152539

PROCESOS DE CONTROL DE MARKOV CON COSTOS Y CONTROLES NO ACOTADOS

Tesis que presenta:

Jesús Adolfo Minjárez Sosa

para obtener el grado de Maestro en Matemáticas

Directores: Dr. Onésimo Hernández Lerma
M.C. Raúl Montes de Oca Machorro

Universidad Autónoma Metropolitana - Iztapalapa
División de Ciencias Básicas e Ingeniería
Departamento de Matemáticas

AGOSTO DE 1993

A mi Fran

A mis padres

A mis hermanos

A mis compañeros

Agradezco al Consejo Nacional de Ciencia y Tecnología por el apoyo brindado para realizar mis estudios de maestría, y a los departamentos de matemáticas de la Universidad de Sonora y de la Universidad Autónoma Metropolitana-Iztapalapa.

También quiero agradecer, de manera especial al Dr. Onésimo Hernández Lerma y al M.C. Raúl Montes de Oca Machorro por aceptar dirigir esta tesis.

PROCESOS DE CONTROL DE MARKOV CON COSTOS Y CONTROLES NO ACOTADOS

INDICE

INTRODUCCION 1

I.- Procesos de Control de Markov

1.- Introducción 4

2.- Modelo de control de Markov 5

3.- Políticas de control 7

4.- Espacio canónico de probabilidad 9

5.- Problema de control óptimo 12

6.- Ejemplos 14

II.- Procesos de Control de Markov con Costos Descontados

1.- Introducción 20

2.- Hipótesis sobre el modelo de control 21

3.- Ecuación de optimalidad α-descontada 28

III.- Procesos de Control de Markov con Costo Promedio

1.- Introducción 42

2.- Hipótesis sobre el modelo de control 43

3.- Condiciones de optimalidad (I) 46

4.- Condiciones de optimalidad (II) 66

CONCLUSIONES 73

APENDICE 77

BIBLIOGRAFIA 82

INTRODUCCION

La Teoría de Control Optimo es una de las áreas de las matemáticas modernas que ha alcanzado un gran desarrollo en los últimos años. Sus aplicaciones están en problemas donde se requiere controlar sistemas dinámicos, los cuales se presentan en campos tan diversos como: Economía, Biología, Medicina y en todas las ramas de la Ingeniería.

Su objetivo es buscar la manera de operación óptima de dichos sistemas, en el sentido de que se deben determinar las acciones o decisiones de control que deben tomarse para optimizar su comportamiento. Dicho comportamiento lo mide una función a la que llamaremos Indice de Funcionamiento, la cual en nuestro caso representará el costo de operación del sistema.

En todos estos campos de aplicación, surge la necesidad de considerar elementos estocásticos que modelen cierta incertidumbre que influye en el comportamiento del sistema. En este trabajo nos ocuparemos de la teoría de control óptimo estocático.

La teoría de control tuvo sus origenes en el cálculo de variaciones, y fué hasta los años 50's cuando se le dió gran impulso. Se propusieron técnicas para resolver el problema de control. Una de ellas fué la Programación Dinámica propuesta por Bellman en el año de 1951, la cual toma gran importancia ya que se puede extender directamente al caso estocástico.

Una clase de sistemas de control estocástico, la constituyen los Procesos de Control de Markov (PCM), cuya teoría se apoya básicamente en la técnica de Programación Dinámica Estocástica.

El estudio de los PCMs se puede dividir, por ejemplo:

1.- Según el tipo de espacios de estados:

(a) espacio numerable;

(b) espacio de Borel (i.e. subconjunto de Borel de un espacio métrico separable y completo);

2.- Segun el tipo de indice de funcionamiento:

(a) costo descontado;

(b) costo promedio;

(c) costo total.

Dentro de esta última clasificación, se da el caso en que los costos pueden ser acotados o no acotados. La literatura más común sobre PCMs se desarrolla en el caso 1(a) con costos acotados.

En este trabajo, estamos interesados en analizar condiciones de optimalidad para los casos 1(b) y 2(b), a tiempo discreto, permitiendo costo y controles no acotados.

Dentro de los caminos para estudiar el caso promedio está el hacerlo por medio del caso descontado. Esto lo propuso inicialmente Blackwell en 1962 y consiste en analizar el caso promedio como un límite del caso descontado. A este esquema se le conoce como el "enfoque del factor de descuento desvaneciente" (vanishing discount factor approach) y es precisamente en el que estamos interesados para desarrollar este trabajo.

En el Capítulo I, presentamos los diferentes elementos de la teoría de control óptimo estocástico, y en particular la de los PCMs. Definimos el índice de funcionamiento y las políticas de control, para después formular el Problema de Control Optimo.

En el Capítulo II, desarrollamos la teoría necesaria sobre los procesos de control de Markov con costos descontados.

En el Capítulo III, analizamos la optimalidad en el caso promedio bajo dos tipos de condiciones.

Las definiciones y resultados que se exponen en este trabajo, se ilustran con los siguientes ejemplos:

1) Sistema de Producción-Inventario,
2) Sistema lineal con costo cuadrático,
3) Sistema de colas,
4) Problemas invariantes.

CAPITULO I

PROCESOS DE CONTROL DE MARKOV

1.- Introducción.
2.- Modelo de control de Markov.
3.- Políticas de control.
4.- Espacio canónico de probabilidad.
5.- Problema de control óptimo.
6.- Ejemplos.

1.- Introducción.

El objetivo de la Teoría de Control Optimo, es determinar las acciones o decisiones de control que deben tomarse para optimizar el comportamiento de un sistema que evoluciona en el tiempo. Dicho comportamiento lo mide una función a la que llamaremos índice de funcionamiento y que en nuestro caso representará el costo de operación del sistema. El hecho de tratar sistemas dinámicos es lo que diferencía a los problemas de control óptimo de los problemas de optimización comunes.

Los problemas de control se pueden clasificar, dependiendo del fenómeno a modelar, en determinísticos o estocásticos (si influyen elementos aleatorios), y a tiempo continuo o discreto. Estos a su vez en problemas con horizonte finito o infinito, dependiendo de que el sistema sea observado durante un período finito o infinito de tiempo. Aquí trabajaremos con problemas con horizonte infinito.

Otra clasificación es con respecto al tipo de índice de funcionamiento. En nuestro caso analizaremos 2 tipos: caso descontado y caso promedio.

En este capítulo, definiremos los Procesos de Control de Markov a tiempo discreto (los cuales constituyen una clase de sistemas de control estocástico a tiempo discreto y cuya teoría se apoya en la técnica de Programación Dinámica Estocástica) y daremos los elementos suficientes para definir el Problema de Control Optimo.

Por último, en la sección 1.6, se darán una serie de ejemplos para ilustrar el problema de control óptimo, los cuales también nos servirán en las secciones posteriores.

2.- Modelo de control de Markov.

DEFINICION 2.1: Un modelo de control de Markov (MCM), estacionario a tiempo discreto, se compone de cuatro componentes, $(\mathbf{X},\mathbf{A},\mathbf{Q},\mathbf{C})$, donde:

- $\mathbf{X}$ es un espacio de Borel [Def. 1, Ap. A] no vacío llamado espacio de estados. A los elementos de $\mathbf{X}$ los llamaremos estados.
- $\mathbf{A}$ es un espacio de Borel no vacío llamado el conjunto de acciones o de control.

A cada $x\in\mathbf{X}$, le asociamos un conjunto no vacío $\mathbf{A}(x)\in\mathcal{B}(\mathbf{A})$, la σ-álgebra asociada al espacio $\mathbf{A}$. A $\mathbf{A}(x)$ lo llamaremos el conjunto de acciones admisibles en el estado x.

Supondremos que el conjunto:

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

de las parejas estado-control admisibles, es un subconjunto de Borel de $\mathbf{X}\times\mathbf{A}$.

- $\mathbf{Q}$ es un kernel estocástico [Def. 1, Ap. B] en $\mathbf{X}$ dado $\mathbb{K}$, al cual llamaremos ley de transición entre los estados.
- $\mathbf{C}$ es una función real definida en $\mathbb{K}$, medible, la cual representa el costo por etapa.

Como podemos observar, la relación $x \to A(x)$ es una multifunción de **X** a **A** [Def. 5, Ap. A]. Algunos autores especifican más la definición 2.1 incluyendo al conjunto $A(x)$: $(X, A, \{A(x),\ x \in X\}, Q, C)$.

El modelo de control (X, A, Q, C) representa un sistema estocástico controlado que se observa en los tiempos $t=0,1,\ldots$. La evolución del sistema por medio de este modelo de control es la siguiente. Sea x_t el estado del sistema al tiempo t y a_t la acción aplicada al tiempo t. Si $x_t = x \in X$ y $a_t = a \in A(x)$, entonces: 1) se incurre en un costo $C(a,x)$ y 2) el sistema avanza a un nuevo estado x_{t+1} de acuerdo a la probabilidad $Q(./x,a)$ en **X**, es decir,

$$Q(B/x,a) = \text{Prob}(x_{t+1} \in B / x_t = x, a_t = a), \quad B \in \mathcal{B}(X).$$

Estando ahora en el nuevo estado, digamos $x_{t+1} = x'$, se aplica un nuevo control $a' \in A(x')$ y se repite el proceso anterior.

Por lo regular, el estado inicial $x_0 \in X$ es un punto dado, o en otro caso, puede ser escogido de acuerdo a una distribución de probabilidad inicial ν dada:

$$\text{Prob}(x_0 \in B) = \nu(B), \quad B \in \mathcal{B}(X).$$

El período de tiempo hasta el cual será observado el sistema se le llama Horizonte o Tiempo Terminal, el cual es un entero no negativo que puede ser fijo o variable como en los problemas con horizonte rodante [17]. Aquí consideraremos problemas con horizonte infinito. El análisis y solución de estos problemas son de gran utilidad para resolver problemas con horizonte finito pero extremadamente grande.

En algunos problemas específicos como control de inventarios, sistema lineal con costo cuadrático, control de reservorios, etc. [ver secc. 6], la evolución del sistema es determinada por ecuaciones en diferencia de la forma:

$$x_{t+1} = F(x_t, a_t, \xi_t) \qquad t=0,1,2,\ldots.$$

donde $\{\xi_t\}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas (i.i.d.) tomando valores en algún espacio de Borel $\mathbf{S}$ y es llamada proceso de ruido o perturbación; F es una función dada, $F:\mathbf{X}\times\mathbf{A}\times\mathbf{S}\to\mathbf{X}$, medible, y el estado inicial x_0 es independiente de $\{\xi_t\}$.

Este tipo de problemas caben dentro del **MCM** $(\mathbf{X},\mathbf{A},\mathbf{Q},\mathbf{C})$. Para obtener la ley de transición $\mathbf{Q}$, sea μ la distribución común de ξ_t, entonces:

$$Q(B/x,a) = \mathrm{Prob}[F(x_t,a_t,\xi_t)\in B/x_t=x,a_t=a] = \mu\{s\in\mathbf{S}:F(x,a,s)\in B\}$$

$$= \int_S I_B[F(x,a,s)]\mu(ds) \qquad B\in\mathcal{B}(\mathbf{X}),\ (x,a)\in\mathbb{K} \tag{2.1}$$

donde $I_B(.)$ es la función indicadora del conjunto B.

Para fines teóricos y/o prácticos, se pide que el **MCM** $(\mathbf{X},\mathbf{A},\mathbf{Q},\mathbf{C})$ cumpla con cierto tipo de restricciones. Estas se imponen en cada una de sus componentes como veremos en los capítulos siguientes.

3.- Políticas de control.

En esta sección veremos la manera de como elegir o seleccionar los controles admisibles en cada etapa. Para esto, primero daremos la siguiente definición.

DEFINICION 3.1: Para cada $t=0,1,\ldots,$ el espacio $\mathbb{H}_t$ de historias admisibles hasta el tiempo t es:

$$\mathbb{H}_0:=\mathbf{X}$$

$$\mathbb{H}_t:=\mathbb{K}^t\times\mathbf{X}=\mathbb{K}\times\mathbb{H}_{t-1} \qquad t=1,2,\ldots$$

Un elemento $h_t\in\mathbb{H}_t$ es un vector o historia de la forma:

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

donde $(x_n, a_n) \in K$ para $n=0,1,\ldots,t-1$ y $x_t \in X$.

La definición general de una política de control es la siguiente:

DEFINICION 3.2: Una política (aleatorizada) es una sucesión $\pi=\{\pi_t\}$ de kérneles estocásticos π_t sobre A dado H_t, con la restricción:

$$\pi_t(A(x_t)/h_t)=1 \quad \forall\ h_t \in H_t,\ t=0,1,2,\ldots$$

Al conjunto de todas las políticas lo denotaremos por Π.

En este caso, la política elige el control con cierta probabilidad, i.e. en cada tiempo t, la acción de control es una variable aleatoria con distribución $\pi_t(./h_t)$. Otro caso es elegir los controles determinísticamente. La definición sería la siguiente:

DEFINICION 3.3: Una política $\pi = \{\pi_t\}$ es determinística si existe una sucesión $\{g_t\}$ de funciones medibles $g_t : H_t \to A$, tal que $\pi_t(./h_t)$ está concentrada en $g_t(h_t)$ para toda $h_t \in H_t$ y $t=0,1,\ldots$, i.e.,

$$\pi_t(B/h_t) := I_B[g_t(h_t)] \qquad B \in \mathcal{B}(A)$$

En este caso nos referiremos a π con la sucesión $\{g_t\}$.

En los casos de las definiciones 3.2 y 3.3, la política elige el control tomando en cuenta toda la historia anterior. Este tipo de políticas toman mayor importancia en los problemas de control adaptado [15].

Dentro de la familia de todas las políticas, existen las que no dependen de la historia anterior; estas son las llamadas políticas de Markov o políticas "sin memoria". Para dar su definición, denotemos por F al conjunto de todas las funciones medibles $f:X \to A$, tal que $f(x) \in A(x)$ para toda $x \in X$.

DEFINICION 3.4: Una política $\pi = \{\pi_t\}$ se dice que es:
(a) de Markov, si existe una sucesión $\{f_t\}$ de funciones $f_t \in F$, tal que $\pi_t(./h_t)$ esta concentrada en $f_t(x_t) \in A(x_t)$ para toda $h_t \in H_t$, $t=0,1,2,\ldots..$
(b) de Markov estacionaria, si existe una función $f \in F$ tal que $\pi_t(./h_t)$ esta concentrada en $f(x_t) \in A(x_t)$ para toda $h_t \in H_t$, $t=0,1,2,\ldots..$

En estos casos, identificaremos a π con la sucesión $\{f_t\}$ o la función f respectivamente.

De la misma forma, cuando usemos la política $\pi = \{f_t\}$, el control aplicado al tiempo t lo denotaremos por $a_t := f_t(x_t)$; el costo obtenido $C(x_t, f(x_t))$ por $C(x_t, f_t)$ y la ley de transición $Q(./x_t, f_t(x_t))$ por $Q(./x_t, f_t)$. Para el caso de una política estacionaria $f \in F$, a dichas expresiones la denotaremos por $C(x,f)$ y $Q(./x,f)$ cuando $x_t = x$.

4.- Espacio canónico de probabilidad.

Cada política $\pi = \{\pi_t\}$ genera un proceso estocástico $\{x_t\}$, $x_t \in X$ $\forall$ $t=0,1,\ldots.$ En esta sección, daremos el espacio canónico de probabilidad para el proceso, así como algunas de sus propiedades.

Sea $(\Omega, \mathcal{F})$ un espacio medible, donde $\Omega := (XA)^{\infty} = XAXA\ldots..$ y $\mathcal{F}$ es la σ-álgebra producto correspondiente. Un elemento $\omega \in \Omega$ es de la forma:

$$\omega = (x_0, a_0, x_1, a_1, \ldots.), \quad x_t \in X,\ a_t \in A \quad \forall\ t=0,1,\ldots.$$

donde las variables x_t y a_t son las proyecciones de Ω en los conjuntos X y A respectivamente.

Se puede ver que Ω contiene al espacio $H_\infty = KKK\ldots.$ de historias admisibles $(x_0, a_0, \ldots..)$ donde $(x_t, a_t) \in K$ $\forall$ t.

Sea $\pi = \{\pi_t\}$ una política de control arbitraria y supongamos que el estado inicial x_0 se elige de acuerdo a una distribución inicial ν, i.e., ν es una medida de probabilidad sobre X. Entonces, por el Teorema de C. Ionescu Tulcea [25], existe una única medida de probabilidad P_ν^π en $(\Omega,\mathcal{F})$ dada por:

$$P_\nu^\pi(dx_0,da_0,dx_1,da_1,\ldots)=\nu(dx_0)\pi_0(da_0/x_0)Q(dx_1/x_0,a_0)\cdot$$

$$\cdot\pi_1(da_1/x_0,a_0,x_1)\ldots\ldots\ldots$$

satisfaciendo:

4.1(a) $P_\nu^\pi(H\infty)=1$.

4.1(b) $P_\nu^\pi(x_0\in B)=\nu(B) \quad B\in\mathcal{B}(X)$.

4.1(c) $P_\nu^\pi(a_t\in C/h_t)=\pi_t(C/h_t) \quad C\in\mathcal{B}(A)$.

4.1(d) $P_\nu^\pi(x_{t+1}\in B/h_t,a_t)=Q(B/x_t,a_t) \quad B\in\mathcal{B}(X)$.

En particular si el estado inicial es $x_0=x$, entonces ν es una medida de probabilidad concentrada en $\{x\}$ y escribiremos P_ν^π como P_x^π. Con esto, $P_x^\pi(x_0=x)=1$.

Al operador esperanza con respecto a P_ν^π lo denotaremos por E_ν^π, y si $x_0=x$, por E_x^π.

Al proceso estocástico $(\Omega,\mathcal{F},P_\nu^\pi,\{x_t\})$ se le conoce como Proceso de Markov Controlado (PMC).

El nombre de PMC es por el hecho de que la propiedad 4.1(d) es esencialmente la propiedad de Markov. De hecho, el proceso $\{x_t\}$ es de Markov si utilizamos una política de Markov (Def. 3.4) como se prueba en la Proposición 4.1.

PROPOSICION 4.1: Si $\pi = \{f_t\}$ es una política de Markov, entonces $\{x_t\}$ es un proceso de Markov con kernel de transición $Q(./x,f_t)$, esto es:

$$P_\nu^\pi(x_{t+1}\in B/x_0,x_1,\ldots x_t) = P_\nu^\pi(x_{t+1}\in B/x_t) = Q(B/x_t,f_t) \qquad (4.2)$$

DEMOSTRACION:

Sea $\pi = \{\pi_t\}$ una política arbitraria. Por las propiedades B4 y B5 en el Apéndice B sobre esperanza condicional, tenemos que:

$$P_\nu^\pi(x_{t+1} \in B/h_t) = E_\nu^\pi\,[P_\nu^\pi(x_{t+1} \in B/h_t, a_t)/h_t]$$

$$= E_\nu^\pi\,[\,Q(B/x_t, a_t)/h_t] \qquad [\text{por } 4.1(d)]$$

$$= \int_A Q(B/x_t, a_t) P_\nu^\pi(da_t/h_t)$$

$$= \int_A Q(B/x_t, a_t) \pi_t(da_t/h_t) \qquad [\text{por } 4.1(c)]$$

Ahora, si $\pi = \{f_t\}$ es una política de Markov, la igualdad anterior quedaría:

$$P_\nu^\pi(x_{t+1} \in B/h_t) = Q(B/x_t, f_t). \tag{4.3}$$

Aplicando de nuevo las propiedades B4 y B5 (Ap. B), primero a la expresión del lado derecho de (4.2) y luego a la del lado izquierdo, obtenemos:

$$P_\nu^\pi(x_{t+1} \in B/x_0, \ldots, x_1) = E_\nu^\pi\,[P_\nu^\pi(x_{t+1} \in B/h_t)/x_0, \ldots, x_t]$$

$$= E_\nu^\pi\,[Q(B/x_t, f_t)/x_0, \ldots, x_t] \qquad [\text{por } (4.3)]$$

$$= Q(B/x_t, f_t), \qquad [\text{por B6, Ap. B}]$$

y además,

$$P_\nu^\pi(x_{t+1} \in B/x_t) = E_\nu^\pi\,[p_\nu^\pi(x_{t+1} \in B/h_t)/x_t]$$

$$= E_\nu^\pi\,[Q(B/x_t, f_t)/x_t] \qquad [\text{por } (4.3)]$$

$$= Q(B/x_t, f_t) \qquad [\text{por B6, Ap. B}]$$

Por lo tanto se cumple (4.2). |

En este resultado, si usamos una política estacionaria $f\in F$, entonces obtenemos que $\{x_t\}$ es un proceso de Markov a tiempo homogeneo con kernel de transición estacionario $Q(./x,f)$.

