Benemérita Universidad Autónoma de Puebla

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS
LICENCIATURA EN MATEMÁTICAS

CARACTERIZACIÓN DE ESTRATEGIAS UMBRAL EN PROCESOS DE DECISIÓN DE MARKOV

TESIS

QUE PARA OBTENER EL TÍTULO DE
LICENCIADO EN MATEMÁTICAS

PRESENTA
JULIETA DEL ROSARIO SEVILLA BRAMBILA

DIRECTOR DE TESIS
DR. HUGO ADÁN CRUZ SUÁREZ

PUEBLA, PUE. 20 DE JUNIO DE 2014

En memoria de mi padre, Carlos.

Agradecimientos

Quisiera usar este espacio para agradecer a las personas que me han acompañado a lo largo de esta etapa.
Primeramente a mis padres, Faviola y Carlos, pues sin su respaldo y paciencia no hubiera llegado hasta aquí, principalmente a mi madre que en todo momento intenta mostrarme el lado positivo de las cosas, a pesar de la adversidad. A mis hermanas, Cinthia y Katya, que aunque la distancia física nos separa, siempre están ahí para recordarme el porqué de esta travesía.

A mis amigos que me han acompañado a lo largo de la carrera aún sin estar relacionados con ella.

A cada uno de los miembros de la LEJ, que han sido mi familia fuera de casa.

A mi asesor por su guía y apoyo para la elaboración de este trabajo. A mis sinodales por tomarse el tiempo de revisar y haber contribuido con sus sabias correcciones.

A la Vicerrectoría de Investigación y Estudios de Posgrado (VIEP), por el apoyo en la elaboración de esta tesis.

Índice general

Notación

I_B función indicadora de un conjunto B, definida por

$$I_B(x) = \begin{cases} 1, & \text{si} \quad x \in B, \\ 0, & \text{en otro caso.} \end{cases}$$

$(x)^+$ representa la parte positiva de x

$$(x)^+ = \begin{cases} x, & \text{si} \quad x > 0, \\ 0, & \text{en otro caso.} \end{cases}$$

$s.c.i.$ semicontinua inferior

$s.c.s.$ semicontinua superior

$\mathbb{F}$ Definición B.6

Φ Definición A.3

$\mathbb{K}$ Definición B.1

Π Definición 1.2

INTRODUCCIÓN

Esta tesis, pertenece al área de Procesos Estocásticos y Teoría de Control, específicamente a los Procesos de Decisión de Markov (*ver* [10]), que son procesos estocásticos a tiempo discreto y qué, como su nombre lo dice, cuentan con la propiedad de Markov; centrándonos en su aplicación en teoría de inventarios y líneas de espera (*ver* [9] y [11]). Estos Procesos también son utilizados en otras áreas como finanzas, planeación de producción, ingeniería, biología y estadística (*ver* [4]).

En los Procesos de Decisión de Markov, es de gran interés poder caracterizar las políticas óptimas, es decir, dar una forma explícita de la estrategia que minimice los costos o maximice las ganancias, dependiendo el caso.

En la tesis nos enfocamos en un tipo particular de políticas, llamadas estrategias umbral (*ver* [9]), que ya cuentan con una caracterización previa. En este tipo de estrategias se elige entre dos acciones posibles, la decisión se toma en base a la observación de una variable que caracteriza el estado del sistema. Nuestro objetivo es, bajo ciertas suposiciones en el modelo de control, poder asegurar la existencia de políticas umbral óptimas.

Las estrategias umbral se pueden encontrar en la teoría de inventarios y de líneas de espera (*ver* [12] *y* [14]) . En ambos casos lo que nos interesa es reducir los costos, ya sea de almacenaje y demanda no cumplida en el caso de inventarios o de servicio y espera de clientes en el caso de líneas de espera. Por la frecuencia con la que las estrategias umbral aparecen en ambas teorías, en la tesis mencionamos la estructura básica de modelos de colas e inventarios.

La tesis se basa en el reporte técnico *On the optimality of base-stock policies for a class of inventory systems with no set-up cost and no backorders* de Óscar Vega-Amaya y Joaquín López Borbón (*ver* [16]), en el cual se encuentra el Teorema que nos asegura la existencia de una política umbral óptima, revisando el reporte, notamos que la hipótesis de estrictamente no acotada que se tiene sobre la función de costos, se puede reemplazar, así que la cambiamos por una hipótesis más débil, la de inf-compacidad sobre una función que nos ayuda a definir el costo, haciendo esto aún se garantiza la conclusión del Teorema. También se planteó un ejemplo que cumple las hipótesis necesarias para asegurar la existencia de una política umbral, éste se puede adaptar a inventarios y líneas de espera. Finalmente, el ejemplo planteado lo adaptamos a ambos casos e hicimos algunas simulaciones para mostrar de manera numérica que la implementación de dichas estrategias ayuda a la reducción de costos.

La tesis está estructurada de la siguiente manera:

En el Capítulo 1, definiremos conceptos que son de suma importancia para el desarrollo y entendimiento de los capítulos posteriores, abarcando desde Modelos de Control y la Propiedad de Markov hasta conceptos de Programación Dinámica en horizonte finito e infinito.

Teniendo el concepto de política, definido en el Capítulo 1, en el Capítulo 2 se definirá lo que es una política umbral, seguido de ésto daremos tres casos, con especificaciones sobre los espacios de estados, acciones y acciones admisibles, la función de costo y la distribución de la variable del modelo, que nos serán suficientes para asegurar la existencia de una política umbral óptima. Finalmente, se enunciará y probará el teorema que nos garantiza la existencia de dicha política para los tres casos mencionados.

En el Capítulo 3, daremos algunos conceptos básicos de dos temas donde se pueden aplicar las políticas umbral, la teoría de inventarios y la de líneas de espera. A continuación, se dará un ejemplo que cumple las hipótesis del teorema que garantiza la existencia de políticas umbral y sus adaptaciones a inventarios y líneas de espera.

En el Capítulo 4, se harán simulaciones en R de sistemas de líneas de espera, para el ejemplo planteado en el capítulo anterior, se harán en el caso controlado, aplicando las políticas obtenidas, y en el caso sin controlar, ésto con el objetivo de mostrar la reducción de costos del sistema al usar estrategias umbral.

Finalmente se darán las conclusiones sobre la utilidad de las estategias umbral, y se planteará posible trabajo futuro en esta línea.

Capítulo 1

PRELIMINARES

1.1. Modelo de Control de Markov

Definición 1.1. Un modelo de control de Markov (MCM), es una quíntupla de la forma:

$$(X,\ A,\ \{A(x) : x \in X\},\ Q,\ c)$$

donde

1. X es un espacio de Borel, llamado espacio de estados.

2. A es un espacio de Borel, llamado espacio de acciones o controles.

3. $\{A(x) : x \in X\}$ es una familia de subconjuntos medibles de A, donde $A(x)$ denota el conjunto de acciones o controles admisibles cuando el sistema se encuentra en el estado $x \in X$, y tienen la propiedad de que el conjunto
$$\mathbb{K} := \{(x, a) |\ x \in X,\ a \in A(x)\}$$
de parejas estado-acción admisible es un subconjunto medible del espacio $X \times A$.

4. Q es un kérnel estocástico (*ver Apéndice A*) sobre X dado $\mathbb{K}$, llamado ley de transición.

5. $c : \mathbb{K} \to \mathbb{R}$ es una función medible llamada función de costo de un paso.

