^{\frac{\lambda+\epsilon}{\epsilon}} \leq 1.$$

y de (6.15) se sigue:

$$\mathrm{I}_1(\epsilon) \leq \frac{2\epsilon}{\lambda} \tag{6.16}$$

De manera similar se tiene:

$$\mathrm{I}_2(\epsilon) := \int_0^\infty \left| \omega e^{-\omega s} - (\omega + \epsilon_1) e^{-(\omega+\epsilon 1)s} \right| \mathrm{ds} \leq \frac{2\epsilon_1}{\lambda} \tag{6.17}$$

Por último, por (4.1), (4.2), (6.11), (6.12), (6.16) y (6.17), se obtienen las siguientes cotas superiores para los índices de estabilidad, para este ejemplo.

i)Para el índice de estabilidad descontado

$$\triangle\beta(U) \leq \frac{2\mathrm{k}_1}{1-\beta} \left[\frac{\epsilon}{\lambda} + \frac{\epsilon_1}{\omega} \right] \left[1 - \ln[2\alpha(\frac{\epsilon}{\lambda} + \frac{\epsilon_1}{\omega})] \right]$$

ii)Para el índice relativo (cuando en (6.4), $\mathrm{c}_1 \geq \kappa > 0$)

$$\sup_{\beta\in(0,1)} \overline{\triangle}\beta(U) \leq \frac{2\mathrm{k}_1}{\kappa} \left[\frac{\epsilon}{\lambda} + \frac{\epsilon_1}{\omega} \right] \left[1 - \ln[2\alpha(\frac{\epsilon}{\lambda} + \frac{\epsilon_1}{\omega})] \right]$$

CAPITULO 7

CONCLUSIONES Y PROBLEMAS ABIERTOS

En este trabajo se han obtenido nuevas estimaciones para los índices de estabilidad (2.3) y (2.4) para los procesos generales controlables de Markov. Las correspondientes desigualdades (4.1) y (4.2), están demostradas bajo las condiciones S.3.1-S.3.3, "de la existencia del estado de regeneración" y de las condiciones del tipo Liapunov, las cuales son bastantes restrictivas y por ésto se reduce el campo de aplicación potencial. Sin embargo, las condiciones mencionadas se satisfacen para muchos modelos de control: colas, inventarios, presas, etc. (véase ejemplos en [11, 16, 24,]).

Por otro lado, la cota para el índice de estabilidad descontado en (4.1) tiene mejores propiedades asíntoticas, cuando el factor de descuento se aproxima a 1, en comparación con los resultados obtenidos en los trabajos previos. Esta característica nos permite obtener la cota (4.2), que es uniforme con respecto al factor de descuento. Es importante también, que nuestras cotas están expresada directamente por la métrica de variación total $\sigma(\mathrm{F},\tilde{\mathrm{F}})$ (entre las distribuciones F y su aproximación $\tilde{\mathrm{F}}$) y, por consiguiente, las desigualdades (4.1) y (4.2) no contienen ningún término que no se exprese en términos de los modelos y de las suposiciones (compare con [1,2]).

Las direcciones a continuar para este trabajo, que parecen ser interesantes e importantes para aplicaciones, son:

A. Relajación de las condiciones S.3.1-S.3.3 para ampliar el campo de aplicaciones

B. Extensión de los resultados a otra clase de procesos controlables.

Comenzemos con un comentario sobre **B**, notemos que los procesos considerados en esta tesis están definidos en espacios muy generales, de estados y controles, de Borel, así, no hay nada para generalizar en esta dirección. Por otro lado, se puede tratar de resolver el problema para los modelos con función de costo por etapa no acotados, estos últimos aparecen naturalmente en ciertos sistemas de control. Hasta donde sabemos, el último trabajo donde se tiene la cota de estabilidad uniforme,con respecto de β, para los modelos Markovianos con costo no acotada, está en el artículo [10], donde se realiza un enfoque distinto del presente trabajo. También parece interesante y no muy difícil, es extender nuestros resultados a la clase de procesos semi-markovianos, que son más adecuados para algunos modelos de control en la práctica. (Acerca de los éxitos de la aplicación de las condiciones de Liapunov, para procesos de este tipo , véase [20]).

Con respecto de la dirección **A**, podriamos observar que si buscamos, como en este trabajo, cotas de estabilidad "casi lineales" en $\sigma(\mathrm{F},\tilde{\mathrm{F}})$, entonces no hay mucha esperanza de generalizar seriamente las suposiciones S.3.2, S.3.3. Parece ser, que lo único que se puede hacer aquí, es reemplazar el punto fijo x_* en S.3.2 y S.3.3 por un "conjunto pequeño" (véase [21]). Después de tal reemplazo, las condiciones como en S.3.2 serian equivalentes a ergodicidad geométrica de los procesos (aplicando las políticas estacionarias). En su momento, parece que tal ergodicidad es importante para obtener cotas "casi lineales " (véase [5,7]).