5.- Problema de control óptimo.

Con cada política $\pi\in\Pi$ obtenemos una sucesión de costos $\{C(x_t,a_t)\}$ que nos dá información sobre el comportamiento del sistema. Para formular el problema de control óptimo, necesitamos una función que mida dicho comportamiento. A esta función se le conoce como índice de funcionamiento y su forma depende de las características del problema y/o del criterio de optimalidad que se emplee. Por ejemplo:

(a) Costo Descontado (CD). Sea $0<\alpha<1$. El Costo Descontado total esperado por usar la política π con estado inicial $x_0=x$ es:

$$V_\alpha(\pi,x) := E_x^\pi \left[\sum_{t=0}^{\infty} \alpha^t C(x_t,a_t) \right]. \tag{5.1}$$

Al número α se le llama "factor de descuento".

(b) Costo Promedio (CP). El Costo Promedio total esperado por usar la política π con estado inicial $x_0=x$ es:

$$J(\pi,x) := \limsup_{n\to\infty} (n+1)^{-1} \sum_{t=0}^{n} E_x^\pi [C(x_t,a_t)]. \tag{5.2}$$

Estos son los índices de funcionamiento más usados en los casos donde el horizonte es infinito. En el caso cuando el horizonte es finito, por lo regular se usa el criterio de costo total esperado (CT), cuyo índice de funcionamiento toma la forma:

$$J_T(\pi,x) := E_x^{\pi}\left[\sum_{t=0}^{N} C(x_t,a_t)\right] \qquad (5.3)$$

Los índices que nos interesan en este trabajo son CD y CP. La diferencia entre estos, es que el primero mide el funcionamiento del sistema en el presente y en el futuro cercano, mientras que el segundo mide el funcionamiento promedio a largo plazo.

Con todos estos elementos podemos formular el problema de control óptimo (PCO) como sigue: Dado un modelo de control $(\mathbf{X},\mathbf{A},\mathbf{Q},\mathbf{C})$, una familia de políticas admisibles Π y un índice de funcionamiento, el PCO consiste en encontrar una política π^* tal que minimice dicho índice.

Por ejemplo, para el caso CD, nuestro problema sería encontrar una política $\pi^*\in\Pi$ tal que:

$$V_\alpha(\pi^*,x) \le V_\alpha(\pi,x) \quad \forall\ \pi\in\Pi,\ x\in X$$

En este caso decimos que π es α-descontada óptima (α-óptima).

Para el caso CP, el PCO sería encontrar una política $\pi^*\in\Pi$, tal que:

$$J(\pi^*,x) \le J(\pi,x) \quad \forall\ \pi\in\Pi$$

y decimos que π^* es óptima en costo promedio (CP-óptima).Similarmente sería para el caso CT.

A las funciones

$$V_\alpha(x) := \inf_\pi V_\alpha(\pi,x) = V_\alpha(\pi^*,x) \qquad (5.4)$$

y

$$J(x) := \inf_\pi J(\pi,x) = J(\pi^*,x) \qquad (5.5).$$

las llamaremos el costo α-descontado óptimo y el costo promedio óptimo respectivamente; también se les llama funciones de "valor".

Para que el problema de control no sea trivial, es natural pedir que $V_\alpha(\pi,x)<\infty$ para al menos una política $\pi\in\Pi$ y $x\in X$. En el caso de CP, puede suceder que el límite no exista. Es por esta razón que tomamos el límite superior en (5.2), aunque también se podría tomar el límite inferior. La diferencia es que con el limsup estamos minimizando el costo mayor, i.e., estaríamos usando un criterio pesimista; y con el liminf, minimizaríamos el costo menor, lo que sería usar un criterio optimista. Bajo ciertas condiciones, ambos límites coinciden, como por ejemplo en [26] y [11].

En este trabajo usaremos el lim sup por facilidad de trabajarlo matemáticamente.

Un resultado interesante [2, 39] dice que, en un gran número de problemas de control estocástico, para cualquier política $\pi\in\Pi$, existe una política de Markov π' tal que:

$$G(\pi',x) \le G(\pi,x), \quad x\in X,$$

donde G es el índice de funcionamiento. Esto lo podremos apreciar en los capítulos siguientes.

6.- Ejemplos.

En esta sección, daremos algunos ejemplos para ilustrar la teoría expuesta en las secciones anteriores. Recurriremos a estos ejemplos cuando tratemos la existencia de soluciones óptimas en los capítulos siguientes. Por el momento solo plantearemos el PCO en cada uno de ellos.

6.1.- Sistema de Producción-Inventario. [8, 23, 29].

Consideremos un sistema de Producción-Inventario, para el cual:

x_t, es la cantidad de artículos disponibles al principio del período t.

a_t, es la cantidad de artículos que se producen o se ordenan al principio del período t.

ξ_t, es la demanda (aleatoria) del artículo durante el período t.

Supondremos que las variables aleatorias ξ_t son i.i.d., acotadas e independientes de x_0, tomando valores en un conjunto $S=[0,D]$.

La dinámica del sistema es:

$$x_{t+1}=(x_t+a_t-\xi_t)^+, \quad t=0,1,\ldots. \tag{6.1}$$

Supondremos que la capacidad de almacenamiento es M, de modo que el espacio de estados y el de control son $\mathbf{X}=\mathbf{A}=[0,M]$. El conjunto de acciones admisibles es $\mathbf{A}(x)=[0,M-x]$ $\forall$ $x\in\mathbf{X}$.

Los costos que intervienen en la etapa t son, el costo por ordenar o producir el artículo, mas el costo que ocasionan los artículos sobrantes ($x_{t+1}>0$). De esta manera, el costo total en el período t será:

$$C(x_t,a_t,\xi_t) = ca_t + hx^+_{t+1}$$

donde c y h son los costos por artículo ordenado o producido y por exceso de artículos respectivamente ($c>0$, $h>0$).

Usando (6.1), los índices de funcionamiento para los casos CD y CP son:

$$V_\alpha(\pi,x) = E^\pi_x\left[\sum_{t=0}^{\infty}\alpha^t[ca_t + h(x_t+a_t-\xi_t)^+]\right]$$

$$J(\pi,x) = \limsup_{n\to\infty}(n+1)^{-1}\sum_{t=0}^{n}E^\pi_x\,[ca_t+h(x_t+a_t-\xi_t)^+]$$

Ahora supongamos que las variables aleatorias $\{\xi_t\}$ tienen distribución común μ. Por (2.1), el kernel estocástico Q tiene la forma:

$$Q(B/x,a) = \int_S I_B[(x+a-s)^+]\mu(ds) \qquad x\in\mathbf{X},\ a\in\mathbf{A}(x),\ B\in\mathcal{B}(\mathbf{X})$$

El problema de control óptimo es encontrar una política α-óptima y CP-óptima que minimice los índices correspondientes.

En estos problemas existen diferentes maneras de escoger el espacio de estados dependiendo de las hipótesis del problema. Por ejemplo, si se permite acumulamiento de demanda y la capacidad es M, el espacio de estados sería $(-\infty,M]$. En nuestro caso se puede mostrar que el problema de Producción-Inventario es equivalente a un problema de control de reservorios [38].

6.2.- Sistemas lineales con costo cuadratico (LQ). [8, 27].

Este tipo de sistemas se presenta en los casos donde se conoce de antemano la trayectoria ideal del proceso, y el problema es encontrar una política tal que la trayectoria generada por esta, diste lo menos posible a la trayectoria deseada.

Aquí supondremos que la trayectoria deseada es el origen. A este tipo de problemas se le conoce como Problema del Regulador Lineal.

La dinámica del sistema es lineal de la forma:

$$x_{t+1}=\gamma x_t+\beta a_t+\xi_t \qquad t=0,1,\ldots..$$

donde x_t, $a_t\in\mathbb{R}$ en el caso más sencillo y las perturbaciones $\{\xi_t\}$ son variables aleatorias i.i.d. tomando valores en un espacio S, con media cero y segundo momento finito $\sigma^2>0$ y que no dependen de x_t y a_t. Además, γ y β son constantes dadas tal que $\gamma\beta\neq0$.

De aquí, los espacios de estados, de control y de acciones admisibles son $\mathbf{X}=\mathbf{A}=\mathbf{A}(x)=\mathbb{R}\ \forall\ x\in\mathbf{X}$.

Suponiendo que μ es la distribución común de las ξ_t, la ley de transición es:

$$Q(B/x,a)=\int_S I_B[\gamma x+\beta a+s]\mu(ds) \qquad x\in X,\ a\in A(x),\ B\in \mathcal{B}(X).$$

El costo en estos problemas es cuadrático de la forma:

$$C(x_t,a_t) = qx_t^2 + ra_t^2 \tag{6.2}$$

donde q y r son constantes tales que $q\geq 0$ y $r>0$.

Los índices para los casos CD y CP los obtenemos sustituyendo el costo (6.2) en (5.1) y (5.2) respectivamente.

Si en el problema de control de reservorios nuestro objetivo es mantener el água cerca de un determinado nivel, podemos usar los sistemas LQ para modelarlo. En este caso, $\beta<0$, x_t sería el nivel de agua en el reservorio al tiempo t, a_t la cantidad de agua liberada al tiempo t y ξ_t el flujo de entrada (aleatorio) en el tiempo t.

El problema de Producción-Inventario y el LQ, pertenecen a la clase de problemas de control estocástico convexo, es decir, la correspondiente función de valor óptimo α-descontada $V_\alpha(x)$ (5.4) es convexa. Un tratamiento de este tipo de problemas se puede encontrar en [14, 21].

6.3.- Sistemas de colas con arribos aleatorios y un solo servidor. [31].

Consideremos un sistema de N colas con un servidor. En este ejemplo supondremos que las colas la forman cierto tipos de trabajos o "clientes" que tienen que esperar para ser procesados o atendidos por un servidor.

En cada etapa, se escoge cierta cantidad de trabajos para cada cola con una determinada distribución F_j, $j=1,2,\ldots,N$. Los estados serán vectores de dimensión N, $x=(x_1,\ldots,x_N)$, donde x_j es el número de trabajos en la cola j. De esta manera, el espacio de estados es $\mathbb{Z}_+^N$, donde $\mathbb{Z}_+=\{0,1,2,\ldots\}$.

Por otra parte, al principio de cada etapa el servidor decide si admite o no objetos en cada una de las colas, con la condición de que podrán ser atendidos hasta el principio de la próxima etapa. Además, el servidor debe decidir si permanece desocupado o atiende a una de las colas no vacías. Si esta última es la decisión, digamos, atender la cola k (no vacía), el servidor puede elegir una tasa de servicio $\beta(k)\in B_k$, donde B_k es un conjunto finito no vacío de tasas de servicio diferentes de cero.

Sea $e=(e_1,e_2,\ldots\ldots,e_N)$ tal que, para $j=1,2,\ldots.N$,

$$e_j = \begin{cases} 1 & \text{si se admiten objetos en la cola } j \\ 0 & \text{en otro caso} \end{cases}$$

y sea i la decisión de permanecer desocupado.

En el estado $x=(0,0,\ldots.0)$, i.e., cuando no existan trabajos esperando en las colas, las posibles decisiones son admitir trabajos en las colas permaneciendo desocupado el servidor. Esto último es por el hecho de que no se puede servir inmediatamente. A esta decisión la denotaremos por (e,i).

Para los estados distintos de cero, i.e., cuando hay trabajos en las colas, las posibles decisiones son (e,i) o admitir objetos en la cola sirviendo con una tasa $\beta(k)$. Esta última decisión la representaremos por $(e,\beta(k))$. Con esto el espacio de acciones sería:

$$A=\{0,1\}^N\times\left(\bigcup_j B_j \cup \{i\}\right)$$

El kernel estocástico Q, está dado en términos de las distribuciones F_j, de hecho sería la distribución conjunta de los estados.

La función de costo, es una función $C:X\times A\to\mathbb{R}$ cuya forma dependería del problema. Por ejemplo, puede incluir un costo de aceptación de trabajos o costos de servicios. De manera general, solo pediremos que para todo e, x, el costo $C(x,(e,i))$ sea creciente en cada coordenada del vector x con las otras fijas, y que para todo e, $\beta(k)$, x, con $x_k>0$, el costo $C(x,(e,\beta(k)))$ sea, similarmente, creciente en cada coordenada de x con las otras fijas.

Los problemas de teoría de colas tienen aplicaciones en redes de telecomunicaciones (e.g. redes telefónicas, de computación), problemas en ingeniería de tráfico y en investigación de operaciones.

6.4.- **Problemas invariantes.** [4, 35].

Un modelo de control markoviano se llama invariante si el kernel de transición no depende del estado actual del sistema, i.e., $Q(./x,a)=Q(./a)$. En nuestro caso pediremos que A sea finito. Algunos problemas que cumplen con esta propiedad son los de remplazamiento y mantenimiento. Aquí no especificaremos las componentes del MCM, ya que estas dependen del problema que se esté trabajando. Algunos ejemplos se pueden encontrar en [4].

CAPITULO II

PROCESOS DE CONTROL DE MARKOV CON COSTOS DESCONTADOS

1.- Introducción.

2.- Hipótesis sobre el modelo de control.

3.- Ecuación de Optimalidad.

1.- Introducción.

De los problemas de control óptimo estocástico con horizonte infinito, el caso con costos descontados (CD) es el más simple de trabajar. Primeramente, si los costos por etapa, $C(x,a)$, son acotados, el índice de funcionamiento [ver (5.1) Cap. I]:

$$V_{\alpha}(\pi,x):=E_x^{\pi}\left[\sum_{t=0}^{\infty}\alpha^t C(x_t,a_t)\right] \tag{1.1}$$

es finito, y por lo tanto también la función de valor óptimo [ver (5.4) Cap. I]:

$$V_{\alpha}(x):=\inf_{\pi} V_{\alpha}(\pi,x) \qquad x\in X \tag{1.2}$$

Esto nos permite apoyarnos en el teorema de punto fijo de Banach, para asegurar la unicidad de la solución de la ecuación de optimalidad o ecuación de Bellman.

En nuestro caso, donde los costos no son acotados, no podemos asegurar lo anterior. Aquí, tenemos que hacer uso de otros artificios para poder caracterizar una solución de la ecuación de optimalidad.

En ambos casos podemos hacer iteración de valores para aproximar a una solución de la ecuación de optimalidad, cosa que no ocurre, en general, con los problemas con costo promedio. Lo anterior es una gran ventaja porque se pueden implementar métodos computacionales para resolver el problema.

Este tipo de índices de funcionamiento tiene, por lo regular, aplicaciones a problemas [e.g. 6, 7] en los que el índice V_α tiene una interpretación económica (o monetaria). En tales casos, se introduce un factor de descuento al costo, por el hecho de que cierta cantidad de dinero en el presente vale menos en el futuro. En muchos problemas el factor de descuento α se interpreta como $\alpha=1/(1+i)$, donde i es la tasa de interes. Así, α^t representa el valor presente de la moneda t períodos después.

La sección principal de este capítulo es la 3, en la cual se da un resultado (Teo. 3.6) que caracteriza a la función de valor óptimo dentro del conjunto de soluciones de la ecuación de optimalidad.

En la sección 2 se darán ciertas hipótesis, junto con una série de resultados que nos garantizarán la existencia de minimizadores. Estas hipótesis las ilustraremos con los problemas expuestos en el capítulo anterior y con ciertos contraejemplos.

2.- Hipótesis sobre el modelo de control.

Sea **(X,A,Q,C)** un modelo de control de Markov [Def. 2.1,I]. En esta sección, estableceremos una serie de hipótesis para el modelo de control, las cuales nos garantizarán resultados interesantes, como la existencia de soluciones óptimas para el PCO. Además, veremos cuales de los ejemplos expuestos en el capítulo anterior cumplen con estas. Para esto, primero daremos la siguiente definición.

DEFINICION 2.1: Una función de valor real definida sobre $\mathbb{K}$, $v:\mathbb{K}\to\mathbb{R}$, se dice que es inf-compacta en $\mathbb{K}$, si el conjunto:

$$\{a\in A(x) \ : \ v(x,a)\leq r\}$$

es compacto $\forall$ $x\in X$ y $r\in\mathbb{R}$.

Denotaremos por $L(X)$ a la clase de funciones en X que son semi-continuas inferiormente (sci) y acotadas por abajo; y con $L(X)_+$ a las funciones no negativas en $L(X)$ [ver Def.2, Ap. A].

Hipótesis 2.2:
(a) $C(x,a)$ es inf-compacta y pertenece a $L(\mathbb{K})_+$
(b) La ley de transición Q es debilmente continua [Def.2, Ap. B], i.e., para cualquier función continua y acotada v en X, la función:

$$v'(x,a) := \int v(y)Q(dy/x,a)$$

es continua en $\mathbb{K}$.
(c) La multifunción $x\to A(x)$ es sci [Def.6, Ap. A], i.e., si $x_n\to x$ en X y $a\in A(x)$, entonces existe $a_n\in A(x_n)$ tal que $a_n\to a$.
(c') La multifunción $x\to A^*(x)$ es sci, donde

$$A^*(x):= \{a\in A(x) \ : \ v(x,a)= \inf_{a\in A(x)} v(x,a)\}$$

para $v\in L(\mathbb{K})$.

Comentario: Por la Proposición 3(b) Ap. B, la hipótesis 2.2(b) es equivalente a pedir que para cualquier $v\in L(X)$ la función $v'(x,a)$ es sci y acotada por abajo en $\mathbb{K}$.

Las hipótesis (c) y (c') nos servirán para probar que la función de valor $V_\alpha(x)$ sea sci y las podemos sustituir pidiendo que la multifunción $x \to A(x)$ sea semicontinua superiormente [Def. 7, Ap. A].

De los ejemplos vistos en el Capitulo I, es claro que la hipótesis 2.2(a) la cumplen el problema de Producción-Inventario y el LQ. Para verificar las hipótesis restantes en estos problemas, daremos los siguientes Lemas.

LEMA 2.3: Sea $\{x_t\}$ un proceso de Markov en X definido por:

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

donde $\{\xi_t\}$ son variables aleatorias i.i.d. con valores en un conjunto S, distribución común μ e independientes del estado inicial x_0. Entonces, si $F(x,s)$ es continua en $x \in X$ para todo $s \in S$, la ley de transición de $\{x_t\}$ es debilmente continua, i.e.,

$$u'(x) = \int u(y)Q(dy/x)$$

es continua y acotada sobre X para toda u continua y acotada.

DEMOSTRACION:

Sea u una función continua y acotada sobre X. Claramente la función $u'(x)$ es acotada. Para demostrar su continuidad, sea $\{x^n\}$ una sucesión en X tal que $x^n \to x$, $x \in X$. Entonces por la continuidad de las funciones F y u, y por el teorema de convergencia acotada tenemos que:

$$u'(x^n) = \int u[F(x^n,s)]\mu(ds) \to u'(x) = \int u[F(x,s)]\mu(ds)$$

lo cual demuestra que $u'(x)$ es continua. |

LEMA 2.4: Si el conjunto K es convexo, entonces, la multifunción $x \to A(x)$ es sci.