Considere un modelo de control de Markov y, para cada $t = 0, 1, \ldots$, definimos el espacio de historias admisibles hasta el tiempo t, $\mathbb{H}_t$, como

$$\begin{aligned} \mathbb{H}_0 &:= X \\ \mathbb{H}_t &:= \mathbb{K}^t \times X \\ &= \mathbb{K} \times \mathbb{H}_{t-1}, \qquad \text{para } t = 1, 2, \ldots. \end{aligned}$$

Un elemento h_t de $\mathbb{H}_t$ llamado t-historia admisible o simplemente t-historia, es de la forma

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

con $(x_i, a_i) \in \mathbb{K}$ para $i = 0, \ldots, t-1$ y $x_t \in X$. Observe que para cada t, $\mathbb{H}_t$ es un subespacio de

$$\begin{aligned} \overline{\mathbb{H}}_t &:= (X \times A)^t \times X = (X \times A) \times \overline{\mathbb{H}}_{t-1} \\ y \quad \overline{\mathbb{H}}_0 &:= \mathbb{H}_0 = X. \end{aligned}$$

1.2. Políticas

Definición 1.2. Una política de control aleatorizada o política de control, es una sucesión $\pi = \{\pi_t, \quad t = 0, 1, \ldots\}$ de kérneles estocásticos π_t sobre el conjunto de acciones A dada $\mathbb{H}_t$, que satisface

$$\pi_t(A(x_t)|h_t) = 1, \quad \forall h_t \in \mathbb{H}_t, \;\; t = 0, 1, \ldots.$$

El conjunto de todas las políticas es denotado por Π.

Definición 1.3. Una política se dice:

- **Markoviana**, si existe una sucesión $\{\psi_t\}$ de kérneles estocásticos $\psi_t \in \Phi$, tales que

  $$\pi_t(\cdot \mid h_t) = \psi_t(\cdot \mid x_t), \qquad \text{para todo } h_t \in \mathbb{H}_t,\ t = 0, 1, \ldots.$$

  Al conjunto de políticas markovianas se le denotará con Π_{RM}.

- **Estacionaria**, si existe $\psi \in \Phi$, tal que

$$\pi_t(\cdot \mid h_t) = \psi(\cdot \mid x_t), \qquad \text{para todo } h_t \in \mathbb{H}_t,\ t = 0, 1, \ldots.$$

  Este conjunto de políticas es denotado por Π_{RS}.

Más aún, $\pi = \{\pi_t\}$ se dice

- **Determinista** (o pura), si existe una sucesión $\{g_t\}$ de funciones medibles $g_t : \mathbb{H}_t \to A$ tal que, para toda $h_t \in \mathbb{H}_t$ y $t = 0, 1, \ldots$, $g_t(h_t) \in A(x_t)$ y $\pi_t(\cdot \mid h_t)$ está concentrada en $g_t(h_t)$, i.e.

$$\pi_t(C| \, h_t) = I_C[g_t(h_t)], \qquad \text{para toda } C \in \mathbb{B}(A).$$

  A este conjunto de políticas se le denota por Π_D.

- **Determinista Markoviana**, si existe una sucesión $\{f_t\}$ de funciones $f_t \in \mathbb{F}$ tal que $\pi_t(\cdot \mid h_t)$ está concentrada en $f_t(x_t) \in A(x_t)$, para toda $h_t \in \mathbb{H}_t$ y $t = 0, 1, \ldots$.
  El conjunto de políticas deterministas markovianas se denota por Π_{DM}.

- **Determinista estacionaria**, si existe una función $f \in \mathbb{F}$ tal que, $\pi_t(\cdot \mid h_t)$ está concentrada en $f(x_t) \in A(x_t)$ para toda $h_t \in \mathbb{H}_t$ y $t = 0, 1, \ldots$.
  Al conjunto de políticas deterministas estacionarias se le denota por Π_{DS}.

1.3. Propiedad de Markov

Sea $\{R_t\}$ una sucesión de kérneles estocásticos en $\mathcal{P}(X|X)$, y sea $\{y_t\}$ un proceso estocástico que toma valores en X. Entonces $\{y_t\}$ se dice Proceso de Markov no homogéneo con kérneles de transición $\{R_t\}$ si, para cada $B \in \mathbb{B}(X)$ y $t = 0, 1 \ldots$

$$P(y_{t+1} \in B| \, y_0, \ldots, y_t) = P(y_{t+1} \in B| \, y_t) = R_t(B|y_t) \tag{1.1}$$

La primera igualdad es llamada propiedad de Markov.

Proposición 1.4. Sea ν una distribución inicial arbitraria. Si $\pi = \{\pi_t\} \in \Pi_{RM}$ entonces, $\{x_t\}$ es un Proceso de Markov no homogéneo, con kérneles de transición $\{Q(\cdot|\cdot, \varphi_t)\}$, es decir, sucede (1.1) para cada $B \in \mathbb{B}(X)$ y $t = 0, 1, \ldots$

$$\begin{aligned} P_\nu^\pi(x_{t+1} \in B)|\, x_0, \ldots, x_t) &= P_\nu^\pi(x_{t+1} \in B|\, x_t) &(1.2)\\ &= Q(B|\, x_t, \varphi_t). &(1.3)\end{aligned}$$

Particularmente, si $\pi = \{f_t\} \in \Pi_{DM}$, (1.3) se tiene para los kérneles de transición $Q(\cdot|\ \cdot, f_t)$. Más aún, para las políticas estacionarias $\varphi^\infty \in \Pi_{RS}$ y $f^\infty \in \Pi_{DS}$, $\{x_t\}$ es un Proceso de Markov homogéneo con kérneles de transición $Q(\cdot|\cdot, \varphi)$ y $Q(\cdot|\cdot, f)$, respectivamente.
Demostración.
Note que, para una política arbitraria $\pi = \{\pi_t\}$

$$P_\nu^\pi(x_{t+1} \in B|\, h_t) = \int_A Q(B|\, x_t, a_t)\, \pi(da_t|\, h_t) \tag{1.4}$$

para toda $B \in \mathbb{B}(X)$ y $t = 0, 1, \ldots$. Ya que por la Proposición C.3

$$\begin{aligned} P_\nu^\pi(x_{t+1} \in B|\, h_t) &= E_\nu^\pi[P_\nu^\pi(x_{t+1} \in B|\, h_t, a_t)|\, h_t]\\ &= E_\nu^\pi[Q(B|\, x_t, a_t)|\, h_t]\\ &= \int_A Q(B|\, x_t, a_t)\, \pi_t(da_t|\, h_t)\end{aligned}$$

y esto pruebla (1.4).