Las condiciones en S.3.1, en nuestra opinión, pueden generalizarse sin perder las propiedades de las desigualdades (4.1) y (4.2). En lugar de la existencia de un "punto de regeneración" (como en S.3.1 o más general un "conjunto de regeneración") con probabilidad positiva de regreso a este punto en una etapa, puede ser posible usar algunas propiedades de recurrencia general de procesos Markovianos (como , por ejemplo, recurrencia de Harris positiva, véase [21]).

Finalmente notemos que la desigualdad (4.1) implica la cota uniforme en (4.2), no sólo en el caso cuando la función de costo es estrictamente positiva. Para relajar esto último, hay que análizar la asíntota del costo óptimo $Q\beta^*(x)$ cuando $\beta \to 1$, usando adecuados teoremas Tauberianos (véase [4.Cap.13]).

REFERENCIAS

[1] Van Dijk N. M. (1988), Perturbation theory for unbounded Markov reward processes with application to queueing. **Adv. Appl. Prob. 20**, 99-111.

[2] Van Dijk N. M. and Puterman M. L. (1988), Perturbation theory for Markov reward processes with

application to queueing systems. **Adv. Appl. Prob. 20**, 79-98.

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

[4] Feller W. (1978), **Introducción en la Teoría de Probabilidad y sus Aplicaciones**. Vol.II, Limusa, México.

[5] Gordienko E. I. (1992), An estimate of the stability of optimal control of certain stochastic and deterministic systems. **J. Soviet Math. 59**, 891-899.

[6] Gordienko E. I. (1988), Stability estimates for controlled Markov chains with a minorant. **J. Soviet Math. 40**, 481-486.

[7] Gordienko E. I. (1994), **Lectures Notes on Stability Estimation in Markov Decision Processes,** Universidad Atónoma Metropolitana, México D. F.

[8] Gordienko E. I. (1985), Adaptive Strategies for certain classes of controlled Markov processes.**Theory Probab. Appl.29**, 504-518

[9] Gordienko E. I. and Salem F. S. (1997), Robustness inequality for Markov control processes with unbounded cost. To appear in **Syst. Control Lett.**

[10] Gordienko E. I. and Salem F. S. (1997), Estimates of Stability of Markov Control Processes with unbounded costs. Submitted in **Kybernetika.**

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

[12] Gordienko E. I., Isauro Martinez M. E. and Marcos Castillo R. M. (1997), Estimation of stability in controlled storage systems. **Reporte de investigación,** 04.0405I.01.001.97, UAM-I, México D. F.

[13] Hernández - Lerma O. (1989), **Adaptive Markov control processes.** Springer - Verlag, New York.

[14] Hernández - Lerma O. and Cavazos - Cadena R. (1990), Density estimation and adaptive control of Markov processes; average and discounted criteria. **Acta Appl. Math, 20**, 285 - 307.

[15] Hernández - Lerma O. and Lasserre J. B. (1990), Average cost optimal policies for Markov control processes with Borel state space and unbounded costs. **Syst. Control Lett. 15**, 349 - 356.

[16] Hernández - Lerma O. and Lasserre J. B. (1996), **Discrete - time Markov control processes: Basic Optimality Criteria.** Springer - Verlag, New York.

[17] Kalashnikov V. (1994), **Mathematical Methods in Queuing Theory**, Kluwer Acad. Publ. Dordrecht.

[18] Kalashnikov V. V. (1994), Regeneration and general Markov chains.

J. Appl. Math. Stoch. Analysis, 7, 357 - 371.

[19] Kalashnikov V.V. (1994), Crossing and comparation of regenerative processes.
Acta. Appl. Math. 34, 151 - 172.

[20] Luque _ Vásquez F. (1997), **Modelos de control semi-Markovianos en Espacios de Borel.** Tesis de Doctorado de Ciencias (Matemáticas), Facultad de Ciencias, UNAM.

[21] Meyn S. P. and Tweedie R. L. (1993), **Markov chains and stochastic Stability.** Springer - Verlag, London.

[22] Porteus E. L. (1990), Stochastic inventory theory. In: Neyman D. P. and Sobel M. J. (Editors) **Stochastic Models**, Vol. 2. North - Holland, Amsterdan, 606 - 652.

[23] Rachev S. T. (1991), **Probability Metrics and the Stability of Stochastic Models.** Wiley, New York.

[24] Vega - Amaya O. and Montes de Oca R. (1996), Application of average dynamic programming to inventory systems. To appear in **Math. Methods Oper. Res.**