DEMOSTRACION:

Sean $k=(x,a)$ y $k'=(x',a')$ elementos de $\mathbb{K}$. Como $\mathbb{K}$ es convexo, el segmento de recta ℓ que una a los puntos k y k' está completamente contenido en $\mathbb{K}$. Ahora, sea $k_1=(x_1,a_1)$ el punto medio del segmento kk', entonces $k'\in\ell\subset\mathbb{K}$. Similarmente, sea $k_2=(x_2,a_2)$ el punto medio de k_1k', entonces $k_2\in\ell\subset\mathbb{K}$. Siguiendo con este procedimiento, obtenemos una sucesión $k_n=(x_n,a_n)$, tal que k_n es el punto medio de $k_{n-1}k'$ y $k_n\in\ell\subset\mathbb{K}$. Se sabe de los cursos de análisis real que la sucesión $\{(x_n,a_n)\}\subset\mathbb{K}$ converge al punto $(x',a')\in\mathbb{K}$. De esta manera, si $x_n\to x$ en X y $a\in A(x)$, existe $a_n\to a$ con lo cual se demuestra que la multifunción $x\to A(x)$ es sci. |

Volviendo a los ejemplos, en el problema de Producción -Inventario tenemos que $x_{t+1}=F(x_t,a_t,\xi_t)=x_t+a_t-\xi_t$. Es claro que F es continua en x_t, a_t para toda t, y por el Lema 2.3 la hipótesis 2.2(b) se cumple. Además, como $A(x)=[0,M-x]$ y $X=[0,M]$, el conjunto $\mathbb{K}$ es convexo, y así por el Lema 2.4, se cumple la hipótesis 2.2(c).

Similarmente para el sistema LQ tenemos que la función $F(x_t,a_t,\xi_t)=\gamma x_t+\beta a_t+\xi_t$ es continua en x_t, a_t para toda t y además como $X=A(x)=\mathbb{R}$, el conjunto $\mathbb{K}$ es convexo. Por lo tanto, por los Lemas 2.3 y 2.4, las hipótesis 2.2(b) y 2.2(c) se cumplen respectivamente.

En el ejemplo de sistema de colas, primero notemos que X y A son numerables, donde $A=A(x)$ para toda x. De aquí, la multifunción $x\to A(x)$ y la función $v'(x,a)$ en 2.2(b) son continuas, y por tanto las hipótesis 2.2(b) y 2.2(c) se cumplen. La verificación de la hipótesis 2.2(a) dependerá del problema específico.

Para los problemas invariantes, es claro que las hipótesis 2.2(b) y 2.2(c) se cumplen, ya que el conjunto $A(x)$ es finito, y de la misma manera que el ejemplo anterior, la hipótesis 2.2(a) dependerá del problema que se esté trabajando.

Las hipótesis 2.2 tienen consecuencias importantes en el garantizar la existencia de selectores medibles [Def.3.4,I]. Los resultados que establecen la existencia de dichos selectores, son llamados Teoremas de Selección Medible como el que enunciamos a continuación.

TEOREMA 2.5: (Teorema de Selección Medible). Si las hipótesis 2.2(c) y 2.2(c') se cumplen y v es una función inf-compacta que pertenece a L(K), entonces, la función:

$$v^*(x) := \inf_{a\in A(x)} v(x,a)$$

pertenece a L(K) y además existe una función f∈F tal que:

$$v^*(x) = v(x,f(x)). \tag{2.1}$$

DEMOSTRACION:

La existencia de f∈F que satisface (2.1) se sigue de [30] Corolario 4.3.

Ahora, sea $x_n \to x$ en X y $a^* \in A^*(x)$ arbitrario. Entonces $v^*(x)=v(x,a^*)$. Por 2.2(c'), existe $a_n \in A^*(x_n)$ tal que $a_n \to a^*$. De aquí, $(x_n,a_n)\to(x,a^*)$ y como v es sci, por la Prop. 3, Ap.A,

$$\liminf_n v^*(x_n)=\liminf_n v(x_n,a_n)\geq v(x,a^*)=v^*(x)$$

y nuevamente por la Prop. 3, Ap. A, v^* es sci.|

Una consecuencia de este teorema son los siguientes resultados, los cuales nos serán de gran utilidad en la próxima sección.

LEMA 2.6: (a) Si las hipótesis 2.2 se cumplen y $u\in L(X)_+$, entonces, la función no negativa

$$u^*(x) := \inf_{a\in A(x)} \left[C(x,a) + \int u(y)Q(dy/x,a)\right]$$

pertenece a L(X) y además existe $f\in F$ tal que:

$$u^*(x) := C(x,f) + \int u(y)Q(dy/x,f) \qquad \forall\ x\in X$$

(b) Para cada n=0,1,..., sea $v_n\in L(K)$ e inf-compacta en K. Si $v_n \nearrow v_0$, entonces:

$$\lim_n \inf_{a\in A(x)} v_n(x,a) = \inf_{a\in A(x)} v_0(x,a) \qquad \forall\ x\in X$$

DEMOSTRACION:

(a) Sea $u\in L(X)_+$. Por las hipótesis 2.2(a), Prop. 4, Ap. A y Prop. 3, Ap. B, tenemos que la función $v(x,a)\in L(K)$, donde,

$$v(x,a) := C(x,a) + \int u(y)Q(dy/x,a)$$

De aquí, $\{a\in A(x) : v(x,a)\le r\}$ es cerrado en K para todo $x\in X$ y además está contenido en el conjunto $\{a\in A(x) : C(x,a)\le r\}$. Así, por la hipótesis 2.2(a) el conjunto $\{a\in A(x) : v(x,a)\le r\}$ es compacto y por lo tanto $v(x,a)$ es inf-compacta.

Por el Teorema 2.5, $u^*(x):=\inf_{a\in A(x)} v(x,a)$ está en $L(X)_+\subset L(X)$ y además, existe $f\in F$ tal que $u^*(x)=v(x,f)$, lo cual prueba lo que se quería.

(b) Para todo $x\in X$, sea

$$\ell(x):= \lim_n \inf_{a\in A(x)} v_n(x,a) \quad \text{y} \quad v_0^*(x):= \inf_{a\in A(x)} v_0(x,a)$$

Por demostrar que $\ell(x)=v_0^*(x)$. Primeramente, como $v_n\uparrow v_0$, entonces:

$$\lim_n \inf_{a\in A(x)} v_n(x,a) \le \inf_{a\in A(x)} v_0(x,a)$$

esto es,

$$\ell(x) \le v_0^*(x) \tag{2.2}$$

Para probar la desigualdad inversa, sea $x\in X$ fijo pero arbitrario. Para cada n, definamos los conjuntos:

$$A_n := \{a\in A(x) : v_n(x,a)\le v_0^*(x)\}$$

Como las funciones v_n son inf-compactas, los conjuntos A_n son compactos, y además, como $v_n \nearrow v_0$, tenemos que $A_n \searrow A_0$, y

$$\inf_{a\in A(x)} v_n(x,a) \le \inf_{a\in A(x)} v_0(x,a) = v_0^*(x)$$

Aplicando el Teorema 2.5 a v_n podemos asegurar que para cada $n\ge 1$, existe $a_n\in A_n$ tal que:

$$v_n(x,a_n) = \inf_a v_n(x,a) \tag{2.3}$$

Por otra parte, por la compacidad de los conjuntos A_n, existe una subsucesión $\{a(n)_i\}$ de $\{a_n\}$ tal que $a(n)_i \to a_0$, $a_0\in A_0$. Como v_n es monotona creciente, se cumple que:

$$v_{n_i}(x,a(n)_i) \ge v_n(x,a(n)_i) \qquad \forall\ n_i>n$$

para cualquier $n\ge 1$. Por la Prop, 3, Ap. A, obtenemos,

$$\liminf_i v_{n_i}(x,a(n)_i) \ge v_n(x,a_0)$$

y por (2.3),

$$\ell(x) \ge \liminf_i \inf_a v_{n_i}(x,a) \ge v_n(x,a_0)$$

Por último, haciendo $n\to\infty$ en esta expresión,

$$\ell(x) \ge v_0(x,a_0) = v_0^*(x) \tag{2.4}$$

Esta última igualdad es por el hecho de que $v_0(x,a_0)=v_0^*(x)$ para $a_0\in A_0$. De (2.2) y (2.4) se prueba lo que se quería.|

En trabajos anteriores al [22] (e.g. 15, 24, 34, etc.) se utilizaban hipótesis más fuertes que las hipótesis 2.2. Por ejemplo, se pedía que los conjuntos $A(x)$ fueran compactos. Notemos que si $v(x,.)$ es sci en $A(x)$ para todo $x\in X$ y además los conjuntos $A(x)$ son compactos, entonces v es inf-compacta en K. La implicación inversa no es válida, solo se puede asegurar que si v es inf-compacta en K, entonces v es sci. Con esto se puede ver la debilidad de las hipótesis 2.2.. Además, si los costos $C(x,a)$ son acotados, entonces la hipótesis de pedir que $C(x,a)$ sea inf-compacta, se reduce a pedir que los conjuntos $A(x)$ sean compactos.

En el capítulo siguiente, analizaremos el caso donde el kernel estocástico Q es fuertemente continuo [Def.2, Ap.B].

Por otro lado, pedir como hipótesis que la multifunción $x\to A(x)$ y $x\to A^*(x)$ sean sci, nos permite probar que las funciones v^* y u^*, en el Teorema 2.5 y en el Lema 2.6 respectivamente, son sci. Esto nos servirá para probar que la función de valor óptimo (1.2) pertenezca a $L(X)$. En [20] se piden las mismas hipótesis que aquí, excepto la 2.2(c) y 2.2(c'). Con ellas, solo se puede asegurar que dichas funciones son medibles.

3.- Ecuación de Optimalidad α-descontada.

En esta sección, trataremos el problema de la existencia de políticas α-óptimas, así como la existencia de soluciones de la ecuación de optimalidad, bajo las hipótesis 2.2. Paralelamente, iremos comparando estos resultados con el caso donde los costos sean acotados.

DEFINICION 3.1: Decimos que una función real $v\in L(X)_+$ es una solución de la Ecuación de Optimalidad (EO) si:

$$v(x)= \min_{a\in A(x)} \left[C(x,a) + \alpha\int v(y)Q(dy/x,a) \right], \quad x\in X. \tag{3.1}$$

En vista de las hipótesis 2.2(a), puede darse el caso en que para algún estado inicial x y una política π, el índice de funcionamiento $V_\alpha(\pi,x)$ sea infinito como lo muestra el siguiente ejemplo.

EJEMPLO 3.2: Consideremos un problema de control donde la dinámica del sistema es:

$$x_{t+1}=\beta x_t+u_t \qquad t=0,1,\ldots.$$

con $\mathbf{X}=\mathbf{A}=\mathbb{R}$, $\mathbf{A}(x)=[-1,1]$ y $C(x,a)=|x|$. De los Lemas 2.3 y 2.4, es claro que las hipótesis 2.2 se satisfacen. Ahora, si $x_0=x$ y consideramos la política $\bar{\pi}=\{0,0,\ldots,0\}$, es fácil ver que $x_t=\beta^t|x|$. De aquí,

$$V_\alpha(\bar{\pi},x) = \sum_{t=0}^{\infty} \alpha^t\beta^t|x|$$

y por lo tanto, $V_\alpha(\bar{\pi},x)=\infty$ si $x\neq 0$ y $|\alpha\beta|\geq 1$. $\blacksquare$

En vista del ejemplo anterior, es necesario que se cumpla la siguiente hipótesis:

Hipótesis 3.3: Existe una política $\hat{\pi}$ tal que $V_\alpha(\hat{\pi},x)<+\infty$ para todo $x\in\mathbf{X}$.

En consecuencia, $V_\alpha(x)<+\infty$ $\forall$ $x\in\mathbf{X}$.

Una manera conveniente para trabajar la EO (3.1), es por medio de operadores. Definimos el operador $T:L(\mathbf{X})_+\to L(\mathbf{X})_+$ como sigue: si $u\in L(\mathbf{X})_+$,

$$Tu(x) := \min_{a\in A(x)}\left[C(x,a) + \alpha\int u(y)Q(dy/x,a) \right] \qquad x\in\mathbf{X}.$$

Así, diremos que $v\in L(\mathbf{X})_+$ es una solución de la EO si es punto fijo de T, i.e. (cf. (3.1)),

$$Tv=v. \tag{3.2}$$

Es importante notar que el operador T es no decreciente en $L(X)_+$, esto es:

$$u \geq v \Rightarrow Tu \geq Tv \qquad u,\ v \in L(X)_+. \tag{3.3}$$

Diferentes autores (e.g. [8], [9], [15]) demuestran que si el costo por estapa $C(x,a)$ es acotado, entonces la única función acotada que satisface la EO (3.1) o (3.2) es la función de valor óptimo (1.2). En nuestro caso no podemos asegurar ésto, como lo muestra el siguiente ejemplo:

EJEMPLO 3.4: Sea $X=[0,\infty)$, $A=A(x)=\mathbb{R}$, $C(x,a)=0$ en $L(X)_+$, y la dinámica del sistema $x_{t+1}=x_t/\alpha$. Nuevamente las hipótesis 2.2 se satisfacen. Ahora, para cada $\beta \in \mathbb{R}$, la función $v(x)=\beta x$, $x \in X$, es una solución de la EO, con lo cual tenemos que el operador T tiene una infinidad de puntos fijos. Sin embargo, la única solución acotada es $v(x)=0$, la cual es la función de valor óptimo. En el Teorema 3.6, mostraremos que la función de valor óptimo $V_\alpha(x)$ es la mínima solución de la EO dentro de la familia $L(X)_+$. |

Antes de presentar el Teorema principal de esta sección, daremos los siguientes Lemas.

LEMA 3.5: Supongase que las hipótesis 2.2 se cumplen.
(a) Si $v \in L(X)_+$ satisface que $v \geq Tv$, entonces $v \geq V^*$.
(b) Si v es una función medible en X tal que Tv está bién definida y satisface que $v \leq Tv$ y

$$\lim_n \alpha^n E_x^\pi v(x_n)=0 \tag{3.4}$$

para toda $\pi \in \Pi$ y $x \in X$, entonces $v \leq V^*$.

DEMOSTRACION:
(a) Sea $v \in L(X)_+$. Por el Lema 2.6(a), existe $f \in \mathbb{F}$ tal que:

$$Tv(x)=\min_{a}\left[C(x,a)+\alpha\int v(y)Q(dy/x,a)\right]=C(x,f)+\alpha\int v(y)Q(dy/x,f) \quad x\in X.$$

Por hipótesis,

$$v(x) \geq C(x,f)+\alpha\int v(y)Q(dy/x,f) \quad \forall\ x\in X. \qquad (3.5)$$

Además,

$$v(x) \geq C(x,f)+\alpha\int v(y)Q(dy/x,f) = E_x^f\ [C(x_0,f)]+\alpha E_x^f[v(x_1)]$$

Por (3.5) tenemos que:

$$\alpha E_x^f\ [v(x_1)] \geq \alpha E_x^f\ [C(x_1,f)] + \alpha^2 E_x^f\ [v(x_2)]$$

y así,

$$v(x) \geq E_x^f\ [C(x_0,f)] + \alpha E_x^f[C(x_1,f)] + \alpha^2 E_x^f[v(x_2)]$$

$$= E_x^f \sum_{t=0}^{1} \alpha^t C(x_t,f) + \alpha^2 E_x^f v(x_2)$$

donde $E_x^f\ [v(x_2)]=\int v(y)Q^2(dy/x,f)$ y $Q^2(B/x,f)=P_x^f(x_2\in B)$ es la probabilidad de transición en 2 pasos de la cadena de Markov $\{x_t\}$ [ver Prop. 4.1,I].

Siguiendo con la iteración de (3.5), obtenemos:

$$v(x) \geq E_x^f \sum_{t=0}^{n-1} \alpha^t C(x_t,f) + \alpha^n E_x^f v(x_n) \quad \forall\ n,\ x\in X \qquad (3.6)$$

y además, $E_x^f\ [v(x_n)]=\int v(y)Q^n(dy/x,f)$ y $Q^n(B/x,f)=P_x^f(x_n\in B)$ es la probabilidad de transición en n pasos de la cadena de Markov $\{x_t\}$.

Por otro lado, por la no negatividad de v, tenemos que:

$$v(x) \geq E_x^f \sum_{t=0}^{n-1} \alpha^t C(x_t, f)$$

y haciendo $n\to\infty$ en esta expresión,

$$v(x) \geq E_x^f \sum_{t=0}^{\infty} \alpha^t C(x_t, f) = V_\alpha(\pi, x) \geq V_\alpha(x) \qquad \forall\ x\in X$$

lo cual prueba lo que se quería. |

(b) Sean $\pi\in\Pi$ y $x\in X$ arbitrarios. Por definición tenemos que:

$$E_x^\pi [\ \alpha^{t+1} v(x_{t+1})/h_t, a_t] = \alpha^{t+1}\int v(y) P_x^\pi[dy/h_t, a_t]$$

$$=\alpha^{t+1}\int v(y) Q(dy/x_t, a_t) \qquad \text{[por 4.1(d), I]}$$

$$=\alpha^t[C(x_t, a_t)+\alpha\int v(y) Q(dy/x_t, a_t) - C(x_t, a_t)]$$

Aparte, por hipótesis tenemos que:

$$v(x_t) \leq Tv(x_t) \leq C(x_t, a_t)+\alpha\int v(y) Q(dy/x_t, a_t) \qquad \forall\ t$$

Así,

$$E_x^\pi [\ \alpha^{t+1} v(x_{t+1})/h_t, a_t] \geq \alpha^t[v(x_t) - C(x_t, a_t)]$$

y de aquí,

$$\alpha^t C(x_t, a_t) \geq \alpha^t v(x_t) - E_x^\pi [\ \alpha^{t+1} v(x_{t+1})/h_t, a_t]$$

$$=-E_x^\pi [\ \alpha^{t+1} v(x_{t+1}) - \alpha^t v(x_t)/h_t, a_t] \qquad \text{[por B6, Ap.B]}$$

$$=E_x^\pi [\ \alpha^t v(x_t) - \alpha^{t+1} v(x_{t+1})/h_t, a_t]$$

Tomando esperanza $E_x^\pi(.)$ en ambos lados de la desigualdad, y usando B7, Ap.B, obtenemos:

$$\alpha^t E_x^\pi[C(x_t,a_t)] \geq E_x^\pi \left[\alpha^t v(x_t) - \alpha^{t+1} v(x_{t+1})/h_t, a_t \right]$$

Ahora, sumando sobre $t=0,1,\ldots,n-1$,

$$\sum_{t=0}^{n-1} \alpha^t E_x^\pi C(x_t,a_t) \geq v(x) - \alpha^n E_x^\pi v(x_n) \qquad \forall\, n$$

y haciendo $n\to\infty$ junto con (3.4), obtenemos que $V_\alpha(\pi,x)\geq v(x)$. Por último, como π y x se tomaron arbitrarios, esta última desigualdad implica que $V_\alpha \geq v$ con lo que se demuestra el Teorema. |

Definimos recursivamente las funciones de Iteración de Valores (IV) como $v_0(.):=0$ y $v_n := Tv_{n-1}=T^n v_0$ para $n=1,2,\ldots$ De aquí,

$$v_n(x) = \min_a \left[C(x,a) + \alpha \int v_{n-1}(y) Q(dy/x,a) \right], \quad n\geq 1. \tag{3.7}$$

Notemos que por el Lema 2.6(a), la función $v_n \in L(X)_+$ para toda $n\geq 0$. Por otro lado, sea

$$Vn(\pi,x) := E_x^\pi \sum_{t=0}^{n-1} \alpha^t c(x_t,a_t) \tag{3.8}$$

el índice de funcionamiento de un problema con n etapas. De la teoría de Programación Dinámica [8, 13], se sabe que $v_n(x)$ en (3.7) es la función de valor óptimo para este problema con costo terminal $v_0(.)=0$ cuando $x_0=x$, esto es:

$$v_n(x) = \inf_\pi V_n(\pi,x). \tag{3.9}$$

TEOREMA 3.6: Supongamos que las hipótesis 2.2 y 3.3 se cumplen. Entonces:

(a) $v_n \nearrow V_\alpha$

(b) V_α es la mínima función en $L(X)_+$ que satisface la EO, i.e.,

$$V_\alpha = TV_\alpha \tag{3.10}$$

(c) Existe una política estacionaria $f^* \in F$ tal que $f^*(x) \in A(x)$ minimiza el lado derecho de (3.10) para toda $x \in X$, i.e.,

$$V_\alpha(x) = C(x, f^*) + \alpha \int V_\alpha(y) Q(dy/x, f^*) \tag{3.11}$$

y f^* es α-óptima. Recíprocamente, si $f^* \in F$ es una política óptima estacionaria, entonces satisface (3.11).