En particular, si $\pi = \varphi_t$ es una política markoviana aleatorizada, entonces (1.4) se convierte en

$$P_\nu^\pi(x_{t+1} \in B|\, h_t) = \int_A Q(B|\, x_t, a_t)\, \varphi(da_t|\, h_t) = Q(B|\, x_t, a_t). \tag{1.5}$$

Considere $x_0^t := (x_0, \ldots, x_t)$, por la Proposición C.3, el lado izquierdo de (1.3) se puede escribir como

$$\begin{aligned} P_\nu^\pi(x_{t+1} \in B|\, h_t) &= E_\nu^\pi[P_\nu^\pi(x_{t+1} \in B|\, h_t)|\, x_0^t]\\ &= E_\nu^\pi[Q(B|\, x_t, \varphi_t)|\, x_0^t]\\ &= Q(B|\, x_t, \varphi_t).\end{aligned}$$

La última igualdad se da por la Proposición C.5. Análogamente, el lado derecho de la igualdad (1.2) satisface

$$\begin{aligned} P_\nu^\pi(x_{t+1} \in B|\, x_t) &= E_\nu^\pi[P_\nu^\pi(x_{t+1} \in B|\, h_t)|\, x_t] \\ &= E_\nu^\pi[Q(B|x_t, \varphi_t)|\, x_t] \\ &= Q(B|\, x_t, \varphi_t). \end{aligned}$$

□

1.4. Procesos de Decisión de Markov

1.4.1. Construcción

Sea $(\Omega, \mathscr{F})$ un espacio medible conformado por $\Omega := \overline{\mathbb{H}}_\infty = (X \times A)^\infty$ y $\mathscr{F}$ su correspondiente σ-álgebra producto. Observe que los elementos de Ω son de la forma $\omega = (x_0, a_0, x_1, a_1, \ldots)$, con $x_t \in X$ y $a_t \in A$ para todo $t \in \mathbb{N}_0$; las proyecciones x_t, a_t de Ω, son llamadas variables de estado y acción, respectivamente.

Note que $\mathbb{H}_\infty = \mathbb{K}^\infty \subset \Omega$, donde $\mathbb{H}_\infty$ es llamado espacio de las historias admisibles, y para cada $(x_0, a_0, x_1, a_1, \ldots) \in \mathbb{H}_\infty$ se tiene que $(x_t, a_t) \in \mathbb{K}$, $\forall\ t \in \mathbb{N}_0$.

Sean $\pi = \{\pi_t\}$ y ν una política de control y una medida de probabilidad sobre X, respectivamente. Por el teorema de Ionescu-Tulcea existe una única medida de probabilidad P_ν^π en $(\Omega, \mathscr{F})$ tal que, para cada $B \in \mathbb{B}(X)$, $C \in \mathbb{B}(A)$ y $h_t \in \mathbb{H}_t$

$$\begin{aligned} P_\nu^\pi(x_0 \in B) &= \nu(B) \\ P_\nu^\pi(a_t \in C \mid h_t) &= \pi_t(C \mid h_t) \\ P_\nu^\pi(x_{t+1} \in B \mid h_t, a_t) &= Q(B \mid x_t, a_t). \end{aligned}$$

El proceso estocástico $(\Omega, \mathscr{F}, P_\nu^\pi, \{x_t\})$ es llamado **Proceso de Decisión de Markov a tiempo discreto**.

1.5. Criterio de Rendimiento

Los procesos de decisión de Markov cuentan con un criterio de rendimiento o función objetivo, el cual nos sirve para medir la calidad de las políticas utilizadas. En esta tesis se utilizará el criterio de costo total descontado, definido a continuación.

Definición 1.5. El criterio de costo total descontado, está dado por

$$V_\alpha(\pi, x) := E_x^\pi\Big[\sum_{t=0}^{\infty} \alpha^t\, C(x_t, a_t)\Big], \tag{1.6}$$

donde $\alpha \in (0,1)$, es llamado factor de descuento.

Definición 1.6. La función de valor óptimo, o simplemente función de valor, se define como

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

El problema de control consiste en encontrar la política $\pi^* \in \Pi$ que satisfaga

$$V^*(x) = V_\alpha(\pi^*, x), \quad \forall x \in X. \tag{1.7}$$

La política $\pi^* \in \Pi$ que satisface (1.7) es llamada **política óptima**.

1.5.1. Horizonte del Problema

Teniendo el criterio de rendimiento de nuestro modelo, podemos definir el horizonte del problema, es decir, hasta qué momento éste será observado.

Si el horizonte de nuestro proceso es finito, entonces nuestro criterio de rendimiento se define como

$$V_\alpha(\pi, x) := E_x^\pi\Big[\sum_{t=0}^{N} \alpha^t\, C(x_t, a_t)\Big], \quad \pi \in \Pi, \quad x \in X, \tag{1.8}$$

donde $\alpha \in (0,1)$ es el criterio de descuento y $N \in \mathbb{N}$ el horizonte del problema.

Cuando no se tiene un tiempo determinado a futuro en el cual terminará nuestro proceso, es conveniente considerar que el horizonte es infinito. En este caso, el criterio de rendimiento se define como

$$V_\alpha(\pi, x) := E_x^\pi \Big[\sum_{t=0}^{\infty} \alpha^t \, C(x_t, a_t) \Big], \quad \pi \in \Pi, \quad x \in X, \tag{1.9}$$

con factor de descuento $\alpha \in (0, 1)$.

1.6. Programación Dinámica

1.6.1. Horizonte Finito

A continuación se enunciará el Teorema de Programación Dinámica para horizonte finito. Para este caso, el criterio de costo total descontado se define como en (1.8).

Teorema 1.7. Sean $V_1, \ldots, V_N$ funciones definidas en X como

$$V_N(x) := C_N(x) \tag{1.10}$$

donde $C_N : X \to \mathbb{R}^+$ denota el costo terminal y para cada $t = N-1, \ldots, 0$,

$$V_t(x) := \min_{a \in A(x)} \{ C(x, a) + \alpha \int_X V_{t+1}(y) \, Q(dy|\, x, a) \}. \tag{1.11}$$

Suponga que estas funciones son medibles y que para cada $t \in \{0, \ldots, N-1\}$ existe un selector $f_t \in \mathbb{F}$ tal que $f_t(x) \in A(x)$ alcanza el mínimo en (1.11), es decir, $\forall x \in X$ y $t \in \{0, \ldots, N-1\}$

$$V_t(x) = C(x, f_t(x)) + \alpha \int_X V_{t+1}(y) \, Q(dy|\, x, f_t(x)).$$

Entonces la política $\pi^* = \{f_0, \ldots, f_{N-1}\}$ es óptima y se tiene que

$$V^*(x) = V_\alpha(\pi^*, x) = V_0(x), \quad \forall x \in X. \tag{1.12}$$

Demostración.
Sea $\pi = \{\pi_t\}$ una política arbitraria y sean

$$\mathscr{C}_t(\pi, x) := E^{\pi}\Big[\sum_{n=t}^{N-1} \alpha^n\ C(x_n, a_n) + \alpha^N C_N(x_N)\Big| x_t = x\Big],$$

$$\mathscr{C}_N(\pi, x) := E^{\pi}[\alpha^N C_N(x_N)|x_N = x] = \alpha^N C_N(x_N),$$

donde $\mathscr{C}_t(\pi, x)$ es el costo total esperado del tiempo t al tiempo terminal N, dado el estado $x_t = x$, al tiempo t, con $t = 0, \ldots, N-1$.