(d) Si π^* es una política tal que $V(\pi^*, .) \in L(X)_+$ y satisface la EO junto con la condición:

$$\lim_n \alpha^n E_x^\pi V_\alpha(\pi^*, x_n) = 0 \quad \forall\ \pi \in \Pi,\ x \in X \tag{3.12}$$

entonces, $V_\alpha(\pi^*, .) = V_\alpha(.)$; de aquí, π^* es α-óptima.

DEMOSTRACION:

(a) Primeramente, por (3.3), las funciones (IV) en (3.7) forman una sucesión $\{v_n\}$ no decreciente en $L(X)_+$ y por lo tanto existe una función $u \in L(X)_+$ tal que $v_n \nearrow u$. Por el Teorema de Convergencia Monótona ([3], [33]) tenemos que:

$$C(x,a) + \alpha \int v_{n-1}(y) Q(dy/x,a) \nearrow C(x,a) + \alpha \int u(y) Q(dy/x,a)$$

Ahora, por el Lema 2.6(b),

$$\lim_n \min_{a \in A(x)} \left\{ C(x,a) + \alpha \int v_{n-1}(y) Q(dy/x,a) \right\}$$

$$= \min_{a \in A(x)} \left\{ C(x,a) + \alpha \int u(y) Q(dy/x,a) \right\}$$

Esto es,

$$u(x) = \lim_n v_n(x) = Tu(x) \tag{3.13}$$

De aquí, $u \in L(X)_+$ satisface la EO. En particular tenemos que $u \geq Tu$, así, por el Lema 3.5(a),

$$u \geq V_\alpha \tag{3.14}$$

Para probar la desigualdad inversa, de (3.8) y (3.9) tenemos que:

$$v_n(x) \leq V_n(\pi,x) \leq V(\pi,x) \quad \forall\ n,\ \pi,\ x$$

Haciendo $n \to \infty$, y como $v_n \nearrow u$, obtenemos:

$$u(x) \leq V_\alpha(\pi,x) \quad \forall\ \pi,\ x$$

Luego,

$$u(x) \leq V_\alpha(x) = \inf_\pi V_\alpha(\pi,x) \tag{3.15}$$

De (3.14) y (3.15) se concluye que $u(x) = V_\alpha(x)$ y por lo tanto, $v_n \nearrow V_\alpha$.

(b) De (3.13) tenemos que $V_\alpha = TV_\alpha$, i.e., V_α satisface la EO. Solo falta demostrar que es la mínima solución de (3.2). Para esto, sea $u' \in L(X)_+$ tal que $u' = Tu'$. Entonces, por el Lema 3.5(a), $u' \geq V_\alpha$ lo cual prueba lo que se quería.

(c) Por el Lema 2.6(a), existe $f^* \in F$ que satisface (3.11), i.e.,

$$V_\alpha(x) = C(x,f^*) + \alpha \int V_\alpha(y) Q(dy/x,f^*)$$

Iterando esta igualdad [ver (3.6)], obtenemos:

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

$$\geq E_x^{f^*}\left[\sum_{t=0}^{n-1} \alpha^t C(x_t, f^*)\right]$$

Esta última desigualded es por el hecho de que $V_\alpha(.)\geq 0$. Haciendo $n\to\infty$, obtenemos que:

$$V_\alpha(x) \geq V(f^*, x)$$

Por (1.2), $V_\alpha(.)=V(f^*,.)$, esto es f^* es α-óptima.

Para el resultado recíproco, primero notemos que si $f\in\mathbb{F}$, por la propiedad de Markov, tenemos:

$$V_\alpha(f,x) = E_x^{f}\left[\sum_{t=0}^{\infty} \alpha^t C(x_t, f\ )\right] = C(x,f) + \alpha\int V_\alpha(y) Q(dy/x,f)$$

En particular, si $f^*\in\mathbb{F}$ es α-óptima, entonces $V_\alpha(f^*,x)=V_\alpha(x)$ y por lo tanto se cumple (3.11).

(d) Sea π^* una política tal que $V_\alpha(\pi^*,x) \in L(X)_+$ satisface la EO, i.e., $V_\alpha(\pi^*,x)=TV_\alpha(\pi^*,x)$. En particular, $V_\alpha(\pi^*,x)\leq TV_\alpha(\pi^*,x)$. Por (3.12) y el Lema 3.5(b), tenemos que $V_\alpha(\pi^*,x)\leq V_\alpha(x)$ y por la parte (b) de este Teorema, $V_\alpha(\pi^*,x)=V_\alpha(x)$ y así, π^* es α-óptima. |

Es importante notar que el Teorema 3.6(a), nos da una manera de aproximar a V_α recursivamente por medio de problemas con un número finito de etapas. Esto es de gran importancia teórica (como las conclusiones del Teorema 3.6(b)) y computacional. Otros tipos de aproximación se pueden ver, por ejemplo en [22].

La convergencia $v_n \nearrow V_\alpha$ puede no ser válida si no se cumple la hipótesis 3.3, como lo muestra el siguiente ejemplo:

EJEMPLO 3.7: Sea $X=[0,\infty)$, $A=A(x)=(0,\infty)$ y la dinámica del sistema es $x_{t+1}=(2/\alpha)x_t+a_t$, con función de costo por etapa $C(x_t,a_t)=x_t+a_t$. Este problema cumple con las hipótesis 2.2. Por otro lado, es fácil probar, recursivamente, que si $x_0=x$, entonces:

$$x_t = (2/\alpha)^t x + \sum_{n=1}^{t} (2/\alpha)^{t-n} a_{n-1} \quad t=1,2,\ldots$$

Así,

$$V_\alpha(\pi,x) = \sum_{t=0}^{\infty} C(x_t,a_t) = \sum_{t=0}^{\infty} \alpha^t x_t + \sum_{t=0}^{\infty} \alpha^t a_t$$

$$= \sum_{t=0}^{\infty} 2^t x + \sum_{t=0}^{\infty} \sum_{n=1}^{t} 2^t (2/\alpha)^{-n} a_{n-1} + \sum_{t=0}^{\infty} \alpha^t a_t$$

Por lo tanto, $V_\alpha(\pi,x)=\infty$ para toda $\pi\in\Pi$ y $x\in X$, y de aquí, $V_\alpha(x)=\infty$ para toda $x\in X$, i.e., no se cumple la hipótesis 3.3.

Ahora, por otro lado, se puede probar por inducción que $v_n(0)=0$ para toda n y por lo tanto v_n no converge a V_α para todo $x\in X$. |

En el caso con costos acotados, se puede ver [8, 9, 15] que una política π^* es α-óptima si y solo si el costo $V_\alpha(\pi^*,.)$ satisface la EO. En nuestro caso esta equivalencia no se satisface, solo podemos asegurar la suficiencia de la α-optimalidad de π^* para que el costo correspondiente a π^* satisfaga la EO. Puede darse el caso en que para alguna política dada π^*, el costo $V_\alpha(\pi^*,.)$ satisfaga la EO, pero π^* no sea α-óptima y por el Teorema 3.6(c) no se satisfaga (3.12). Esto lo ilustra el siguiente ejemplo.

EJEMPLO 3.8: Sea $X=A=\mathbb{R}$, $A(x)=[0,1]$ y la dinámica del sistema es $x_{t+1}=\alpha^{-1}a_tx_t$, con costo por etapa $C(x_t,a_t)=|x_t|$. Claramente el problema satisface las hipótesis 2.2. Consideremos la política $\pi^*=\{1,1,\ldots\}$, entonces, es fácil probar que si $x_0=x$, tenemos que $x_t=\alpha^{-t}x$, $t=0,1,\ldots$ y de aquí,

$$V_\alpha(\pi^*,x) = E_x^{\pi^*}\left[\sum_{t=0}^{\infty}\alpha^t|x_t|\right] = E_x^{\pi^*}\left[\sum_{t=0}^{\infty}|x|\right] = \infty \quad \forall\ x\neq 0$$

y $V_\alpha(\pi^*,0)=0$. Por otra parte, $TV_\alpha^*(\pi,x)=\infty$, i.e., $V_\alpha^*(\pi,x)$ satisface la EO, y por otra parte, $V_\alpha(x)=|x|$. Por lo tanto, π^* no es α-óptima. |

Si tenemos que π^* es una política tal que $V_\alpha(\pi^*,.)$ satisface la EO, por el Teorema 3.6(d), podemos garantizar que π^* es α-óptima si se cumple la condición (3.12). Claramente, si los costos por etapa son acotados, digamos $0\leq C(x,a)\leq M$ para todo $(x,a)\in K$, se cumple (3.12) ya que de (1.1) tenemos que $0\leq V(\pi,.)\leq M/(1-\alpha)$ para todo $\pi\in\Pi$. Condiciones suficientes para la condición (3.12) se dan en [22].

Para concluir esta sección, ilustraremos los elementos expuestos con los ejemplos dados en el Capítulo I.

La EO para el problema de Producción-Inventario es:

$$V_\alpha(x) = \min_{a\in A(x)} E\left\{ ca + h(x+a-\xi)^+ + \alpha V_\alpha(x+a-\xi)^+ \right\} \qquad (3.16)$$

Para verificar la hipótesis 3.3, consideremos la política estacionaria $\{\hat{f}\}$ definida por $\hat{f}(x)=x$ $\forall$ $x\in X$. Como $x\in[0,M]$, la política $\hat{f}$ es acotada. Por lo tanto, el costo cuando se usa $\hat{f}$ es acotado, y por la presemcia del factor de descuento α, tenemos que $V_\alpha(\hat{f},x)<+\infty$ $\forall$ $x\in X$ y así, $V_\alpha(x)<+\infty$ $\forall$ $x\in X$.

En [8] pag. 244, se muestra que una política α-óptima $f\in F$ para este problema es:

$$f(x)=\begin{cases} S^{*}-x & \text{si } x\le S^{*} \\ 0 & \text{si } x>S^{*} \end{cases} \tag{3.17}$$

donde S^{*} minimiza a la función:

$$G^{*}(y) = cy + L(y) + E\,\{V_{\alpha}(y-\xi)\}$$

con $L(y) = hE\,\{(y-\xi)^{+}\}$.

El punto S^{*} existe, ya que $\lim_{|y|\to\infty} G^{*}(y)=\infty$ y además $L(y)$ y $V_{\alpha}(x)$ son convexas, lo que implica que $G^{*}(y)$ tenga esta propiedad. La convexidad de V_{α} se demuestra por inducción (ver [8] pag. 66).

Para los sistemas LQ, la EO es:

$$V_{\alpha}(x) = \min_{a} \left\{ qx^{2} + ra^{2} + \alpha E\, V_{\alpha}(\gamma x + \beta a + \xi)\right\}.$$

En [8] pag. 242 se muestra que:

$$V_{\alpha}(x) = k(\alpha)x^{2} + k(\alpha)\alpha\sigma^{2}/(1-\alpha) \quad \forall\ x\in\mathbb{R},\ \alpha\in(0,1), \tag{3.18}$$

donde $k(\alpha)$ es la única solución positiva de la ecuación:

$$k = [1 - \alpha k\beta^{2}(r + \alpha k\beta^{2})^{-1}]\alpha k\gamma^{2} + q \tag{3.19}$$

y

$$k(\alpha)\to k^{*} \text{ cuando } \alpha\nearrow 1, \tag{3.20}$$

donde k^{*} es la única solución positiva de la ecuación que resulta cuando hacemos $\alpha\nearrow 1$ en (3.19).

Claramente, $V_{\alpha}(x)<+\infty$ para toda $x\in\mathbf{X}$, con lo que se verifica la hipótesis 3.3, y además, una política α-óptima $f\in\mathbf{F}$ es:

$$f(x) = -\alpha\beta\gamma k(\alpha)x/(\alpha k(\alpha)\beta^{2} + r) \tag{3.21}$$

Para obtener la EO para el problema de sistema de colas, primero introduciremos alguna notación.

Sea $y=(y_1,y_2,\ldots,y_N)$ el vector que indica el número de nuevos trabajos en cada cola. Denotaremos por ey al vector que resulta de multiplicar los vectores e (p. 18) y y coordenada a coordenada [ver Ej. 6.3, Cap.I], es decir, ey es el vector tal que su j-ésima componente es:

$$(ey)_j = \begin{cases} 0 & \text{si } e_j=0 \\ y_j & \text{si } e_j=1 \end{cases} \qquad j=1,2,\ldots,N$$

Esto es, el vector ey indica el número de trabajos esperando en la cola sin atenderlos.

Sea $E(e)=\{j \ : \ e_j=1\}$ y además, si $x_k>0$, definimos el vector $x(\beta(k))$ cuyas coordenadas son:

$$[x(\beta(k))]_j = \begin{cases} x_j & \text{si } j\neq k \\ (x_k-\beta(k))^+ & \text{si } j=k \end{cases} \qquad j=1,2,\ldots,N$$

Esto es, el vector $x(\beta(k))$ indica el número de trabajos en la cola tomando la decisión de servirlos.

Con todo esto, la EO es:

$$V_\alpha(0) = \min_e\left\{C(0,(e,1)) + \alpha\int\ldots\int V_\alpha(ey)\prod_{j\in E(e)} dF_j(y_j)\right\} \qquad (3.22)$$

$$V_\alpha(x) = \min_e\left\{C(0,(e,1)) + \alpha\int\ldots\int V_\alpha(x+ey)\prod_{j\in E(e)} dF_j(y_j)\right\} \wedge$$

$$\min_e \min_k \min_{\beta(k)}\left\{C(x,(e,\beta(k)))+\alpha\int\ldots\int V_\alpha(x(\beta(k))+ey)\prod_{j\in E(e)} dF_j(y_j)\right\}$$

donde $T\wedge S = \min\{T,S\}$ y si E(e) es vacío la integral es el valor del integrando.

Para verificar la hipótesis 3.3, sea $\hat{f}$ la política que siempre escoge $e^*=(0,0,\ldots,0)$ y que sirve a los objetos en las colas en un orden fijo, digamos, la primer cola hasta que

152539

quede vacía, luego la segunda hasta que también quede vacía, etc.. Sea $x=x_0$, un estado inicial arbitrario. Aplicando la política $\hat{f}$, el proceso alcanza el estado cero en un número finito de pasos con costo finito.

De la ecuación (3.22), tenemos que:

$$V_\alpha(0) \leq C(0,(e^*,i)) + \alpha\int\ldots\int V_\alpha(e^*y) \prod_{j\in E(e^*)} dF_j(y_j)\Big\}$$

$$= C(0,(e^*,i)) + \alpha V_\alpha(0) \qquad (3.23)$$

Esta última igualdad es por el hecho de que $E(e^*)$ es vacío. De la expresión (3.23) obtenemos que:

$$V_\alpha(0) \leq C(0,(e^*,i))/(1-\alpha) < +\infty$$

y por el comentario hecho en relación a la política $\hat{f}$, $V_\alpha(x)<+\infty$ para toda $x\in X$, lo cual verifica la hipótesis 3.3.

Para los problemas invariantes la EO, en su forma general, es:

$$V_\alpha(x) = \min_a \left\{ C(x,a) + \alpha\int v(y)Q(dy/a) \right\}$$

La hipótesis 3.3 se cumple, como verificaremos en el capítulo siguiente.

Como $A(x)$ es finito, de antemano podemos decir que existe una política α-óptima para estos problemas.

CAPITULO III

PROCESOS DE CONTROL DE MARKOV CON COSTO PROMEDIO

1.- Introducción.

2.- Hipótesis sobre el modelo de control.

3.- Condiciones de optimalidad (I).

4.- Condiciones de optimalidad (II).

1.- Introducción.

Matemáticamente hablando, los problemas de control óptimo con costo promedio son los más difíciles de analizar. A diferencia del caso con costos descontados (Cap.II), aquí no tenemos en general una solución de la ecuación de optimalidad. Por esta razón, existen varios caminos para resolver el problema dependiendo de las restricciones en el modelo de control. Abundaremos más sobre este problema en la sección 3. El objetivo de este capítulo es obtener la optimalidad en costo promedio de una política bajo condiciones generales, como las hipótesis 2.2 del capítulo anterior, vía el caso descontado.

Analizaremos la optimalidad en costo promedio bajo dos tipos de hipótesis.

En la sección 3, trabajaremos con las hipótesis 2.1, que definiremos en la sección 2, además daremos condiciones (C1) para garantizar la existencia de una política CP-óptima (Teo. 3.4). Posteriormente, definiremos dos condiciones más (C2, C3), las cuales son suficientes para C1 (Teo. 3.7).

En la sección 4, supondremos que se cumplen la hipótesis 2.2 del capítulo anterior junto con la condición C1, y además impondremos condiciones a la función de valor óptimo α-descontada y al espacio de estados **X**, lo cual nos garantizará la existencia de una política CP-óptima (Teo. 4.2).

La diferencia entre los dos bloques de hipótesis está en el kernel de transición Q. En la primera (Hip. 2.1) se pide que Q sea fuertemente continuo, y en la segunda (Hip. 2.2, II), Q es débilmente continuo. Claramente la Hipótesis 2.1 implica la Hipótesis 2.2,II.

Por último, para una rápida referencia, enumeramos de nuevo el índice de funcionamiento para el caso promedio (5.2, I), y la función de valor CP-óptima (5.5, I).

$$J(\pi,x):=\limsup_{n} (n+1)^{-1} \sum_{t=0}^{n} E_x^{\pi} [C(x_t,a_t)] \qquad (1.1)$$

$$J(x):=\inf_{\pi} J(\pi,x) \qquad (1.2)$$

2.- Hipótesis sobre el modelo de control.

Sea **(X,A,Q,C)** un modelo de control de Markov. Al igual que en el capítulo anterior, impondremos condiciones para validar los teoremas de selección apropiados.

Hipótesis 2.1:

(a) $C(x\ ,a\ )$ es inf-compacta y pertenece a $\mathbf{L(K)}_+$

(b) La ley de transición Q es fuertemente continua [Def.2(a), Ap.B]

(c) La multifunción $x\to A(x)$ es semicontinua inferiormente (sci).

(c') La multifunción $x\to A^{*}(x)$ es sci, donde $A^{*}(x)$ es como en la hipótesis 2.2(c') del Cap. II.

Para verificar que los ejemplos del capitulo I cumplen con estas hipótesis, solo basta probar que 2.1(b) se cumple. Para esto daremos el siguiente criterio:

LEMA 2.2: Sea $\{x_t\}$ un proceso de markov definido por:

$$x_{t+1}=G(x_t)+\xi_t$$

donde $G:X\to X$ es una función dada y $\{\xi_t\}$ son variables aleatorias i.i.d. con valores en un espacio de Borel S e independientes del estado x_0. Además, supongamos que su distribución común μ tiene una densidad g respecto a la medida de Lebesgue, i.e., $\mu(ds)=g(s)ds$. Entonces, si $X=S=\mathbb{R}^n$, G es continua y si g es una densidad continua y acotada, el kernel estocástico Q es fuertemente continuo.

DEMOSTRACION:

Sea v una función medible y acotada en X. Deseamos demostrar que la función

$$x \to \int v(y)Q(dy/x)$$

es continua y acotada. Claramente esta función es acotada. Para demostrar la continuidad, primero tenemos que:

$$\int v(y)Q(dy/x) = \int v[G(x)+s]g(s)ds = \int v(x)g[s-G(x)]ds \qquad (2.1)$$

para $x\in\mathbb{R}^n$. Ahora, supongase que $x^n \to x$. Por (2.1) tenemos que:

$$\left|\int v(y)Q(dy/x) - \int v(y)Q(dy/x)\right| = \left|\int v(x^n)g[s-G(x^n)]ds\right.$$

$$\left.-\int v(x)g[s-G(x)]ds\right| \le |v(x^n)-v(x)| \int |g[s-G(x^n)]-g[s-G(x)]|ds$$

Por la continuidad de v y el teorema de Scheffé [15, p.p. 125], esta última expresión converge a cero, lo cual demuestra lo que se quería.|

Para los problemas LQ, si suponemos que las variables aleatorias ξ_t tienen densidad normal con media cero y varianza σ^2, se sigue del lema 2.2 que Q es fuertemente continua, i.e., la hipótesis 2.1(b) se cumple. Para el problema de Producción-Inventario, necesitamos pedir que la demanda ξ_t tenga densidad continua y acotada para garantizar que la hipótesis 2.1(b) se cumple.

Para los problemas invariantes y de sistema de colas, la prueba es similar a la del capítulo anterior. En ambos casos tenemos continuidad de Q.

Siguiendo el mismo esquema del capítulo anterior, daremos el siguiente Teorema de Selección Medible bajo las hipótesis 2.1.

TEOREMA 2.3: Supongamos que las hipótesis 2.1 se cumplen, y sea u una función medible no negativa en X. Entonces la función $\bar{u}$ es medible, donde:

$$\bar{u}(x) := \inf_{a \in A(x)} \left[C(x,a) + \int u(y) Q(dy/x,a) \right] \qquad x \in X$$

Además, existe $f \in F$ tal que

$$\bar{u} = C(x,f(x)) + \int u(y) Q(dy/x,f(x)) \quad \forall\ x.$$

DEMOSTRACION:

La medibilidad de $\bar{u}$ se sigue de 2.1(a) y 2.1(b) y la existencia del selector $f \in F$ se sigue de [30] Cor. 4.3.|

De este resultado, podemos ver que el Teorema 3.6(c) del capítulo II también es válido bajo las hipótesis 2.1.

En trabajos anteriores al [27], las hipótesis usuales eran mas fuertes que las 2.1 y las 2.2 cap. II, como se vió en la sección 2 del capítulo anterior. El suponer estas hipótesis nos permite generalizar los resultados para el caso promedio, pero pagando el precio de no tener soluciones de la ecuación de optimalidad correspondiente al caso promedio como veremos en la próxima sección.

3.- Condiciones de Optimalidad (I).

En esta sección, daremos condiciones que nos garanticen la existencia de políticas costo promedio óptimas (CP-óptima), en el marco de las hipótesis 2.1.

DEFINICION 3.1: Un par $(j^*, h(.))$ que consiste de un número real j^* y una función medible h sobre **X**, se dice que es una solución de la Ecuación de Optimalidad con Costo Promedio (EOCP), si para toda $x \in \mathbf{X}$,

$$j^* + h(x) = \min_{a \in A(x)} \left[C(x,a) + \int h(y) Q(dy/x,a) \right]. \tag{3.1}$$

Diremos que una solución $(j^*, h(.))$ a (3.1) es acotada si h es acotada.

En el caso en que los costos por etapa $C(x,a)$ son acotados y/o el espacio de estados **X** es numerable, se puede garantizar, bajo ciertas condiciones, la existencia de una solución a (3.1) y de una política CP-óptima (e.g. [8,15,19,32]). Por ejemplo, en [19] se imponen condiciones de ergodicidad en el modelo de control, para garantizar la existencia de una solución acotada a la EOCP junto con una política CP-óptima.

En nuestro caso (espacio de estados de Borel y costos posiblemente no acotados), esto no lo podemos asegurar directamente. En este mismo caso, en los últimos 4 años, diferentes autores [16,20,27] han dado condiciones para garantizar la existencia de políticas CP-óptimas, pero no de una solución a (3.1). Estos autores extendieron a espacios de estados y acciones de Borel los resultados de [35,36, 37], en los cuales **X** es numerable y **A** es finito. Sin embargo solo se pudo garantizar la existencia de soluciones de la desigualdad:

$$j^{*} + h(x) \geq \min \left[C(x,a) + \int h(y) Q(dy/x,a) \right] \qquad (3.2)$$

A (3.2) se le conoce como Desigualdad de Optimalidad con Costo Promedio (DOCP).

Mas aún, en [10], se demuestra que es posible obtener la desigualdad estricta en (3.2) bajo las condiciones que se usan en [37].

El hecho de no tener igualdad es una desventaja, ya que se dificultaría implementar esquemas de iteración de valores como en el caso descontado, para aproximar a la función de valor CP-óptima (1.2).

Sin embargo, anexando condiciones como convexidad [14] y equicontinuidad [28] de ciertas funciones se puede tener la igualdad. Comentaremos este caso en las conclusiones.

Existen varias maneras de obtener la CP-optimalidad. Una es imponiendo condiciones de ergodicidad al modelo [15,19]. Otra es imponer condiciones a la función de valor óptimo α-descontada [16,20,27,35,36,37].

Esta última, es la que se conoce como el enfoque del factor de descuento desvaneciente, y es el que se usará en este trabajo.

Para esto, primero enunciaremos la condición de optimalidad que nos garantizará la existencia de la política CP-óptima (Teo. 3.4), y daremos algunos lemas que nos relacionan los índices de funcionamiento del caso descontado (1.1,II) y el del caso promedio (1.1).

Sea $V_{\alpha}(.)$ la función de valor α- descontada (1.2,II) y sea $\bar{x} \in X$ un estado fijo pero arbitrario. Definimos la función:

$$h_\alpha(x) := V_\alpha(x) - V_\alpha(\bar{x}), \qquad x\in X,\ \alpha\in(0,1). \tag{3.3}$$

CONDICION 1 (C1): (Condición de Optimalidad). Existen constantes no negativas N y M, una función no negativa b sobre X, no necesariamente medible, y $\alpha_0\in(0,1)$ tal que:

(a) $V_\alpha(x) < +\infty$ para todo $x\in X$ y $\alpha\in(0,1)$;

(b) $(1-\alpha)V_\alpha(\bar{x}) \le M$ para todo $\alpha\in[\alpha_0,1)$;

(c) $-N \le h_\alpha(x) \le b(x)$ para todo $x\in X$ y $\alpha\in[\alpha_0,1)$.

LEMA 3.2: Bajo la hipotesis 2.1(a), sean $J(\pi,x)$, $J^*(x)$, $V_\alpha(\pi,x)$ y $V_\alpha(x)$ como en (1.1), (1.2), (1.1,II) y (1.2,II) respectivamente. Entonces para toda π y x,

(a) $\limsup_{\alpha\nearrow 1} (1-\alpha)V_\alpha(\pi,x) \le J(\pi,x)$

(b) $\limsup_{\alpha\nearrow 1} (1-\alpha)V_\alpha(x) \le J(\pi,x)$

(c) $\limsup_{\alpha\nearrow 1} (1-\alpha)V_\alpha(x) \le J^*(x)$

DEMOSTRACION:

(a) La parte (a) se sigue del Teorema 1, Ap. C, tomando

$$C_t := E_x^\pi\, C(x_t,a_t) \qquad\qquad S_n := J_n(\pi,x) = \sum_{t=0}^{n-1} E_x^\pi\, C(x_t,a_t)$$

(b) Como $V_\alpha(x) \le V_\alpha(\pi,x)$ para toda π y x, se sigue de (a) que:

$$\limsup_{\alpha\nearrow 1} (1-\alpha)\, V_\alpha(x) \le J(\pi,x) \qquad \forall\ \pi,\ x.$$

(c) Se obtiene de (b) tomando el infimo sobre π.

∎

LEMA 3.3: Bajo C1, existe una constante $j^*\ge 0$ y una sucesión de factores de descuentos $\alpha(n)\nearrow 1$ tal que:

$$\lim_n\, (1-\alpha(n))\, V_{\alpha(n)}(x) = j^* \qquad \forall x\in X \tag{3.4}$$

DEMOSTRACION:

Sea $\bar{x}$ el estado fijo en (3.3). Por C1(b), existe el límite de $(1-\alpha)V_{\alpha}(\bar{x})$ cuando $\alpha \nearrow 1$, digamos que es j^{*}.

Por otro lado, sea $\alpha_n \nearrow 1$ tal que:

$$\lim_{n} (1-\alpha(n))V_{\alpha(n)}(\bar{x}) = j^{*} \tag{3.5}$$

Ahora, por C1, para cada x,

$$\left|(1-\alpha(n))V_{\alpha(n)}(x)-j^{*}\right|=\left|(1-\alpha(n))V_{\alpha(n)}(x)-(1-\alpha(n))V_{\alpha(n)}(\bar{x})+\right.$$

$$\left.+(1-\alpha(n))V_{\alpha(n)}(\bar{x})-j^{*}\right|$$

$$= \left|(1-\alpha(n))(V_{\alpha(n)}(x)-V_{\alpha(n)}(\bar{x}))+(1-\alpha(n))V_{\alpha(n)}(\bar{x})-j^{*}\right|$$

$$\leq \left|(1-\alpha(n))h_{\alpha(n)}(x)\right|+\left|(1-\alpha(n))V_{\alpha(n)}(\bar{x})-j^{*}\right| \qquad \text{[por (3.3)]}$$

$$\leq (1-\alpha(n))\max\{N,\ b(x)\}+\left|(1-\alpha(n))V_{\alpha(n)}(\bar{x})-j^{*}\right| \qquad \text{[por C1(c)]}$$

$$\to 0 \quad \text{cuando} \quad n\to\infty \qquad \text{[por (3.5)]}$$

Lo cual prueba lo que se quería. |

A continuación, daremos el teorema principal de esta sección, que garantiza la existencia de una política CP-óptima pero no la igualdad en (3.1), esto es, obtendremos la DOCP (3.2).

TEOREMA 3.4: Supongamos que las hipótesis 2.1 y la condición C1 se cumplen. Sea j^{*} la constante obtenida en el lema 3.3. Entonces:

(a) Existe una función medible h en **X**, tal que $-N\leq h(.)\leq b(.)$ y

$$j^{*}+h(x) \geq \min_{a\in A(x)}\left[C(x,a) + \int h(y)Q(dy/x,a) \right] \quad \forall\ x\in X. \tag{3.6}$$

(b) Existe una política estacionaria $f^{*}\in F$ tal que:

$$j^{*}+h(x) \geq C(x,f^{*}(x)) + \int h(y)Q(dy/x,f^{*}(x)) \quad \forall\ x\in X \tag{3.7}$$

(c) f^{*} es CP-óptima y $J(f^{*},x)=j^{*}$ $\forall$ $x\in X$.

DEMOSTRACION:

(a) En vista del Lema 3.3, sea $\{\alpha(n)\}$ tal que $\alpha(n) \nearrow 1$. Definimos la función $h:X\to R$, como:

$$h(x) := \liminf_{n} h_{\alpha(n)}(x) = \lim_{n} H_n(x), \quad x\in X, \tag{3.8}$$

donde $H_n(x) := \inf_{k\geq n} h_{\alpha(k)}(x)$.

Aparte sea $x\in X$ arbitrario. Entonces por el Teorema 3.6 del capítulo II, para cada n existe $a_n\in A(x)$ tal que:

$$V_{\alpha(n)}(x) = C(x,a_n) + \alpha(n)\int V_{\alpha(n)}(y)Q(dy/x,a_n)$$

De aquí, restando la constante $\alpha(n)V_{\alpha(n)}(\bar{x})$,

$$V_{\alpha(n)}(x)-\alpha(n)V_{\alpha(n)}(\bar{x})=$$

$$= C(x,a_n) + \alpha(n)\int [V_{\alpha(n)}(y)-V_{\alpha(n)}(\bar{x})]Q(dy/x,a_n)$$

$$= C(x,a_n) + \alpha(n)\int h_{\alpha(n)}(y)Q(dy/x,a_n) \quad \forall\ n \quad \text{[por (3.3)]}$$

Sumando y restando $V_{\alpha(n)}(\bar{x})$ y por (3.3), tenemos que:

$$(1-\alpha(n))V_{\alpha(n)}(\bar{x}) + h_{\alpha(n)}(x) =$$

$$= C(x,a_n) + \alpha(n)\int h_{\alpha(n)}(y)Q(dy/x,a_n) \quad \forall n \tag{3.9}$$

Por otro lado, de (3.5) y (3.8),

$$j^{*} + h(x) = \lim_{n}\inf \left[(1-\alpha(n))V_{\alpha(n)}(\bar{x}) + h_{\alpha(n)}(x)\right]$$

Así, para cualquier $\varepsilon > 0$ dado, existe un entero $n(\varepsilon)$ y una subsucesión $\{\alpha(n_i)\}$ de $\{\alpha(n)\}$ tal que:

$$j^*+h(x)+\varepsilon \ge (1-\alpha(n_i))V_{\alpha(n_i)}(x) + h_{\alpha(n_i)}(x) \quad \forall\ n_i \ge n(\varepsilon) \qquad (3.10)$$

y por (3.9) y la definición de $Hn(x)$, tenemos:

$$j^* + h(x) + \varepsilon \ge C(x,a_{n_i}) + \alpha(n_i)\int h_{\alpha(n_i)}(y)Q(dy/x,a_{n_i})$$

$$\ge C(x,a_{n_i}) + \alpha(n)_i\int Hn_i(y)Q(dy/x,a_{n_i}) \quad \forall\ n_i \ge n(\epsilon) \qquad (3.11)$$

Sumando $\alpha(n_i)N$ en ambos lados de (3.11),

$$j^*+h(x)+\alpha(n_i)N+\epsilon \ge$$

$$C(x,a_{n_i})+\alpha(n_i)\int[Hn_i(y)+N]Q(dy/x,a_{n_i}) \quad \forall\ n_i \ge n(\epsilon)$$

Como $\alpha(n_i)<1$, entonces $N > \alpha(n_i)N$, lo cual implica

$$j^* +h(x) +N +\epsilon \ge C(x,a_{n_i})+\alpha(n_i)\int[Hn_i(y)+N]Q(dy/x,a_{n_i}) \qquad (3.12)$$

para toda $n_i \ge n(\epsilon)$.

Aparte, para $i=1,2,3,\ldots$, definamos los conjuntos:

$$D_i(x):=\{a\in A(x):C(x,a)+\alpha(n_i)\int[Hn_i(y)+N]Q(dy/x,a)\le j^*+h(x)+N+\varepsilon\},$$

es decir, $D_i(x)$ es el conjunto de $a\in A(x)$ tal que (3.12) se cumple.

Sea $r:=j^*+h(x)+N+\varepsilon$ y $A_r(x):=\{a\in A(x):\ C(x,a) \le r\}$. De C1(c) tenemos que $Hn_i(.)+N\ge 0$ y por 2.1(b), $D_i(x)$ es un subconjunto cerrado de $A_r(x)$. Así, por la hipótesis 2.1(a), los conjuntos $D_i(x)$ son compactos. Además, por (3.12),

$a_{n_i} \in D_i(x)$ $\forall$ $n_i \geq n(\varepsilon)$ y como $\alpha(n_i)(H_{n_i}(.)+N) \nearrow h(.)+N$, entonces, $D_i \subseteq D_j$ $\forall$ $j<i$, y por lo tanto, los conjuntos $D_i(x)$ forman una sucesión no decreciente de conjuntos compactos, distintos del vacío, que converge al conjunto compacto no vacío:

$$D(x):=\{a \in A(x): C(x,a) + \int [h(y)+N]Q(dy/x,a) \leq j^{*}+h(x)+ N + \varepsilon\}.$$

De aquí, existe una subsucesión de $\{n_i\}$, la cual la denotaremos igual por $\{n_i\}$ para no complicar la notación, y un control $a_x \in D(x)$ tal que $a_{n_i} \to a_x$ cuando $i \to \infty$.

Aparte, sea $L>n(\varepsilon)$ un entero arbitrario. Entonces de (3.12),

$$j^{*}+h(x)+N+\varepsilon \geq C(x,a_{n_i})+\alpha(n_i)\int [H_L(y)+N]Q(dy/x,a_{n_i}) \ \forall \ n_i>L \qquad (3.13)$$

Haciendo $i \to \infty$, por las hipótesis 2.1(a)-(b), obtenemos:

$$j^{*}+ h(x) + N + \varepsilon \geq C(x,a_x) + \int [H_L(y) + N]Q(dy/x,a_x)$$

y ahora haciendo $L \to \infty$ y por el Lema de Fatou:

$$j^{*}+ h(x) + N + \varepsilon \geq C(x,a_x) + \int [h(y) + N]Q(dy/x,a_x)$$

Restando N en ambos lados de la igualdad,

$$j^{*}+ h(x) + \varepsilon \geq C(x,a_x) + \int h(y)Q(dy/x,a_x)$$

$$\geq \min_{a \in A(x)} \left[C(x,a) + \int h(y)Q(dy/x,a) \right]$$

y como ε lo tomamos arbitrario, obtenemos lo que se quería.

(b) Primeramente, definamos $u(.):=h(.)+N$. Por C1(c), $u(.) \geq 0$, y además es medible. Entonces por el Teorema 2.3, existe $f^{*} \in F$ tal que para toda x:

$$\min_{a}\left[C(x,a)+\int[h(y)+N]Q(dy/x,a)\right] = C(x,f^{*})+\int[h(y)+N]Q(dy/x,f^{*})$$

Restando N en ambos lados de la igualdad y por (3.6), obtenemos:

$$j^{*}+h(x)\geq \min_{a}\left[C(x,a)+\int h(y)Q(dy/x,a)\right] = C(x,f^{*})+\int h(y)Q(dy/x,f^{*})$$

para toda $x\in X$. De aquí obtenemos (3.7).