Observe que $V_\alpha(\pi, x) = \mathscr{C}_0(\pi, x)$.
Para probar el teorema, debemos mostrar que para todo $x \in X$ y $t \in \{0, \ldots, N-1\}$

$$\mathscr{C}_t(\pi, x) \geqslant V_t(x). \tag{1.13}$$

Y que la igualdad en (1.13) se da cuando $\pi = \pi^*$, es decir,

$$\mathscr{C}_t(\pi^*, x) = V_t(x). \tag{1.14}$$

Particularmente, para $t = 0$, $V_\alpha(\pi, x) \geqslant V_0(x)$, con $V_\alpha(\pi^*, x) = V_0(x)$, $\forall x \in X$. De esto que tengamos la igualdad (1.12), así

$$\mathscr{C}_0(\pi^*, x) = V_0(x) = V_\alpha(\pi^*, x) = V^*(x),$$

notemos que para $t = N$, se cumplen trivialmente (1.13) y (1.14), pues

$$\mathscr{C}(\pi, x) = V_N(x) = C_N(x).$$

Ahora, procederemos a probar (1.13) y (1.14), por inducción hacia atrás, para $t = N-1, \ldots, 0$.

Supongamos que para alguna $t \in \{0, \ldots, N-1\}$,

$$\mathscr{C}_{t+1}(\pi, x) \geqslant V_{t+1}(x), \quad x \in X,$$

entonces

$$
\begin{aligned}
\mathscr{C}_t(\pi, x) &= E^{\pi}\Big[C(x_t, a_t) + \sum_{n=t+1}^{N-1} \alpha\, C(x_n, a_n) \;+\; C_N(x_N)\Big| x_t = x\Big] \\
&= E^{\pi}[\alpha^t C(x_t, a_t)|x_t = x] \\
&\quad + E^{\pi}\Big[\sum_{n=t+1}^{N-1} \alpha^n\, C(x_n, a_n) \;+\; \alpha^N C_N(x_n)\Big| x_t = x\Big] \\
&= \int_A \alpha^t C(x, a)\, \pi(da|\, x) \\
&\quad + E^{\pi}\Big[\sum_{n=t+1}^{N-1} \alpha^n\, C(x_n, a_n) \;+\; \alpha^N C_N(x_n)\Big| x_t = x\Big].
\end{aligned}
$$

Note que

$$
\begin{aligned}
&E^{\pi}\Big[\sum_{n=t+1}^{N-1} \alpha^n\, C(x_n, a_n) \;+\; \alpha^N C_N(x_N)\Big| x_t = x\Big] = \\
&E^{\pi}\Big[E^{\pi}[\sum_{n=t+1}^{N-1} \alpha^n\, C(x_n, a_n) \;+\; \alpha^N C_N(x_n)|\, x_{t+1} = y]\Big| x_t = x\Big].
\end{aligned}
$$

Ya que x_{t+1} está en función de x_t y por la Propiedad C.3 (*ver Apéndice C*).

Por otro lado,

$$
E^{\pi}\Big[\sum_{n=t+1}^{N-1} \alpha^n\, C(x_n, a_n)\, \pi(da|x) \;+\; \alpha^N C_N(x_n)\Big| x_{t+1} = y\Big] = \mathscr{C}_{t+1}(\pi, y).
$$

Así,

$$
\begin{aligned}
\mathscr{C}_t(\pi, x) &= \int_A \alpha^t\, C(x,a)\, \pi(da|\, x) \;+\; E^{\pi}[\mathscr{C}_{t+1}(\pi, y)|\, x_t = x] \\
&= \int_A \Big\{\alpha^t C(x,a) \;+\; \int_X \mathscr{C}_{t+1}(\pi, y)\, Q(dy|\, x, a)\Big\}\, \pi_t(da|\, x) \\
&\geqslant \int_A \Big\{C(x,a) \;+\; \int_X \alpha\, V_{t+1}(\pi, y)\, Q(dy|\, x, a)\Big\}\, \pi_t(da|\, x) \\
&\geqslant \min_{A(x)} \Big[C(x,a) + \int_X \alpha\, V_{t+1}(y)\, Q(dy|\, x, a)\Big] \\
&= V_t(x).
\end{aligned}
$$

De esto se cumple que:

$$\mathscr{C}_t(\pi, x) \geqslant V_t(x),$$

para toda $t \in \{N-1, \ldots, 0\}$.

Por otra parte, si la igualdad se cumple para $\pi = \pi^*$, es decir,

$$\mathscr{C}_{t+1}(\pi^*, x) = V_{t+1}(x),$$

se tiene que

$$
\begin{aligned}
\mathscr{C}_t(\pi^*, x) &= C(x, f_t(x)) \;+\; E^{\pi}[\mathscr{C}_{t+1}(\pi^*, y)|\, x_t = x] \\
&= C(x, f_t(x)) \;+\; \alpha \int_X V_{t+1}(y)\, Q(dy|x, f_t(x)) \\
&= V_t(x).
\end{aligned}
$$

De esto que se cumpla (1.14) y esto sucede para toda $t \in \{N-1, \ldots, 0\}$. Así, se concluye lo deseado.

□

1.6.2. Horizonte Infinito

En este caso, el criterio a minimizar es

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

donde $\alpha \in (0,1)$ es el factor de descuento. Una política π^* que satisface

$$V_\alpha(\pi^*, x) = \inf_{\Pi} V_\alpha(\pi, x) =: V^*(x), \quad \forall x \in X, \tag{1.15}$$

se dice óptima α-descontada, y V^* es la función de valor α-descontado.

Suponga que la función de costo en un paso C es no negativa. Denotamos por V_n el costo en el n-ésimo estado, dado por

$$V_n(\pi, x) := E_x^\pi \Big[\sum_{t=0}^{n-1} \alpha^t \, C(x_t, a_t) \Big]. \tag{1.16}$$

Así, por el Teorema de Convergencia Monótona (*ver* [3]), podemos escribir $V_\alpha(\pi, x)$ como

$$V_\alpha(\pi, x) = \lim_{n \to \infty} V_n(\pi, x).$$

Una función medible $\nu : X \to \mathbb{R}$ se dice solución a la ecuación de optimalidad de costo α-descontado, si satisface

$$\nu(x) = \min_{A(x)} [C(x,a) + \alpha \int_X \nu(y) \, Q(dy|\, x, a)], \quad \forall x \in X. \tag{1.17}$$

Probaremos que la función de valor α-descontado V^*, definida en (1.15), es solución de la ecuación de optimalidad de costo α-descontado, es decir,

$$V^*(x) = \min_{A(x)} [C(x,a) + \alpha \int_X V^*(y) \, Q(dy|\, x, a)], \quad \forall x \in X. \tag{1.18}$$

Considere las funciones de iteración de valores definidas como

$$\nu_n(x) = \min_{A(x)} [C(x,a) + \alpha \int_X \nu_{n-1}(y) \, Q(dy|\, x, a)], \tag{1.19}$$

para toda $x \in X$ y $n = 1, 2, \ldots$, con $\nu_0 \equiv 0$, y se probará que

$$V^*(x) = \lim_{n \to \infty} \nu_n(x), \quad \forall x \in X.$$

Se espera tener el resultado ya que ν_n es la función de valor de costo descontado en el n-ésimo estado V_n dado en (1.16), con costo terminal cero, así

$$\nu_n(x) = \inf_{\Pi} V_n(\pi, x), \quad \forall x \in X, \tag{1.20}$$

haciendo $n \to \infty$ en (1.19) obtenemos (1.18) si logramos justificar el intercambio entre límite y mínimo.
Este método es llamado método de las aproximaciones sucesivas.

Hipótesis 1.8. Considere

- La función de costo en un paso C es semicontinua inferior (s.c.i.), no negativa e inf-compacta en $\mathbb{K}$.
- Q es fuertemente continuo.

Hipótesis 1.9. Existe una política π tal que $V_\alpha(\pi, x) < \infty$ para cada $x \in X$.

Denotamos por Π^0 a la familia de políticas que cumplen la Hipótesis 1.9.

Teorema 1.10. Suponga que se cumplen las Hipótesis 1.8 y 1.9. Entonces

a) La función de valor α-descontado es la solución mínima de la ecuación de optimalidad de costo α-descontado, es decir,

$$V^*(x) = \min_{A(x)}[C(x,a) + \alpha \int_X V^*(y)\, Q(dy|\, x, a)] \tag{1.21}$$

y, si u es otra solución de la ecuación de optimalidad, entonces $u(x) \geqslant V^*(x)$ para todo $x \in X$,

b) Existe un selector $f_* \in \mathbb{F}$ tal que $f_* \in A(x)$ alcanza el mínimo de (1.21), es decir,

$$V^*(x) = C(x, f_*) + \alpha \int_X V^*(y)\, Q(dy|\, x, f_*) \tag{1.22}$$

para toda $x \in X$, y la política determinista estacionaria f_*^∞ es óptima; así mismo si $f_*^\infty \in \Pi_{DS}$ es óptima, entonces satisface (1.22),

c) Si π^* es una política tal que $V_\alpha(\pi^*, \cdot)$ es solución de la ecuación de optimalidad y satisface

$$\lim_{n\to\infty} \alpha^n E_x^\pi V_\alpha(\pi^*, x) = 0, \quad \forall \pi \in \Pi^* \text{ y } x \in X,$$

entonces, $V_\alpha(\pi^*, \cdot) = V^*(\cdot)$, así π^* es óptima si y sólo si $V_\alpha(\pi^*, \cdot)$ satisface la ecuación de optimalidad de costo α-descotado,

d) Si existe una política óptima α-descontada entonces, existe una que es determinista estacionaria.

Para la demostración de este Teorema, serán requeridos algunos lemas.

Lema 1.11. Sean u y u_n, $n = 1, 2, \ldots$ multifunciones s.c.i., acotadas inferiormente e inf-compactas en $\mathbb{K}$. Si $u_n \uparrow u$ entonces,

$$\lim_{n\to\infty} \min_{A(x)} u_n(x,a) = \min_{A(x)} u(x,a), \quad \forall x \in X.$$

Demostración.
Para cada $x \in X$, definimos

$$\ell(x) := \lim_{n} \min_{A(x)} u_n(x,a) \quad y \quad u^*(x) := \min_{A(x)} u(x,a).$$

Como $u_n \uparrow u$, tenemos que $\ell(\cdot) \leqslant u^*(\cdot)$. Para probar la otra desigualdad, fijemos $x \in X$ y considere los conjuntos

$$\begin{aligned} A_n &:= \{a \in A(x) |\ u_n(x) \leqslant u^*(x)\}, \quad n \in \mathbb{N}, \\ A_0 &:= \{a \in A(x) |\ u(x,a) = u^*(x)\}. \end{aligned}$$

Note que cada uno de estos conjuntos es compacto y no vacío. Observe que los A_n decrecen a un conjunto compacto, pues si $a \in A_{n+1}$ entonces, $u_{n+1}(x,a) \leqslant u^*(x)$ pero $u_n(x,a) \leqslant u_{n+1}(x,a)$, de esto que $a \in A_n$, así $A_n \downarrow A_0$.

Para cada $n \geqslant 1$ sea $a_n \in A_n$ tal que $u_n(x,a_n) = \min_{A(x)} u_n(x,a)$, podemos asegurar la existencia de a_n por la Proposición C.6 (*ver Apéndice C*). Dado que $A_n \downarrow A_0$ y son compactos, existen $a_0 \in A_0$ y una subsucesión $\{a_{n_i}\} \subset \{a_n\}$ tal que $a_{n_i} \to a_0$.

Ahora, se tiene que

$$u_{n_i}(x, a_{n_i}) \geqslant u_n(x, a_{n_i}), \quad \forall n_i > n,$$

haciendo $n \to \infty$

$$\begin{aligned} \lim_{n\to\infty} u_{n_i}(x,a_{n_i}) &\geqslant u_n(x,a_{n_i}) \\ \lim_{n\to\infty} \min_{A(x)} u_n(x,a_{n_i}) &\geqslant u_n(x,a_0) \\ \ell(x) &\geqslant u_n(x,a_0), \end{aligned}$$

y como $u_n \uparrow u$, se tiene que $\ell(x) \geqslant u(x, a_0) = u^*(x)$ y esto sucede para cualquier $x \in X$, así, queda probado el Lema 1.11.

$\square$

Definición 1.12. $M(X)^+$ denota el conjunto de las funciones medibles no negativas en X, y para cada $u \in M(X)^+$, Tu es la función en X definida como

$$Tu(x) := \min_{A(x)} \{C(x,a) + \alpha \int_X u(y)\, Q(dy|\, x, a)\}, \quad x \in X.$$

Lema 1.13. Suponga que se cumple la Hipótesis 1.8, T es un operador sobre $M(X)^+$ en sí mismo, es decir, para cada $u \in M(X)^+$, $Tu \in M(X)^+$ y más aún, existe un selector $f \in \mathbb{F}$ tal que

$$Tu(x) = C(x,f) + \alpha \int_X u(y)\, Q(dy|\, x, f), \quad \forall x \in X.$$

Observe que usando el operador T podemos reescribir la ecuación de optimalidad de costo α-descontado y las funciones de iteración de valores como

$$V^* = TV^* \quad \text{y} \quad \nu_n = T\nu_{n-1}, \quad \text{para } n \geqslant 1,$$

y $\nu_0 := 0$.

Se probará que la función de valor V^* es un punto fijo del operador T.

Lema 1.14. Suponga que se cumplen las Hipótesis 1.8 y 1.9

a) Si $u \in M(X)^+$ y es tal que $u \geqslant Tu$ entonces, $u \geqslant V^*$.

b) Si $u : X \to \mathbb{R}$ es una función medible tal que Tu está bien definida, con $u \leqslant Tu$, y
$$\lim_{n\to\infty} \alpha^n E[u(x_n)] = 0, \quad \forall \pi \in \Pi^0, \;\; x \in X,$$
entonces $u \leqslant V^*$.