(c) De la desigualdad (3.7) tenemos que:

$$j^{*}+ h(x)+ N \geq C(x,f^{*})+\int [h(y)+N]Q(dy/x,f^{*}) \qquad \forall\ x\in X$$

Iterando esta desigualdad,

$$nj^{*} + h(x) + N \geq E_x^{f^{*}}\left[\sum_{t=0}^{n-1} C(x_t,a_t)\right] + E_x^{f^{*}}\left[h(x_n)+N\right]$$

Còmo $h(.)+N$ es no negativa, entonces:

$$nj^{*}+h(x)+N \geq E_x^{f^{*}}\left[\sum_{t=0}^{n-1} C(x_t,a_t)\right]$$

lo cual implica:

$$j^{*}+ h(x)/n + N/n \geq n^{-1} E_x^{f^{*}}\left[\sum_{t=0}^{n-1} C(x_t,a_t)\right]$$

Tomando limsup cuando $n\to\infty$ y por (1.1) y (1.2),

$$j^{*} \geq J(f^{*},x) \geq J^{*}(x) \tag{3.14}$$

Por otro lado, del Lema 3.2(c) y Lema 3.3, tenemos que:

$$j^{*} \leq J^{*}(x) \leq J(f^{*},x) \tag{3.15}$$

Así, por (3.14) y (3.15), $j^*=J^*(x)=J(f^*,x)$, lo cual demuestra lo que se quería. |

Como podemos observar en la prueba del Teorema 3.4, la condición de optimalidad C1 juega un papel importante para obtener las políticas CP-óptimas. Existen otras condiciones que también dan las mismas conclusiones de este teorema, como las que se usan en [37] o [16] y las que se usan en [20] o [35]. Dichas condiciones son:

CONDICION 2 (C2): [37,16]. Existe una constante $N\geq 0$, una función $b(x)$, medible no negativa y una política estacionaria $\hat{f}\in\mathbf{F}$, tal que para todo $x\in\mathbf{X}$ y $\alpha\in(0,1)$,

(a) $V_\alpha(x)< +\infty$

(b) $h_\alpha(x)\geq -N$

(c) $h_\alpha(x)\leq b(x)$ y $\int b(y)Q(dy/x,\hat{f})< \infty$.

Notemos que, a diferencia de la condición C1, aquí si pedimos que la función b sea medible en X.

Definamos, $m_\alpha := \inf_x V_\alpha(x)$ y $g_\alpha(x) := V_\alpha(x)-m_\alpha$, para $x\in\mathbf{X}$ y $\alpha\in(0,1)$.

CONDICION 3 (C3): [20,35].

(a) Existe una política $\hat{\delta}$ y un estado inicial $\hat{x}$ tal que $J(\hat{\delta},\hat{x})<+\infty$.

(b) Existe $\beta\in[0,1)$ tal que $\sup_\alpha g_\alpha(x)<+\infty$ para todo $x\in\mathbf{X}$, donde el sup se toma sobre las α en $(\beta,1)$.

Demostraremos que cada condición (C2, C3), implica a la condición C1. Antes, demostraremos el siguiente Lema, con el cual nuevamente obtenemos una relación entre los índices de funcionamiento del caso descontado y el del caso promedio.

Definamos:

$$j := \inf_x \inf_\pi J(\pi,x)$$

$$g^L := \liminf_{\alpha \nearrow 1} (1-\alpha) m_\alpha$$

$$g^U := \limsup_{\alpha \nearrow 1} (1-\alpha) m_\alpha$$

LEMA 3.5: Supongase que C3(a) se cumple. Entonces:

$$0 \le g^L \le g^U \le j < \infty \tag{3.16}$$

DEMOSTRACION:

Notemos que, como $V_\alpha(x)<\infty$ para todo $x \in X$, $m_\alpha<\infty$ y de aquí, $g^L \ge 0$. Además, $g^L \le g^U$ por definición, y por C3(a), $j<\infty$. Así, solo basta probar que $g^U \le j$. Para esto, por el Lema 3.2(c) tenemos que, para todo x,

$$\limsup_{\alpha \nearrow 1} (1-\alpha) V_\alpha(x) \le J^*(x) = \inf_\pi J(\pi,x)$$

lo cual implica que:

$$\limsup_{\alpha \nearrow 1} (1-\alpha) \inf_x V_\alpha(x) \le \inf_x \inf_\pi J(\pi,x)$$

y de aquí, $g^U \le j$. |

LEMA 3.6: Supongase que C3 se cumple. Entonces $m_\alpha<\infty$ $\forall \alpha \in (0,1)$.

DEMOSTRACION:

Primero notemos que, como $C(x,a) \ge 0$ $\forall$ $(x,a) \in K$, si $\alpha' \le \alpha''$, donde α', $\alpha'' \in (0,1)$, entonces:

$$m_{\alpha'} \le m_{\alpha''} \tag{3.17}$$

esto es, el mapeo $\alpha \to m_\alpha$ es no decreciente.

Ahora supongamos que existe $\alpha(1) \in (0,1)$ tal que $m_{\alpha(1)}=+\infty$. Entonces por (3.17), $m_\alpha=+\infty$ $\forall$ $\alpha \ge \alpha(1)$. De aquí,

$$\limsup_{\alpha \nearrow 1} (1-\alpha) m_\alpha = +\infty$$

lo cual contradice a (3.16). Por lo tanto, bajo C3, $m_\alpha<+\infty$ $\forall$ $\alpha\in(0,1)$. |

TEOREMA 3.7: Supongamos que la hipótesis 2.1 (o la 2.2, II) se cumple. Entonces cada una de las condiciones C2 y C3 implican C1.

DEMOSTRACION:

(1) C2$\Rightarrow$C1.

Es claro que bajo C2 se cumple C1(a) y C1(c). Para demostrar que C1(b) se cumple, de la ecuación de optimalidad α-descontada (3.1,II), tenemos:

$$(1-\alpha)V_\alpha(\bar{x})=V_\alpha(\bar{x})-\alpha V_\alpha(\bar{x})\le C(\bar{x},\hat{f}(\bar{x}))+\alpha\int V_\alpha(y)Q(dy/\bar{x},\hat{f}(\bar{x}))-\alpha V_\alpha(\bar{x})$$

$$= C(\bar{x},\hat{f}(\bar{x})) +\alpha\int h_\alpha(y)Q(dy/\bar{x},\hat{f}(\bar{x})) \qquad \text{[por (3.3)]}$$

$$\le C(\bar{x},\hat{f}(\bar{x})) + \int b(y)Q(dy/\bar{x},\hat{f}(\bar{x})) \qquad \text{[por C2(c)]}.$$

Definiendo $M:= C(\bar{x},\hat{f}(\bar{x})) + \int b(y)Q(dy/\bar{x},\hat{f}(\bar{x}))$ obtenemos C1(b). Por lo tanto, C1 se cumple.

(2) C3$\Rightarrow$C1

Por el Lema 3.6, tenemos que,

$$V_\alpha(x) = (V_\alpha(x)-m_\alpha) + m_\alpha = g_\alpha(x) + m_\alpha$$

$$\le \sup_\alpha g_\alpha(x) + m_\alpha < +\infty \tag{3.18}$$

donde el sup se toma sobre $\alpha\in(\beta,1)$, y esta última desigualdad es por C3(b). Por lo tanto se cumple C1(a).

Ahora, definamos $\bar{x} = \hat{x}$, donde $\bar{x}$ es el estado fijo en (3.3) y $\hat{x}$ es el estado en C3(a). Por C3(a) y el Lema 3.2(c), tenemos:

$$\limsup_{\alpha\nearrow 1} (1-\alpha)V_\alpha(\bar{x}) \le J^*(\bar{x}) < +\infty$$

es decir, lim sup de $(1-\alpha)V_\alpha(\bar{x})$ está acotado por una constante que no depende de α pero sí de $\bar{x}$. Por lo tanto, existe una constante $L:=L(\bar{x})>0$ y $M:=M(\bar{x})\in(0,1)$ tal que:

$$0 \le (1-\alpha)V_\alpha(\bar{x}) \le L \qquad \forall\ \alpha\in[M,1)$$

De aquí, definiendo $\alpha_0:=M$, obtenemos C1(b).

Por último, para obtener C1(c), primero notemos que:

$$h_\alpha(x) \le g_\alpha(x) \qquad \forall\ x\in X,\ \alpha\in(0,1). \tag{3.19}$$

Ahora, como $g_\alpha(x) \le \sup_\alpha g_\alpha(x)$ donde el sup se toma sobre los $\alpha\in(\beta,1)$ para algún $\beta\in[0.1)$, por (3.19) tenemos:

$$h_\alpha(x) \le \sup_\alpha g_\alpha(\bar{x}). \tag{3.20}$$

Por otro lado, de la definición de $g_\alpha(x)$, podemos decir que $g_\alpha(\bar{x}) \ge V_\alpha(\bar{x}) - V_\alpha(\bar{x})$, para todo $x\in X$. De aquí,

$$-\sup_\alpha g_\alpha(\bar{x}) \le h_\alpha(x) \tag{3.21}$$

para $\alpha\in(\beta,1)$. Definiendo $b(x)=\sup_\alpha g_\alpha(x)$, $x\in X$, y $N:=\sup_\alpha g_\alpha(\bar{x})$, de (3.20) y (3.21) obtenemos

$$-N \le h_\alpha(x) \le b(x) \qquad \forall\ x\in X,\ \alpha\in(0,1) \tag{3.22}$$

Por lo tanto, se cumple C1(c) para $\alpha_0\in(0,1)$ arbitrario, y así C3 implica C1. |

En [27] se da una condición adicional, que junto con C1, C2 y C3 se obtienen condiciones más fuertes (C1', C2' y C3'), con las cuales se obtienen relaciones adicionales entre C1, C2 y C3. Dicha condición es:

CONDICION (C): Existe un número $\theta\in(0,1)$, una función no negativa φ, medible sobre X, y una política θ-óptima $f_\theta\in F$ tal que, para toda $x \in X$,

$$\sup_{\theta<\alpha<1} h_\alpha(x) \leq h_\theta(x) + \varphi(x) \quad \text{y} \quad \int \varphi(y) Q(dy/x, f_\theta) < +\infty$$

Ahora definamos:

CONDICION 1' (C1'): C1 y C se cumplen.

CONDICION 2' (C2'): C2 y C se cumplen.

CONDICION 3' (C3'): C3 y C se cumplen.

Las relaciones adicionales entre C1, C2 y C3 las da el siguiente Teorema:

TEOREMA 3.8: Supongamos que se cumplen las hipótesis 2.1. Entonces:

(a) C1 implica C3;

(b) C2 implica C3;

(c) C2' implica C3;

(d) C3' implica C2

(e) C1' implica C2.

DEMOSTRACION:

(a) Como C1 y las hipótesis 2.1 se cumplen, por el Teorema 3.4, existe una política CP-óptima f^* para toda $x \in X$. Tomando $\hat{\pi}=f^*$, obtenemos C3(a). Por otro lado, para $\bar{x}$ fijo como en (3.3),

$$g_\alpha(x) := V_\alpha(x) - m_\alpha = (V_\alpha(x) - V_\alpha(\bar{x})) + V_\alpha(\bar{x}) - m_\alpha$$

$$= h_\alpha(x) + V_\alpha(\bar{x}) - m_\alpha \qquad \text{[por (3.3)]}$$

$$= h_\alpha(x) + V_\alpha(\bar{x}) - \inf_x V_\alpha(x) = h_\alpha(x) + V_\alpha(\bar{x}) + \sup_x(-V_\alpha(x))$$

$$\leq b(x) + \sup_x (V_\alpha(\bar{x}) - V_\alpha(x))$$

$$< +\infty \quad \forall\ \alpha\in(\alpha_0,1),\ x\in X \quad [\text{por C1}].$$

Por lo tanto, definiendo $\beta:=\alpha_0$, tenemos que $\sup_{\alpha} g_\alpha(x)<+\infty$ $\forall$ $x\in \mathbf{X}$, donde el sup se toma sobre las $\alpha\in(\beta,1)$. Con esto queda demostrado C3(b).

(b) Se sigue del Teorema 3.7 y la parte (a).

(c) Se sigue de las siguientes implicaciones:
C2' $\Rightarrow$ C2 (por definición); C2 $\Rightarrow$ C1 (Teo. 3.7); C1 $\Rightarrow$ C3 (por(a)).

(d) Supongamos que se cumple C3'. Por (3.18) y (3.22) en la prueba del Teorema 3.7, obtenemos C2(a) y C2(b).

Ahora por C, tenemos que $\sup_{\alpha} h_\alpha(x) <+\infty$ con $\alpha\in(\theta,1)$. Además, por el Teorema 3.6 del capítulo II, esta función pertenece a L(**X**), y por lo tanto es medible. Por C nuevamente,

$$\int \sup_{\alpha} h_\alpha(y)Q(dy/x,f_\theta)$$

$$\leq \int h_\theta(y)Q(dy/x,f_\theta) + \int \varphi(y)Q(d/y/x,f_\theta) < +\infty \tag{3.23}$$

donde $\alpha\in(\theta,1)$. Definiendo $b(x):= \sup_{\alpha} h_\alpha(x)$ y $\hat{f}:=f_\theta$, por (3.23) tenemos:

$$h_\alpha(x) \leq b(x) \qquad\qquad \int b(y)Q(dy/x,\hat{f}) < +\infty$$

con lo cual obtenemos C2(c).

(e) Por (a), C1, implica C3, y de aquí, C1, C implican C3, C, esto es , C1' implica C3', y por (d), C3' implica C2. Por lo tanto, C1' implica C2, con lo que se demuestra (d). |

De los teoremas 3.7 y 3.8 comcluimos las suiguientes relaciones:

$$
\begin{array}{ccccc}
C3' & & & & \\
 & \searrow & & & \\
C1' \to & C2 & \searrow & & \\
 & \downarrow & & C1 & \\
C2' \to & C3 & \nearrow & &
\end{array}
$$

El tener condiciones suficientes para la condición de optimalidad C1, es una gran ventaja, ya que abre la posibilidad de asegurar la existencia de políticas CP-óptimas en problemas específicos. De hecho, las condiciones C2 y C3 son esencialmente mas fáciles de verificar que la condición C1.

A continuación ilustraremos algunas de estas condiciones con los ejemplos que hemos venido manejando.

(A) El problema de Producción-Inventario cumple con la Condición C2.

Primeramente, del capitulo anterior sección 3, la Ecuación de Optimalidad α-descontada es:

$$V_\alpha(x) = \min_{a\in A(x)} E\,\{ca + h(x+a-\xi)^+ + \alpha V_\alpha(x+a-\xi)^+\}$$

y una politica α-óptima para este problema es:

$$f(x)=\begin{cases} S^*-x & \text{si } x\le S^* \\ 0 & \text{si } x>S^* \end{cases}$$

Además, $V_\alpha(x)$ es convexa y $x\in[0,M]$, $\xi\in[0,D]$.

En el capitulo II se probó C2(a) (hip. 3.3, II). La verificación de C2(b) y C2(c) se hará en dos casos:

(i) $x \le S^*$.

Sea $\bar{x} = 0$ el estado fijo en (3.3). De la EO α-descontada tenemos:

$$V_\alpha(x) = E\,\{c(S^*-x) + h(S^*-\xi)^+ + \alpha V_\alpha(S^*-\xi)^+\}$$

$$V_\alpha(0) = E\,\{cS^* + h(S^*-\xi)^+ + \alpha V_\alpha(S^*-\xi)^+\}$$

De aquí,

$$h_\alpha(x) = V_\alpha(x) - V_\alpha(0) = -cx$$

Como $x \in [0,M]$, entonces:

$$-cM \le -cx \le 0 \qquad (3.24)$$

Por otro lado, como $cM \ge 0$, tenemos que:

$$-cx \le cM-cx \qquad (3.25)$$

Ahora, tomando $N=cM$ y $b(x)=cM-cx \ge 0$, por (3.24) y (3.25), se cumple C2(b) y C2(c), y de aquí se cumple la condición C2.

(ii) $x > S^*$.

Nuevamente, sea $\bar{x}=0$ el estado fijo en (3.3). En este caso $V_\alpha(x)$ y $V_\alpha(0)$ son:

$$V_\alpha(x) = E\,\{\, h(x-\xi)^+ + \alpha V_\alpha(x-\xi)^+\}$$

$$V_\alpha(0) = E\,\{\, h(-\xi)^+ + \alpha V_\alpha(-\xi)^+\} = \alpha V_\alpha(0)$$

De aquí $V_\alpha(0) = 0$, y ahora,

$$h_\alpha(x) = hE\{(x-\xi)^+\} + E\{V_\alpha(x-\xi)^+\}$$

Observemos que $(x-\xi)^+ \le M$. Con esto,

$$h_\alpha(x) \le hM + \alpha E\{V_\alpha(x-\xi)^+\} \tag{3.26}$$

Por otro lado, es claro que $V_\alpha(x)$ es continua y $(x-\xi)^+$ pertenece al conjunto compacto $[0,M]$. De aquí, $V_\alpha(x-\xi)^+$ alcanza su máximo y su mínimo en $[0,M]$ y así podemos decir que $V_\alpha(y) \le B$ para toda $y \in [0,M]$ y alguna constante B. Entonces, de (3.26),

$$h_\alpha(x) \le hM + B \tag{3.27}$$

Por otro lado, tenemos que:

$$h_\alpha(x) \ge 0 \tag{3.28}$$

Definiendo $N := 0$ y $b(x) := hM + B \; \forall \; x \in [0,M]$, por (3.27) y (3.28) se cumple C2(b) y C2(c) y por lo tanto C2.

Así, por el Teorema 3.7, las conclusiones del Teorema 3.4 son válidas.

(B) El sistema LQ cumple con la condición C1.

La Ecuación de Optimalidad α-descontada es (enumerando nuevamente las ecuaciones para fácil referencia):

$$V_\alpha(x) = k(\alpha)x^2 + k(\alpha)\alpha\sigma^2/(1-\alpha) \quad \forall \; x \in R, \; \alpha \in (0,1), \tag{3.29}$$

donde $k(\alpha)$ es la única solución positiva de la ecuación:

$$k = [1-\alpha k\beta^2(r + \alpha k\beta^2)^{-1}]\alpha\kappa\gamma^2 + q \tag{3.30}$$

y

$$k(\alpha) \to k^\bullet \text{ cuando } \alpha \nearrow 1, \tag{3.31}$$

donde $k^\bullet$ es la única solución positiva de la ecuación que resulta cuando hacemos $\alpha\nearrow 1$ en (3.30).

Para ver que se cumple C1, solo falta probar C1(b) y C1(c). Para esto, tomamos $\bar{x}=0$ como el estado fijo en (3.3).

De (3.29) tenemos:

$$(1-\alpha)V_{\alpha}(0) = k(\alpha)\alpha\sigma^2 \qquad \forall\ \alpha\in(0,1) \tag{3.32}$$

y de aquí,

$$h_{\alpha}(x):=V_{\alpha}(x) - V_{\alpha}(0) = k(\alpha)x^2 \quad \forall\ x\in X,\ \alpha\in(0,1) \tag{3.33}$$

En vista de (3.31), sea $\varepsilon>0$ fijo pero arbitrario. Escogemos $\alpha_0\in(0,1)$ tal que:

$$k(\alpha) \le k^* + \varepsilon \qquad \forall\ \alpha\in[\alpha_0,1) \tag{3.34}$$

Por lo tanto, de (3.32), (3.33) y (3.34), obtenemos:

$$(1-\alpha)V_{\alpha}(0) \le (k^* + \varepsilon)\sigma^2$$

y

$$0 \le h_{\alpha}(x) \le (k^*+\varepsilon)x^2$$

Definiendo $M:=(k^*+\varepsilon)\sigma^2$, $b(x):=(k^*+\varepsilon)x^2$ y por estas dos expresiones, concluimos que se cumplen C1(b) y C1(c), y por lo tanto C1.

De hecho, el sistema LQ cumple con todas las condiciones como se muestra en [27], [20].

(C) El sistema de colas cumple con la condición C2.