Demostración.
a) Sea $u \in M(X)^+$ con $u \geqslant Tu$ y considere $f \in \mathbb{F}$ tal que

$$u(x) \geqslant C(x,f) + \alpha \int_X u(y)\, Q(dy|\, x, f), \quad \forall x \in X,$$

iterando esta desigualdad, tenemos que

$$\begin{aligned}
u(x) &\geqslant C(x,f)+\alpha\int_X [C(y,f)+\alpha\int_X u(z)\,Q(dz|\,y,f)]\,Q(dy|\,x,f)\\
&= C(x,f)+\alpha\int_X C(y,f)\,Q(dy|\,x,A)\\
&\quad +\alpha^2\int_X\int_X u(z)\,Q(dz|\,y,f)\,Q(dy|x,f)\\
&= E_x^f\Big[\sum_{t=0}^{1}\alpha^t\,C(x_t,f)\Big]+\alpha^2 E_x^f[u(x_2)],
\end{aligned}$$

siguiendo la iteración, se obtiene

$$u(x)\geqslant E_x^f\Big[\sum_{t=0}^{n-1}\alpha^t\,C(x_t,f)\Big]+\alpha^n E_x^f[u(x_n)],$$

donde

$$E_x^f[u(x_n)]=\int_X u(y)\,Q^n(dy|x,f)$$

y Q^n denota el kérnel de transición en el n-ésimo paso del Proceso de Markov usando f. Como u es no negativa,

$$u(x)\geqslant E_x^f\Big[\sum_{t=0}^{n-1}\alpha^t\,C(x_t,f)\Big]$$

para cualquier $n\in\mathbb{N}$ y $x\in X$, haciendo $n\to\infty$

$$u(x)\geqslant V_\alpha(f,x)\geqslant V^*(x)\quad\forall x\in X.$$

Así, tenemos el inciso *a)*.

b) Considere $\pi\in\Pi$, $x\in X$ y que $Tu\geqslant u$. Entonces, por la Propiedad de

Markov, tenemos

$$\begin{aligned} E_x^\pi[\alpha^{t+1}\, u(x_{t+1})|\, h_t, a_t] &= \alpha^{t+1}\int_X u(y)\, Q(dy|\, x_t, a_t) \\ &= \alpha^t\, [C(x_t,a_t) + \alpha\int_X u(y)\, Q(dy|\, x_t, a_t) \\ &\quad -C(c_t, a_t)] \\ &= \alpha^t[C(x_t,a_t) + \alpha\int u(y)\, Q(dy|x_t,a_t)] \\ &\quad -\alpha^t C(x_t,a_t) \\ &\geqslant \alpha^t[u(x_t) - C(x_t,a_t)]. \end{aligned}$$

Así,

$$\begin{aligned} \alpha^t C(x_t,a_t) &\geqslant -E_x^\pi[\alpha^{t+1}\, u(x_{t+1}) - \alpha^t\, u(x_t)|\, h_t, a_t] \\ &= E_x^\pi[\alpha^t\, u(x_t) - \alpha^{t+1}\, u(x+1)|\, h_t, a_t]. \end{aligned}$$

Tomando esperanzas y sumando desde $t = 0$ a $n-1$, tenemos

$$\begin{aligned} E_x^\pi\Big[\sum_{t=0}^{n-1}\alpha^t\, C(x_t,a_t)\Big] &\geqslant E_x^\pi\Big[\sum_{t=0}^{n-1}\alpha^t\, u(x_t) - \alpha^{t+1}u(x_{t+1})\Big|\, h_t, a_t\Big] \\ &= E_x^\pi[u(x_0) - \alpha^n\, u(x_n)] \\ &= u(x) - \alpha^n E_x^\pi[u(x_n)], \end{aligned}$$

haciendo $n \to \infty$ y considerando las Hipótesis de este inciso, tenemos que

$$E_x^\pi\Big[\sum_{t=0}^{\infty}\alpha^t\, C(x_t,a_t)\Big] \geqslant u(x),$$

así,

$$V_\alpha(\pi, x) \geqslant u(x),$$

y esto fue para x y π arbitrarios, de esto que

$$V^* \geqslant u.$$

□

Lema 1.15. Suponga que se cumplen las Hipótesis 1.8 y 1.9, entonces $\nu_n \uparrow V^*$ y V^* satisface la ecuación de optimalidad de costo α-descontado.
DEMOSTRACIÓN.
Observe que para cada $x \in X$, $\pi \in \Pi$ y $n \in \mathbb{N}_0$

$$\nu_n(x) \leqslant V_n(x) \leqslant V_\alpha(\pi, x).$$

Por lo tanto,

$$\nu_n(x) \leqslant V^*(x). \tag{1.23}$$

El operador T es monótono, es decir, si u y $u' \in M(X)^+$ y $u \geqslant u'$ entonces, $Tu \geqslant Tu'$. Como $\nu_0 := 0$ y $\nu_n := T\nu_{n-1}$ para $n \geqslant 1$, las funciones de iteración de valores forman una sucesión creciente en $M(X)^+$, lo cual implica que $\nu_n \uparrow \nu^*$, para alguna $\nu^* \in M(X)^+$.

Considere

$$u_n(x,a) := C(x,a) + \alpha \int_X \nu_n(y)\, Q(dy|\, x,a),$$
$$u(x,a) := C(x,a) + \alpha \int_X \nu^*(y)\, Q(dy|\, x,a),$$

$(x,a) \in \mathbb{K}$. Por el Teorema de la Convergencia Monótona (*ver* [3]), tenemos que $u_n \uparrow u$.
Por otro lado, las funciones u y u_n son s.c.i. e inf-compactas en $\mathbb{K}$ (*ver* [9]), así, por el Lema 1.11

$$\nu^* = \lim_{n\to\infty} \nu_n = \lim_{n\to\infty} T\nu_{n-1} = T\nu^*,$$

es decir, ν^* es una solución de la ecuación de optimalidad de costo α descontado. □

Falta ver que $\nu^* = V^*$. Por el Lema 1.14, $\nu^* \geqslant V^*$ y la otra desigualdad se sigue de (1.23) y el hecho de que $\nu_n \uparrow \nu^*$.

A continuación, demostraremos el Teorema 1.10.
DEMOSTRACIÓN.
a) Del Lema 1.15 V^* es una solución de la ecuación de optimalidad de costo

descontado y por el Lema 1.14, V^* es la mínima solución de la Ecuación de Optimalidad de Costo Descontado, es decir, si $u \geqslant Tu$, entonces $u \geqslant V^*$.

b) La existencia de un selector $f_* \in \mathbb{F}$ que satisfaga que para cada $x \in X$

$$V^*(x) = C(x, f_*) + \alpha \int_X V^*(y)\, Q(dy|\, x, f_*)$$

se asegura por el Lema 1.13.

Iterando la relación anterior, se tiene que, para cada $x \in X$ y $n \geqslant 1$

$$\begin{aligned} V^*(x) &= E_x^{f_*}\Big[\sum_{t=0}^{n-1} \alpha^t\, C(x_t, f_*)\Big] + \alpha^n E_x^{f_*}[V^*(x_n)] \\ &\leqslant E_x^{f_*}\Big[\sum_{t=0}^{n-1} \alpha^t\, C(x_t, f_*)\Big]. \end{aligned}$$

Así, cuando $n \to \infty$ de (1.15), $V^* \leqslant V_\alpha(f_*^\infty, x)$, así $V^* = V_\alpha(f_*^\infty, x)$, de ésto que f_*^∞ sea óptima.
De manera recíproca, si $f_* \in \Pi_{DS}$, entonces

$$\begin{aligned} V_\alpha(f_*, x) &= E_x^{f_*}\Big[\sum_{t=0}^{\infty} \alpha^t C(x_t, f_*)\Big] \\ &= E_x^{f_*}\Big[C(x_0, f_*) + \sum_{t=1}^{\infty} \alpha^t\, C(x_t, f_*)\Big] \\ &= C(x, f_*) + E_x^{f_*}\Big[\sum_{t=1}^{\infty} \alpha^t C(x_t, f_*)\Big]. \end{aligned}$$

Usando la Proposición C.4 y la Propiedad de Markov, podemos reescribir

$$\begin{aligned} E\Big[\sum_{t=1}^{\infty} \alpha^t\, C(x_t, f_*)\Big] &= \int_X E_x^{f_*}\Big[\sum_{t=1}^{\infty} \alpha^t\, C(x_t, f_*)\Big| x_1 = y\Big]\, Q(dy|x, f_*) \\ &= \int_X V_\alpha(f_*, y)\, Q(dy|\, x, f_*). \end{aligned}$$

Particularmente, si f_* es óptima entonces, $\nu^*(x) = V_\alpha(f_*, x), \quad x \in X$.

c) Si $V_\alpha(\pi^*, x)$ satisface la Ecuación de Optimalidad entonces, por el Lema 1.14, tenemos que $V_\alpha(\pi^*, \cdot) = V^*(\cdot)$.

d) Finalmente, podemos concluir que si una política óptima existe, entonces existe una poítica óptima determinista estacionaria (*ver* [10]).