La condición C2(a) se probó en el Capítulo II. Para las restantes, recordemos que la Ecuación de Optimalidad α-descontada para $x\neq 0$ es:

$$V_\alpha(x) = \min_e \left\{C(x,(e,1)) + \alpha\int_0^\infty\cdots\int_0^\infty V_\alpha(x+ey) \prod_{j\in E(e)} dF_j(y_j)\right\} \wedge$$

$$\min_e \min_k \min_{\beta(k)} \left\{C(x,(e,\beta(k))+\alpha\int_0^\infty\cdots\int_0^\infty V_\alpha(x(\beta(k))+ey) \prod_{j\in E(e)} dF_j(y_j)\right\}$$

Ahora consideremos las funciones de IV, $V_n(x)$ (3.7,II). En este caso tenemos:

$$v_n(x) = \min_e \left\{C(x,(e,1)) + \alpha\int_0^\infty\cdots\int_0^\infty v_{n-1}(x+ey) \prod_{j\in E(e)} dF_j(y_j)\right\} \wedge$$

$$\min_e \min_k \min_{\beta(k)} \left\{C(x,(e,\beta(k))+\alpha\int_0^\infty\cdots\int_0^\infty v_{n-1}(x(\beta(k))+ey) \prod_{j\in E(e)} dF_j(y_j)\right\}$$

Por la suposición de que $C(x,(e,1))$ y $C(x,(e,\beta(k)))$ son crecientes, se puede demostrar por inducción en n que $v_n(x)$ es creciente en cada coordenada del vector x con las otras fijas, y como $v_n \nearrow V_\alpha$ [Teo. 3.6(a),II], esto también se cumple para V_α. Así, $V_\alpha(x) \geq V_\alpha(0) \geq 0$, y definiendo $\bar{x}=0$ como el estado fijo en (3.3), esto implica que $h_\alpha(x) := V_\alpha(x) - V_\alpha(0) \geq 0$. Por lo tanto, con $N:=0$, se cumple C2(b).

Aparte, al igual que en la prueba de C2(a) [hip. 3.3,II], consideremos la política $\hat{f}$ que siempre escoge a $e^*=(0,0,\ldots,0)$ y sirve a los trabajos en las colas en el siguiente orden: la primera cola hasta que quede vacía, luego la segunda hasta que quede vacía, etc..

Como se vió en el Capítulo II, si el proceso empieza en un estado x, arbitrario, entonces el proceso alcanza el estado cero en un número finito de pasos con costo finito, digamos, $C_0(x)$.

Ahora, supongamos que después de que el proceso llegue al estado cero bajo la política $\hat{f}$, usamos una política α-óptima $\bar{f}$. Denotemos por δ a la política que resulta de combinar $\hat{f}$ y $\bar{f}$, y sea T el tiempo en que el proceso alcanza por primera vez el estado cero (notemos que T no es aleatorio). Como no podemos asegurar que δ sea α-óptima, entonces:

$$V_\alpha(x) \le E_x^\delta\left[\sum_{t=0}^{T-1} \alpha^t C(x_t, a_t)\right] + E_x^\delta\left[\sum_{t=T}^{\infty} \alpha^t C(x_t, a_t)\right] \le C_0(x) + V_\alpha(0)$$

De aquí, $h_\alpha(x) := V_\alpha(x) - V_\alpha(0) \le C_0(x)$ y como $C_0(x)$ es creciente se cumple C2(c). Por lo tanto, se cumple la condición C2 y por el Teorema 3.7 las conclusiones del Teorema 3.4 son válidas.

(D) Los problemas Invariantes cumplen con la condición C3.

Recordemos que los problemas invariantes cumplen que $Q(B/x,a) = Q(B/a)$ y además, $A=A(x)$ es finito $\forall\ x \in X$.

Claramente la condición C3(a) se cumple. Mostraremos C3(b) como sigue. Notemos primeramente que por el Teorema 3.6 del Cap. II, existe $\hat{f} \in F$ tal que:

$$V_\alpha(x) = C(x,\hat{f}) + \alpha\int V_\alpha(y) Q(dy/\ \hat{f})$$

En este caso, $\hat{f}$ no depende de x y además:

$$m_\alpha := \inf_x V_\alpha(x) = \alpha\int V_\alpha(y) Q(dy/\ \hat{f}) + \inf_x C(x,\hat{f})$$

Ahora,

$$g_\alpha(x) := V_\alpha(x) - m_\alpha = V_\alpha(x) - \inf_x V_\alpha(x)$$

$$= C(x,\hat{f}) + \alpha\int V_\alpha(y) Q(dy/\ \hat{f}) - \alpha\int V_\alpha(y) Q(dy/\ \hat{f}) - \inf C(x,\hat{f})$$

$$= C(x,\hat{f}) - \inf_x C(x,\hat{f}) \le C(x,\hat{f}) \le \max_a C(x,a) < +\infty \quad \forall\ \alpha \in (0,1),\ x \in X$$

Por lo tanto,

$$\sup_{\alpha} g_{\alpha}(x) < +\infty \qquad \forall\ x\in X$$

Lo cual prueba la condición C3.

En [35] se dan criterios para verificar la condición C3. Inclusive se muestra, usando estos criterios, que el problema de Producción-Inventario y los Problemas Invariantes cumplen la condición C3.

4.- Condiciones de optimalidad (II).

En esta sección trataremos la existencia de la CP-optimalidad bajo las hipótesis 2.2,II y C1. Presentaremos dos maneras de lograr esto, las cuales corresponden a pedir cada una de las siguientes hipótesis adicionales:

(i) $H_n(.) := \inf\limits_{m>n} h_{\alpha_m}(.)$ es sci.

(ii) **X** es un subconjunto abierto de Borel convexo de un espacio normado localmente compacto y $h_{\alpha_n}(.)$ es una función continua y convexa.

Antes de presentar el teorema principal de esta sección, daremos el siguiente Lema:

LEMA 4.1: Supóngase que se cumple (ii) y sea $\{\alpha(n)\}$ la sucesión en el Lema 3.3. Entonces existe, una subsucesión $\{\alpha(n)'\}$ de $\{\alpha(n)\}$, con $\alpha(n)'\nearrow 1$, y una función continua $h:X\to\mathbb{R}$ tal que:

$$h_{\alpha(n)'}(x)\to h(x) \qquad \forall\ x\in X$$

DEMOSTRACION:

Sea $x_0 \in X$ un estado arbitrario y $\varepsilon > 0$ suficientemente pequeño, tal que la bola cerrada $\bar{B}(x_0, 2\varepsilon)$, con centro en x_0 y radio 2ε, está contenida en X, i.e.,

$$\bar{B}(x_0, 2\varepsilon) := \{y' \in X : ||x_0 - y'|| \le 2\varepsilon\} \subset X$$

Sea x, y $\in B(x_0, \varepsilon) := \{y' \in X : ||x_0 - y|| < \varepsilon\}$ con $x \neq y$ y definamos:

$$z := x + \varepsilon(x-y)/||x-y|| \tag{4.1}$$

Ahora,

$$||x_0 - z|| = ||x_0 - x - \varepsilon(x-y)/||x-y|| \,|| \le ||x_0 - x|| + \varepsilon < 2\varepsilon$$

y así, $z \in B(x_0, 2\varepsilon)$. De (4.1),

$$||x-y||z = ||x-y||x + \varepsilon(x-y)$$

y despejando x,

$$x = z||x-y||/(||x-y||+\varepsilon) + y\varepsilon/(||x-y||+\varepsilon) \tag{4.2}$$

Como $y \in B(x_0, \varepsilon) \subset B(x_0, 2\varepsilon)$, por (4.2), x es una combinación convexa de z, y $\in B(x_0, 2\varepsilon)$.

Por otro lado tenemos que:

$$h_{\alpha(n)}(x) \le \sup_{\alpha(n)} h_{\alpha(n)}(x) := b'(x) \qquad \forall\ x \in X$$

Por (ii), la función $b'(x)$ es continua y además, como X es localmente compacto, alcanza su máximo en $\bar{B}(x_0, 2\varepsilon)$. De aquí, existe una constante $M(x_0, 2\varepsilon)$ tal que:

$$|h_{\alpha(n)}(y')| \le M(x_0, 2\varepsilon) \qquad \forall\ y' \in \bar{B}(x_0, 2\varepsilon) \tag{4.3}$$

y, por lo tanto, $\{h_{\alpha(n)}(x)\}$ es uniformemente acotada localmente.

Por otro lado, de la convexidad de $h_{\alpha(n)}(.)$, tenemos:

$$h_{\alpha(n)}(x) \le ||x-y||h_{\alpha(n)}(z)/(||x-y||+\varepsilon)+$$

$$+ \varepsilon h_{\alpha(n)}(y)/(||x-y||+\varepsilon) \qquad (4.4)$$

donde x es como en (4.2). De aquí, sumando y restando $||x-y||h_{\alpha(n)}(y)/(||x-y||+\varepsilon)$ y factorizando términos apropiados en el lado derecho de (4.4), nos queda:

$$h_{\alpha(n)}(x) \le ||x-y||h_{\alpha(n)}(z)/(||x-y||+\varepsilon)$$

$$-||x-y||h_{\alpha(n)}(y)/(||x-y||+\varepsilon) + h_{\alpha(n)}(y)$$

y de aquí,

$$h_{\alpha(n)}(x)-h_{\alpha(n)}(y) \le (||x-y||+\varepsilon)^{-1}||x-y||/\left[h_{\alpha(n)}(z)-h_{\alpha(n)}(y)\right]$$

$$\le (||x-y||+\varepsilon)^{-1}||x-y||/[2M(x_0,2\varepsilon)] \qquad \text{[por (4.3)]}$$

$$< 2M(x_0,2\varepsilon)/\varepsilon\ ||x-y|| \qquad (4.5)$$

Intercambiando y por x en el procedimiento anterior,

$$h_{\alpha(n)}(y) - h_{\alpha(n)}(x) < 2M(x_0,2\varepsilon)/\varepsilon\ ||x-y|| \qquad (4.6)$$

Así, de (4.5) y (4.6),

$$||h_{\alpha(n(}(y) - h_{\alpha(n)}(x)|| < 2M(x_0,3\varepsilon)/\varepsilon\ ||x-y||$$

y como x_0 fué arbitrario, la familia $\{h_{\alpha(n)}(.)\}$ es localmente equilipschitziana y en particular localmente equicontinua. De aquí y por el Teorema de Arzelá-Ascoli [33, pp. 169], existe una subsucesión $\{\alpha(n)'\}$, $\alpha(n)'\nearrow 1$, y una función continua $h:X\to\mathbb{R}$ tal que $h_{\alpha(n)'}(x)\to h(x)$, cuando $\alpha(n)'\nearrow 1$, para toda $x\in X$, y la convergencia es uniforme en subconjuntos compactos de X.|

TEOREMA 4.2: Supongamos que se cumplen las hipótesis 2.2,II y la condición C1. Sea $j^*\geq 0$ y $\{\alpha(n)\}$, $\alpha(n)\nearrow 1$, como se obtuvo en el Lema 3.3. Si una de las condiciones (i), (ii) se cumplen, entonces las conclusiones del Teorema 3.4 son válidas.

DEMOSTRACION:

Primero notemos que en los 2 casos, bajo (i) o bajo (ii), basta probar que existe una función medible h en X tal que j^* y h satisfacen la desigualdad (3.6). Con esto se prueba la parte (a) del Teorema 3.4; la parte (b) se sigue del Lema 2.6(a) del Capítulo II, y la parte (c) es consecuencia de la expresión (3.7).

CASO (i).

Como en la demostración del Teorema 3.4(a), podemos obtener la desigualdad (3.13):

$$j^*+h(x)+N+\varepsilon \geq C(x,a_{n_i})+\alpha(n_i)\int (H_L(y)+N)Q(dy/x,a_{n_i}) \quad \forall\ n_i>L$$

donde, $h(.)$ y $H_L(.)$ son como en (3.8) y $a_{n_i}\to a_x\in A(x)$ cuando $i\to+\infty$. Ahora, haciendo $i\to+\infty$, por las hipótesis 2.2,II y (i) junto con las propiedades de sci, obtenemos:

$$j^*+ h(x)+ N+ \varepsilon \geq C(x,a_x)+\int (H_L(y)+N)Q(dy/x,a_x)$$

y haciendo $L\to\infty$, por el Lema de Fatou,

$$j^*+ h(x)+ N+ \varepsilon \geq C(x,a_x)+\int (h(y)+N)Q(dy/x,a_x)$$

De aquí,

$$j^*+h(x)+\varepsilon \geq C(x,a_x)+\int h(y)Q(dy/x,a_x) \geq \min_a \left[C(x,a)+\int h(y)Q(dy/x,a)\right]$$

y como ε se tomó arbitrario, se demuestra lo que se quería.

CASO (ii).

Para no complicar la notación, denotemos por $\{\alpha(n)\}$ a la subsucesión $\{\alpha(n)'\}$ obtenida del lema 4.1 y sea $h:X\to\mathbb{R}$ la función contínua, tal que:

$$h_{\alpha(n)}(x)\to h(x) \qquad \forall\ x\in X \tag{4.7}$$

como en el Lema 4.1.

Por otro lado, como X es separable (por ser espacio de Borel) y localmente compacto, por [5] pp. 180, existe una sucesión de conjuntos abiertos G_ℓ, $\ell=1,2,\ldots,$ los cuales tienen cerradura compacta $\bar{G}_\ell$ y $G_\ell \nearrow X$.

Aparte, por (3.4), dado $\varepsilon>0$, existe $T>0$ tal que:

$$(1-\alpha(n))V_{\alpha(n)}(\bar{x}) \leq j^* + \varepsilon \qquad \forall\ n\geq T \tag{4.8}$$

Fijemos G_ℓ y un estado $x\in G_\ell$. Es claro que (4.7) se cumple en cada $\bar{G}_\ell$, así, tomando el mismo $\varepsilon>0$ que en (4.8),

$$h(x) - \varepsilon \leq h_{\alpha(n)}(x) \leq h(x) + \varepsilon \tag{4.9}$$

para toda $n\geq T$ y $x\in G_\ell$. De (4.8) y (4.9), obtenemos:

$$j^*+h(x)+2\varepsilon \geq (1-\alpha(n))V_{\alpha(n)}(\bar{x})+h_{\alpha(n)}(x) \qquad \forall\ n\geq T. \tag{4.10}$$

Del Teorema 3.6, para cada n existe $a_n\in A(x)$ tal que:

$$V_{\alpha(n)}(x) = C(x,a_n) + \alpha(n)\int V_{\alpha(n)}(y)Q(dy/x,a_n) \tag{4.11}$$

Por otro lado, para todo $n\geq T$,

$$(1-\alpha(n))V_{\alpha(n)}(\bar{x})+h_{\alpha(n)}(x)=$$

$$=(1-\alpha(n))V_{\alpha(n)}(\bar{x})+V_{\alpha(n)}(x) - V_{\alpha(n)}(\bar{x}) \quad [\text{por } (3.3)]$$

$$= V_{\alpha(n)}(x) - \alpha(n)V_{\alpha(n)}(\bar{x})$$

$$= C(x,a_n) + \alpha(n)\int h_{\alpha(n)}(y)Q(dy/x,a_n) \quad [\text{por } (3.3) \text{ y } (4.11)]$$

De aquí, junto con (4.10),

$$j^{*}+ h(x) + 2\varepsilon \geq C(x,a_n) + \alpha(n)\int h_{\alpha(n)}(y)Q(dy/x,a_n) \quad \forall\ n\geq T$$

Sumando $\alpha(n)N$ en ambos lados de la desigualdad, y como $\alpha(n)>1$, tenemos:

$$j^{*}+h(x)+N+2\varepsilon\geq C(x,a_n)+\alpha(n)\int [h_{\alpha(n)}(y)+N]Q(dy/x,a_n) \ \forall\ n\geq T \quad (4.12)$$

De la primera desigualdad en (4.9),

$$j^{*}+ h(x) +N + 2\varepsilon\geq C(x,a_n) + \alpha(n)\int [h(y)-\varepsilon+N]Q(dy/x,a_n) \quad \forall\ n\geq T$$

y de aquí, usando el hecho de que $h(.) + N \geq 0$,

$$j^{*}+h(x)+N+3\varepsilon\geq C(x,a_n)+\alpha(n)\int I_{G_\ell}(y)[h(y)+N]Q(dy/x,a_n) \quad (4.13)$$

donde I_{G_ℓ} es la función indicadora de G_ℓ.

Ahora definamos los conjuntos:

$$B_n(x)=\{a\in A(x):C(x,a)+\alpha(n)\int I_{G_\ell}(y)[h(y)+N]Q(dy/x,a)\leq j^{*}+h(x)+N+3\varepsilon\}$$

para $n=1,2,\ldots$. Por los mismos argumentos de la demostración del Teorema 3.4(a), tenemos que $B_n(x)$ es no vacío y compacto para cada n, y $B_n(x)\searrow B(x)$, donde:

$$B(x)=\{a\in A(x):\ C(x,a)+ \int I_{G_\ell}(y)[h(y)+N]Q(dy/x,a)\leq j^{*}+h(x)+N+3\varepsilon\}$$

y $B(x)$ es no vacío. Por lo tanto, existe una subsucesión

$\{a(n_i)\}$ de $\{a(n)\}$ y $a_x \in B(x)$, tal que $a(n_i) \to a_x$ cuando $i \to \infty$. De (4.13), con $n=n_i$, tenemos:

$$j^* + h(x) + N + 3\varepsilon \geq C(x, a_{n_i}) + \alpha(n_i) \int I_{G_\ell}(y)[h(y)+N]Q(dy/x, a_{n_i}) \quad \forall\ n_i \geq T$$

Como $I_{G_\ell}(.)(h(.)+N)$ es sci, haciendo $i \to \infty$, por Prop. 3, Ap.A,

$$j^* + h(x) + N + 3\varepsilon \geq C(x, a_x) + \int I_{G_\ell}(y)[h(y)+N]Q(dy/x, a_x) \quad (4.14)$$

Ahora, haciendo $\ell \to \infty$ y usando el hecho de que $G_\ell \nearrow X$,

$$j^* + h(x) + N + 3\varepsilon \geq C(x, a_x) + \int (h(y)+N)Q(dy/x, a_x)$$

y cuando $\varepsilon \to 0$,

$$j^* + h(x) + N \geq C(x, a_x) + \int (h(y)+N)Q(dy/x, a_x).$$

Concluyendo como en la prueba del Teorema 3.4 o la del caso (i) se demuestra el caso (ii), y por lo tanto el Teorema 4.2. |

De los ejemplos que hemos venido manejando, el de Producción-Inventario y el LQ cumplen con (i) y (ii).

Notemos en (ii), que si X es abierto y $h_{\alpha n}(.)$ es convexa, entonces es continua. En [21] se dan condiciones para que la función $V_\alpha(x)$ sea convexa, y por lo tanto se da la convexidad de $h_\alpha(x)$.

CONCLUSIONES

En este trabajo, se revisaron condiciones recientes de [16, 20, 27] para obtener la optimalidad en costo promedio de una política de control. Para esto, primero se revisaron los elementos suficientes para plantear el problema de control óptimo y definimos el modelo de control de markov.

Se impusieron restricciones al modelo de control más débiles que las que se pedían en trabajos anteriores [15,24,34, etc.]. Como podemos ver en las hipótesis 2.2,II y 2.2,III, pedimos que los espacios de estados y de control sean de Borel. Además, en lugar de pedir que los conjuntos de acciones admisibles para un estado x, $A(x)$, sean compactos para cada x, pedimos que los conjuntos $A_r(x):=\{a\in A(x) : C(x,a)\le r\}$ sean compactos para toda $x\in X$, $r\in\mathbb{R}$, i.e., $C(x,a)$ es inf-compacta, y con costos posiblemente no acotados. Lo anterior nos permite generalizar resultados anteriores [35,37].

En nuestro caso, para obtener la CP-optimalidad, nos basamos en la teoría de los problemas de control con costos descontados. Esto es, analizamos el caso con costo promedio, por medio del enfoque del "factor de descuento desvaneciente".

La relación entre el caso promedio y el caso descontado la obtuvimos basándonos en un Teorema Tauberiano y sus consecuencias [Lema 3.2 Cap. III].