$\square$

Capítulo 2

POLÍTICAS UMBRAL EN PROCESOS DE DECISIÓN DE MARKOV

Las políticas umbral son estrategias en las cuales se tienen dos posibles acciones a elegir, se escoge alguna de las acciones en base a la observación de una variable que caracteriza el estado del sistema. Estas políticas se encuentran comúnmente en líneas de espera y teoría de inventarios (*ver* [12] *y* [14]).
La importancia del estudio de las políticas umbral surge de que estas políticas ya poseen una caracterización previa, de ésto que su implementación sea más fácil.

2.1. Política Umbral

Considere X subconjunto de $\mathbb{R}^+ \cup \{0\}$, de la forma $[0, \infty)$ o $[0, K]$, $\;K \in \mathbb{R}^+$ y $A \subseteq X$.

Definición 2.1. Una política (o estrategia) umbral pura, con umbral $S \in X$, es una política determinista, $f \in \mathbb{F}$, tal que para cualquier estado del sistema

$x \in [0, S]$, asigna una acción a_1 y para cualquier otro estado asigna una acción a_2.

Particularmente, nos enfocaremos en las políticas umbral de la forma

$$f_S(x) = \begin{cases} S - x, & \text{si} \quad x \in [0, S], \\ 0, & \text{en otro caso.} \end{cases}$$

con umbral $S \geqslant 0$.

2.2. Existencia de una Política Umbral Óptima

En esta sección se planteará un Teorema que asegura la existencia de políticas umbral óptimas para tres casos que serán enunciados más adelante, a continuación se planteará el Modelo de Control de Markov sobre el cual vamos a trabajar, seguido de esto se darán las especificaciones e hipótesis de cada caso, el enunciado y la demostración del Teorema mencionado.

2.2.1. Modelo de Control de Markov

Considere un Modelo de Control de Markov cuya dinámica está dada por

$$x_{t+1} = (x_t + a_t - \xi_t)^+, \quad t \in \mathbb{N}_0 \quad \text{y} \quad x_0 = x,$$

donde $\{\xi_t\}$, son variables aleatorias independientes e idénticamente distribuidas. La dinámica anterior fue estudiada por D. V. Lindley para el caso de líneas de espera (*ver* [13]).
Definimos su función de costo en un paso como

$$C(x, a) := H(x + a) + ba, \tag{2.1}$$

donde H es una función Borel medible no negativa y b es una constante positiva, de esto que C sea una función no negativa.

Definimos la función

$$L(y) := H(y) + by + \alpha E[V^*(y - \xi)^+], \quad y \in X, \tag{2.2}$$

observe que la ecuación de optimalidad (1.17), se puede reescribir como

$$V^*(x) = \min_{a \in A(x)} L(x+a) - bx, \quad \forall x \in X.$$

2.2.2. Caso 1. Espacios de Estados y Acciones Acotados

En este caso consideramos $X = A = [0, K]$, para algún $K \geqslant 0$, y para toda $x \in X$, $A(x) = [0, K - x]$.

Hipótesis 2.2. La función H es continua en X.

Teorema 2.3. Considere el modelo y función de costos descritos en la Sección 2.2.1. Si los espacios de estados y acciones son como los descritos arriba y la función H cumple la Hipótesis 2.2. Entonces la función de valor V_α es continua y existe una política umbral óptima f_S, con umbral $S \in X$.
Demostración.
La función de costo definida como $C(x, a) = H(x + a) + ba$ es continua en $\mathbb{K}$, pues H es continua por hipótesis, además es acotada en $\mathbb{K}$, ya que $\mathbb{K}$ es un conjunto compacto.

Por la Proposición B.5 se tiene que la multifunción $x \to A(x)$ es continua (*ver Apéndice B*).

Por otro lado, de D.1 se tiene que V^*, la función de valor, es continua (*ver Apéndice D*).

Es claro que $L(y) = H(y) + by + \alpha E\big[V^*((y - \xi)^+)\big]$ es continua, ya que H lo es por hipótesis y la continuidad de $E\big[V^*((y - \xi)^+)\big]$ se tiene de la Proposición C.2 (*ver Apéndice C*).

Así, podemos definir el conjunto $B = \{y \in X : \ L(y) = \inf_{x \in X} L(x)\}$, observe que éste es no vacío, pues L es continua y X es compacto, de esto que la función alcance su ínfimo en X.

Ahora, considere $\{y_n\} \subset B$ tal que $y_n \to y$, para algún $y \in X$; sabemos que L es una función continua, así para $\epsilon > 0$, existen $\delta > 0$ y $N \in \mathbb{N}$ tales que, si

$n \geqslant N$, entonces

$$|y - y_n| < \delta,$$

así

$$|L(y) - L(y_n)| < \epsilon,$$

haciendo $\epsilon \to 0$, tenemos

$$L(y) - L(y_n) \leqslant 0,$$

entonces

$$L(y) \leqslant L(y_n) = \inf_{x \in X} L(x,)$$

de esto que

$$L(y) = \inf_{x \in X} L(x).$$

Se concluye que $y \in B$, así, tenemos que el conjunto B es cerrado y como $B \subset X$ y éste es compacto, se tiene que B también es compacto.

Por otro lado, por la continuidad de L, podemos definir

$$S_\alpha := \max\{y \in X : \ L(y) = \min_{z \in X} L(z)\}.$$

Dada la definición de S_α, se considerarán los siguientes casos:
(i) $S_\alpha = K$, entonces la política f_{S_α} es óptima pues, para cualquier $(x, a) \in \mathbb{K}$, se tiene

$$L(S_\alpha) \leqslant L(x + a).$$

(ii) $S_\alpha < K$, por definición de S_α, existe $S \in (S_\alpha, K]$ tal que L es estrictamente creciente en (S_α, S) pues, si L fuera decreciente en algún $S > S_\alpha$, entonces S_α no sería el máximo del conjunto B.
Ahora, definimos

$$\hat{S} := \sup\{S \in (S_\alpha, K] : \ L \ \text{ es creciente en } \ (S_\alpha, S)\}.$$

Si $\hat{S} = K$, entonces, L es creciente en $(S_\alpha, K]$, y se tiene el resultado deseado.

Si $\hat{S} < K$, suponga que L es no creciente en $(S_\alpha, K]$. Es decir, la función L alcanza máximos y mínimos en puntos interiores de cualquier subintervalo de

$(S_\alpha, K]$.