Existen otros caminos para obtener la CP-optimalidad, como por ejemplo, imponiendo condiciones de ergodicidad [15,19]. Pero el analizar estos problemas vía el caso descontado es de gran ventaja por el hecho de que aquí tenemos una teoría mas completa, inclusive en el caso general tenemos una solución, completamente determinada, de la Ecuación de Optimalidad α-descontada. Como se puede ver, las condiciones de optimalidad para el caso promedio están básicamente en la función de valor óptimo α-descontada $V_\alpha(x)$, excepto en el caso de la condición C3, donde se pide que el índice de funcionamiento para el caso promedio sea finito para alguna política $\pi\in\Pi$ y $x\in X$.

Por otro lado, a diferencia del caso descontado, cuando estamos trabajando con hipótesis generales como las 2.2,II y 2.1,III, en el caso promedio no tenemos una solución de la ecuación de optimalidad respectiva; en otros casos (por ejemplo, pidiendo que los costos sean acotados, el espacio de estados numerable y/o el conjunto de acciones finito), sí se tiene una solución de la ecuación de optimalidad [8,19,32,36, etc.].

En ambos casos (CD, CP), el tener igualdad en la EO, es de gran ventaja, ya que se pueden implementar métodos de aproximación a la función de valor óptimo respectiva. Por ejemplo, para el caso CD, podemos usar las funciones de iteración de valores (IV) definidas en la sección 3 del capítulo II. A este esquema de aproximación se le conoce como Método de Iteración de Valores o de Aproximaciones Sucesivas.

Las funciones IV están definidas recursivamente, y es importante notar, que dicha aproximación se hace por medio de problemas con un número finito de etapas, de los cuales existe bastante literatura al respecto [8,32, etc.]. Otros métodos de aproximación para los problemas con costos descontados, bajo las hipótesis 2.2 Cap. II, son analizados en [22].

Para los problemas con costo promedio, el método de Iteración de Valores en el caso de X y A finitos, es muy similar al del caso descontado. En [8] Cap VII, se hace un análisis de esta técnica, inclusive se obtienen cotas del error para la función de valor óptimo. También se da otro método de aproximación, el llamado método de Iteración de Políticas, el cual también se estudia, para el caso descontado, en [22] bajo hipótesis generales.

En el marco en que se desarrolló este trabajo, no es muy directa la implementación de métodos de aproximación debido a que no se cuenta con una solución de la Ecuación de Optimalidad en Costo Promedio, solo podemos asegurar la existencia de políticas CP-óptimas bajo las condiciones de optimalidad C1, C2 y C3. Estas dos últimas, fueron introducidas inicialmente en [37] y [35] respectivamente, y C1 en [27]. No se conoce ningún trabajo donde se obtenga la igualdad en la ecuación de optimalidad bajo estas condiciones; por lo regular se piden adicionales. Por ejemplo, en [14], se obtiene una solución de la EO bajo las siguientes condiciones: hipótesis 2.1,III, $\mathbf{X}$ subconjunto abierto de $\mathbb{R}^n$, condición C2 con $b(x)$ semicontinua superiormente (scs) y $V_{\alpha n}(x)$ convexa $\forall\ x\in\mathbf{X}$ y αn es como en el Lema 3.3.

Notemos que en este caso, los problemas de Producción-Inventario y el LQ, cumplen con estas condiciones y por lo tanto podemos asegurar que existen soluciones a la ecuación de optimalidad respectiva.

Otro resultado, que extiende el anterior, es el que se da en [28]. Aquí, para obtener una solución de la EO, se piden las siguientes condiciones: hipoptesis 2.2,II, $\mathbf{X}$ localmente compacto, condicion C1 y $h_{\alpha n}$ una familia equicontinua.

Como podemos ver en la demostración del Lema 4.1 del Capítulo III, la convexidad de las $h_{\alpha n}$ implica que $h_{\alpha n}$ sea localmente equicontinua, además, como C2 implica C1 y las hipótesis 2.1,III son suficientes para las hipótesis 2.2,II, las condiciones en [28] son mas débiles que las condiciones en [14]. Un ejemplo donde se ilustra esto último se da precisamente en [28] y es el siguiente:

-Sea $\mathbf{X}=\mathbf{A}=\mathbb{R}$, $\mathbf{A}(x)=\{a\}$, $\forall\ x\in\mathbf{X}$. La función de costo es:

$$C(x,a)=\begin{cases} x^{1/2} & \text{si } x\geq 0 \\ 0 & \text{si } x<0 \end{cases}$$

y la ley de transición es $Q(\{0\}/x,a)=1\ \forall\ x\in\mathbf{X}$. Como $\mathbf{A}(x)$ consta de un solo punto, la EO α-descontada es:

$$V_\alpha(x) = C(x,a) + \alpha\int V_\alpha(y)Q(dy/x,a) = C(x,a) + \alpha V_\alpha(0)$$

Tomando x=0, tenemos que $V_\alpha(0)=\alpha V_\alpha(0)$, y por lo tanto $V_\alpha(0)=0$. Entonces $V_\alpha(x)=C(x,a)$.

Sea $\bar{x}$ un estado fijo y $\alpha(n)\nearrow 1$, entonces:

$$h_{\alpha(n)}(x) := V_{\alpha(n)}(x) - V_{\alpha(n)}(0) = C(x,a)$$

Así, por la continuidad de la función C la familia $\{h_{\alpha(n)}\}$ es equicontinua. Es claro que C no es convexa y se puede probar que la familia $\{h_{\alpha(n)}\}$ no es equilipschitz localmente.|

Como podemos observar, las condiciones C1, C2, y C3, juegan un papel importante para garantizar la CP-optimalidad. Dichas condiciones son la base para extender estos resultados a otros problemas de control, como por ejemplo al caso semi-markoviano [40]. En dicho trabajo se extiende la condición C3 al caso semi-markoviano, pero de igual forma se obtiene una desigualdad de optimalidad. Un interesante trabajo puede ser extender las condiciones de [14] o [28] a este caso y así, obtener una solución a la ecuación de optimalidad correspondiente.

Por otro lado, en el marco de los trabajos [14] y [28], no se conocen métodos de aproximación a la función de valor óptimo en costo promedio, lo cual abre la posibilidad de trabajar en este sentido. Otro trabajo, también interesante, sería extender las condiciones que aquí se trabajaron, o inclusive las de [14] y [28], al caso de Procesos de Control de Markov Adaptables [15].

Por último, haremos un comentario de la bibliografía que se utilizó. Los elementos de la Teoría de Control estocástico se revisaron, esencialmente en [8,9,15]. El articulo que nos sirvió de base para el Capitulo II fué [22] y para el Capítulo III fueron [14], [20] y [27]. El resto de las referencias nos sirvieron para complementar la teoría y los artículos base.

A P E N D I C E

A.- ESPACIOS, FUNCIONES Y MULTIFUCIONES.

A un espacio topológico siempre lo dotaremos de la σ-álgebra de Borel $\mathcal{B}(X)$, i.e., la σ-álgebra más pequeña de subconjuntos de X que contiene todos los conjuntos abiertos de X.

DEFINICION 1: Un espacio de Borel es un subconjunto de un espacio métrico, completo y separable.

Un subconjunto de Borel de un espacio de borel es también un espacio de Borel. Algunos ejemplos de espacios de Borel son:

- $\mathbb{R}^n$ con la topología usual.
- Un conjunto numerable con la topología discreta.
- Un espacio métrico compacto.
- El producto, finito o numerable, de una sucesión de espacios de Borel.

DEFINICION 2: Sea X un espacio topológico y v una función definida en X tomando valores en los reales extendidos. La función v se dice que es:

(a) Semicontinua inferiormente (sci) si el conjunto $\{x\in X: v(x)\le r\}$ es cerrado en X para cada $r\in\mathbb{R}$.
(b) Semicontinua superiormente (sci) si el conjunto $\{x\in X: v(x)\ge r\}$ es cerrado en X para cada $r\in\mathbb{R}$.

Notemos que v es SCI si y solo si -v es SCS. Además, v es continua si y solo si v es SCI y SCS.

Denotaremos por L(X) al conjunto de todas las funciones SCI definidas en X y acotadas por abajo.

PROPOSICION 3: Sea **X** un espacio métrico y v una función definida en **X** tomando valores en los reales extendidos. Entonces:

(a) v es SCI si y solo si para cada sucesión $\{x_n\}$ en **X**, tal que $x_n \to x \in \mathbf{X}$, tenemos que:

$$\liminf_n v(x_n) \geq v(x)$$

(b) $v \in L(\mathbf{X})$ si y solo si existe una sucesión de funciones contínuas y acotadas v_n tal que $v_n \nearrow v$.

DEMOSTRACION: Ash(1972) Apéndice A6.

PROPOSICION 4: Si v, v_1, $v_2, \ldots, v_n$ pertenecen a $L(\mathbf{X})$, entonces:

(a) $\alpha v(.) \in L(\mathbf{X})$ para cualquier $\alpha > 0$.
(b) $v_1 + v_2 + \ldots + v_n$ y $\min_i v_i$ pertenecen a $L(\mathbf{X})$.

DEMOSTRACION: Ash(1972) apéndice A6.

Sean **X** y **A** espacios de Borel.

DEFINICION 5: Una multifunción o correspondencia φ de **X** a **A** es una función sobre **X** cuyo valor $\varphi(x)$, para cada $x \in \mathbf{X}$, es un subconjunto no vacío de **A**.

DEFINICION 6: Decimos que φ es sci, si para cada sucesión $\{x_n\}$ en **X** y $a \in \mathbf{A}(x)$, tal que $x_n \to x \in \mathbf{X}$, existe $a_n \in \mathbf{A}(x_n)$ tal que $a_n \to a$.

DEFINICION 7: Decimos que φ es scs, si para todo conjunto abierto $G \subset \mathbf{A}$, el conjunto $\{x : \mathbf{A}(x) \subset G\}$ es abierto en **X**.

B.- KERNELES ESTOCASTICOS Y ESPERANZA CONDICIONAL.

Sean **X** y **Z** espacios de Borel.

DEFINICION 1: Un kernel estocástico sobre **X** dado **Z** es una función Q(./.), o Q(dx/z), tal que:

(a) Q(./z) es una medida de probabilidad en **X** para cada $z \in \mathbf{Z}$, y

(b) Q(B/.) es una función medible sobre **Z** para cada $B \in \mathcal{B}(\mathbf{X})$.

DEFINICION 2: Sea Q(dx/z) un kernel estocástico sobre **X** dado **Z**. Decimos que:

(a) Q es fuertemente continuo si la función:

$$z \to \int v(x)Q(dx/z) \tag{B.1}$$

es acotada y continua para cada función medible y acotada v.

(b) Q es debilmente continuo si la función en (B.1) es acotada y continua para cada función continua y acotada v.

(c) Q es sci si la función en (B.1) es sci para cada $v \in L(\mathbf{X})$.

PROPOSICION 3: Sea Q(dx/z) un kernel estocástico en **X** dado **Z**. Entonces:

$$(a) \Rightarrow (b) \Leftrightarrow (c)$$

La demostración de la Proposición 3 se sigue directamente de las definiciones.

DEFINICION 4: Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, y sea $\mathcal{G} \subseteq \mathcal{F}$ una sub σ-álgebra de $\mathcal{F}$. La esperanza condicional de la variable aleatoria X, $E|X|<\infty$, dado $\mathcal{G}$, se define como la única c.s. variable aleatoria denotada por $E[X/\mathcal{G}]$ tal que sea $\mathcal{G}$-medible y que:

$$\int_A E[X/\mathcal{G}]dP = \int_A XdP \qquad \forall\ A\in\mathcal{G}$$

Algunas propiedades de la esperanza condicional son:
Sean X, Y variables aleatorias; a, b, c números reales y $(\Omega,\mathcal{F},P)$ un espacio de probabilidad. Entonces:

B1.- $E\,[aX+bY+c/\mathcal{G}] = aE\,[X/\mathcal{G}] + bE\,[Y/\mathcal{G}] + c.$ c.s.

B2.- Si $X\le Y$ c.s. $E\,[X/\mathcal{G}] \le E[Y/\mathcal{G}]$ c.s.

B3.- Si $X_n \ge 0$ para toda n y $X_n \nearrow X$ c.s., entonces

$$E\,[X_n/\mathcal{G}] \nearrow E\,[X/\mathcal{G}].$$

B4.- $E\,[I_B/G] = P(B/G)$, donde I_B es la función indicadora del conjunto B.

B5.- Sean $\mathcal{F}_1$, $\mathcal{F}_2$ dos sub σ-álgebras de $\mathcal{F}$ tales que $\mathcal{F}_1 \subset \mathcal{F}_2 \subset F$. Entonces:

$$E\,[\,E\,[X/\mathcal{F}_1]/\mathcal{F}_2\,] = E\,[X/\mathcal{F}_1\,]$$

$$E\,[\,E\,[X/\mathcal{F}_2]/\mathcal{F}_1\,] = E\,[X/\mathcal{F}_1\,]$$

B6.- Si X es G-medible $E\,[X/G] = X$.

B7.- $E\,[\,E\,[X/G]\,] = E(X)$.

Las demostraciones de todas estas propiedades se pueden encontrar en Ash(1972) Capítulo 6.

C.- TEOREMA TAUBERIANO.

TEOREMA 1: Sea $\{C_n\}$ una sucesión de números reales no negativos y $\alpha\in(0,1)$. Entonces:

$$\liminf_{N\to\infty} N^{-1}\sum_{m=0}^{N-1} C_m \le \liminf_{\alpha\nearrow 1} (1-\alpha)\sum_{n=0}^{\infty} \alpha^n C_n$$

$$\le \limsup_{\alpha\nearrow 1} (1-\alpha)\sum_{n=0}^{\infty} \alpha^n C_n \le \limsup_{N\to\infty} N^{-1}\sum_{m=0}^{N-1} C_m$$

La demostración de este Teorema se puede encontrar en [36].

BIBLIOGRAFIA

[1] Arapostathis, A., Borkar, V., Fernandez-Gaucherand, E., Ghosh, M. K. and Marcus S. I., Discrete-time controlled Markov processes with average cost criterion: a survey. SIAM J. Control Optim., 31(1993), 282-344.

[2] Arkin, V. I., Extension of the class of Markov controls. Lectures Notes Math., 1091 (1984), 55-65.

[3] Ash, R. B., Real Analysis and Probability. Academic Press, New York (1972).

[4] Assaf, D., Invariant problems in dynamic programming -average reward criterion. Stoch. Processes Appl., 10 (1980), 313-322.

[5] Aubin, J. P., Applied Abstract Analysis. John Wiley and Sons, Inc. (1977).

[6] Bensoussan, A., Stochastic control in discrete time and applications to the theory of production. Math. Program. Study, 18 (1982), 43-60.

[7] Bhattacharya, R. N. and Majumdar, M., Stochastic models in mathematical economics: A review. In Statistics: Applications and New Directions (J. K. Ghosh, G. Kallianpur and J. Roy, Eds.), Eka Press, Calcuta (1984), 55-99.

[8] Bertsekas, D. P., Dynamic Programming and Stochastic Control. Academic Press (1976).

[9] Bertsekas, D. P. and Shreve, S. E., Stochastic Optimal Control: The Discrete Time Case. Academic Press (1978).

[10] Cavazos-Cadena, R., A counterexample on the optimality equation in Markov decision chains with the average cost criterion. Syst. and Control Lett., 16 (1991), 387-392.

[11] Cavazos-Cadena, R., Solution to the optimality equation in a class of average Markov decision chains with unbounded cost. Kybernetika (Prague), 27(1991), 23-37.

[12] Cavazos-Cadena, R. and Senott, L. I., Comparig recent assumptions for the existence of average optimal stationary policies. Oper. Res. Lett. 11(1992), 33-37.

[13] Dynkin, E. B. and Yushkevich, A. A., Controlled Markov Processes. Springer-Verlag (1979).

[14] Fernandez-Gaucherand, E., Marcus, S. I. and Arapostathis, A., Convex stochastic control problems. SIE Working Paper #92-006, Systems and Industrial Eng. Dept., University of Arizona, February 1992.

[15] Hernández-Lerma, O., Adaptive Markov Control Processes. Springer-Verlag (1989).

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

[17] Hernández-Lerma, O. and Lasserre, J. B., Error bounds for rolling horizon policies in discrete-time Markov control processes. IEEE Trans. Automat. Control, 35 (1990), 1118-1124.

[18] Hernández-Lerma, O., Lectures Notes on Discrete-Time Markov Control Processes. IV CLAPEM (1990).

[19] Hernández-Lerma, O., Montes de Oca, R. and Cavazos-Cadena, R., Recurrence conditions for Markov decision processes with Borel state space: a survey. Ann. Oper. Res. Special Volume on MDP's (Hrnández-Lerma, O. and Lasserre, J. B., Eds.), 28 (1991), 29-46.

[20] Hernández-Lerma, O., Average optimality in dynamic programming on Borel spaces unbounded cost. Syst. Control Lett., 17 (1991), 237-242.

[21] Hernández-Lerma, O. and Runggaldier W., Monotone approximations for convex stochastic control problem. J. Math. Syst., Estimation and Control. (En prensa).

[22] Hernández-Lerma, O and Muñoz M., Discrete-time Markov control processes with discounted unbounded cost: optimality criteria. Kybernetika (Prague), 28 (1992), 191-212.

[23] Heyman, D. P. and Sobel, M. J., Stochastic Models in Operations Research, Vol. 2, Mc Graw-Hill (1984).

[24] Himmelberg, C. J., Parthasarathy, T. and Van Vleck, F. S.,Optimal plans for dynamic programming problems. Math. Oper. Res., 1 (1976), 390-394.

[25] Hinderer, K., Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter. Lecture Notes Oper. Res., 33, Springer Verlag New York (1970).

[26] Hordijk, A., A sufficient condition for the existence of an optimal policy with respect to the average cost criterion in markovian decision processes. Trans. 6th. Prague Conf. on Information Theory, (1973), 263-274.

[27] Montes de Oca, R. and Hernández-Lerma, O., Conditions for average optimality in Markov control processes with unbounded cost and controls. Syst., Estimation and Control. (En prensa).

[28] Montes de Oca, R., The average cost optimality equation for Markov control processes on Borel spaces. (Sometido para su publicación).

[29] Prawda, J., Métodos y Modelos de Investigación de Operaciones, Vol. 2, Modelos Estocásticos, Limusa (1991).

[30] Rieder U., Measurable selection theorems for optimization problems. Manuscripta Math., 24 (1978), 115-131.

[31] Ritt, R., K. and Sennott, L. I., Optimal stationary policies in general state space Markov decision chains with finite action sets. (sometido para publicación).

[32] Ross, S. M., Introducction to Stochastic Dynamic Programming. Academic Press (1963).

[33] Royden, H. L., Real Analysis. Maxwell Macmillan Int. (1989).

[34] Schal, M., Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal. Z. Wattrs. Verw. Geb., 32 (1975), 179-196.

[35] Schal, M., Average optimality in dynamic programming with general state space, (en prensa), Institut F. Angewandte Mathematik, Universitat Bonn, Wegelerstr. 6, D-5300, FRG (1990).

[36] Sennot, L. I., A new condition for the existence of optimun stationary policies in average cost Markov decision processes unbounded case. Proc. 25th IEEE Conf. Decision and Control (Athens, Grece, Dec. 1986) 1719-1721.

[37] Sennot, L. I., Average cost optimal stationary policies in infinite state Markov decision processes with unbouded cost. Oper. Res., 37 (1989), 626-633.

[38] Sobel, M. J., Reservoirs management models. Water Resources Research, Vol. 11, 6 (1975), 767-776.

[39] Strauch, R. E., Negative dynamic programming. Ann. Math. Statist., 37 (1966), 871-890.

[40] Vega-Amaya, O., Average optimality in semi-markov control models on Borel spaces: Unbounded cost and controls. (sometido para publicación).

[41] Yakowitz, S., Dynamic programming applications in water resources. Water Resources Research, Vol. 18, 4 (1982), 673-696.