Ahora, definimos

$$S^* := \inf\{y \in [\hat{S}, K] : \; L(y) = \inf_{y \in [\hat{S}, K]} L(y)\}.$$

Note que L es estrictamente decreciente en una vecindad derecha de $\hat{S}$, de lo contrario $\hat{S}$ no sería el supremo del conjunto sobre el cual se define.
Observe que, el conjunto $\{y \in (S_\alpha, \hat{S}) : \; L(y) = \inf_{x \in [\hat{S}, K]} L(x)\}$ es no vacío, ya que

$$L(S_\alpha) < L(S^*) < L(\hat{S}),$$

la primera desigualdad se tiene por definición de S_α, pues éste es el máximo elemento en X que minimiza la función L, así $L(S_\alpha) < L(S) \;\; \forall S \in (S_\alpha, K]$. Particularmente $L(S_\alpha) < L(S^*)$, ya que $S_\alpha < S^*$. Para la segunda desigualdad se tiene que $S^* > \hat{S}$ y S^* es el menor elemento de $[\hat{S}, K]$ que hace a L alcanzar su mínimo en $[\hat{S}, K]$, entonces $L(S^*) < L(\hat{S})$.

Dado que L es continua, por el teorema del valor intermedio, al menos existe un elemento $s \in (S_\alpha, \hat{S})$ tal que $L(s) = L(S^*)$, es decir, $L(s) = \inf_{x \in [\hat{S}, K]} L(x)$, entonces el conjunto $\{y \in (S_\alpha, \hat{S}) : \; L(y) = \inf_{x \in [\hat{S}, K]} L(x)\}$ es no vacío.
Además el conjunto es acotado y cerrado, pues son las preimágenes de un singular, así podemos definir

$$S_* := \sup\{y \in (S_\alpha, \hat{S}) : \; L(y) = \inf_{z \in [\hat{S}, K]} L(z)\}.$$

Sean $\{z_n\}$ y $\{z'_n\}$ sucesiones tales que $z_n \uparrow S_*$ y $z'_n \downarrow S_*$. Se tiene que

$$\begin{aligned} V^*(z_n) &= \min_{a \in A(z_n)} \big\{L(z_n + a) - b z_n\big\} \\ &= L(z_n) - b z_n \\ &= H(z_n) + \alpha E\Big[V^*((z_n - \xi)^+)\Big], \end{aligned}$$

pues L es creciente en $(S_\alpha, \hat{S})$, entonces la acción que minimiza es $f(z_n) = 0$. Aplicando límite tenemos,

$$\lim_{n \to \infty} V^*(z_n) = H(S_*) + \alpha E\Big[V^*((S_* - \xi)^+)\Big]$$

Por otro lado, por las definiciones de S_* y S^*, se tiene que en (S_*, S^*) no hay minimizadores de L. Ahora

$$\begin{aligned} V^*(z'_n) &= \min_{a\in A(z'_n)} \left\{L(z'_n + a) - bz'_n\right\}, \quad \text{haciendo } u = z'_n + a \\ &= \min_{u\in[0,K]} \{L(u) - bz'_n\}, \end{aligned}$$

en la segunda igualdad u toma valores en $[0, K]$ ya que $0 \leqslant a \leqslant K - z'_n$, es decir, $0 \leqslant a + z'_n \leqslant K$, donde $a + z'_n = u$.

Note que $u \in (S_*, K]$ y el minimizador en este intervalo es S^*, de ésto que $u = S^*$, así $a = S^* - z'_n$, entonces

$$\begin{aligned} V^*(z'_n) &= L(z'_n + S^* - z'_n) - bz'_n \\ &= L(S^*) - bz'_n \\ &= H(S^*) + bS^* + \alpha E\Big[V^*((S^* - \xi)^+)\Big] - bz'_n. \end{aligned}$$

Aplicando límite, tenemos

$$\lim_{n\to\infty} V^*(z'_n) = H(S^*) + b(S^* - S_*) + \alpha E\Big[V^*((S^* - \xi)^+)\Big].$$

Por definición $S_* < S^*$, así

$$\lim_{n\to\infty} V^*(z_n) \neq \lim_{n\to\infty} V^*(z'_n),$$

lo cual contradice la continuidad de V^*, de esto que L sea creciente en $(S_\alpha, K]$. □

2.2.3. Caso 2. Espacio de Estados no Acotado y Espacio de Acciones Acotado

En este caso consideramos $X = [0, \infty)$ y $A = A(x) = [0, K]$, donde $K \geqslant 0$, y ésto sucede para todo $x \in X$.

Hipótesis 2.4. Se cumple lo siguiente:

1. La función H es continua en X.

2. La función L, definida como (2.2), es inf-compacta.

Teorema 2.5. Considere el modelo y la función de costos dados en la Sección 2.2.1, los espacios de estados y acciones descritos arriba y se cumplen las Hipótesis 2.4. Entonces la función de valor V^* es continua y existe una política umbral óptima f_S con umbral $S \in X$.
DEMOSTRACIÓN.
La multifunción $x \to A(x)$ es compacto valuada y continua, pues $A(x)$ es compacto y la multifunción es constante para todo $x \in X$.

La función de costo en un paso es inf-compacta en $\mathbb{K}$, ya que

$$\{a \in A \,:\, C(x,a) \leq r\} \subset \{a \in A \,:\, ba \leq r\},$$

y el subconjunto $\{a \in A \,:\, ba \leq r\}$ es compacto para toda $r \in \mathbb{R}$, pues es una bola cerrada en $\mathbb{R}$.

Por otra parte, el kérnel estocástico $Q(\cdot|\cdot,\cdot)$ es fuertemente continuo, por la Observación A.4 (*ver Apéndice A*) y la función de valor es continua por D.1 (*ver Apéndice D*).

Note que $L(y) = H(y) + by + \alpha E\Big[V^*\Big((y-\xi)^+\Big)\Big]$, $y \in X$, es continua pues, H es continua por hipótesis y por la Proposición C.2 (*ver Apéndice C*), $E\Big[V^*\Big((y-\xi)^+\Big)\Big]$ también lo es, y así se tiene el resultado.

Más aún, observe que la función L tiene ínfimo en X pues es continua y acotada inferiormente, ya que es no negativa, de ésto que exista $r = \inf_{y\in X} L(y)$, con $r \in \mathbb{R}$, así el conjunto $\{y \in X \,:\; L(y) \leqslant \inf_{x\in X} L(x)\}$ es compacto, pues L es inf-compacta. Además, el conjunto es no vacío, ya que L es continua, particularmente s.c.i., acotada inferiormente e inf-compacta, entonces existe un selector $f^* \in \mathbb{F}$ tal que

$$L(x, f^*(x)) = L^*(x) = \min_{a\in A(x)} L(x,a).$$

Así, podemos definir

$$S_\alpha := \max\{y \in X : \ L(y) = \min_{z \in X} L(z)\}.$$

Note que la función L es no decreciente en $X \cap [S_\alpha, \infty) = [S_\alpha, \infty)$, de lo contrario, se tendría que, para $x > S_\alpha$, $L(x) \leqslant L(S_\alpha)$, lo cual contradice la definición de S_α.

Ahora, definimos

$$\hat{S} := \sup\{S > S_\alpha : \ L \text{ es creciente en } (S_\alpha, S)\}.$$

Si $\hat{S} = \infty$, se tiene que L es creciente en (S_α, ∞) y obtenemos el resultado deseado.
En caso de que $\hat{S} < \infty$, suponga que L es no creciente y definimos

$$S^* := \inf\{y \in (\hat{S}, \infty) \mid L(y) = \inf_{y \in [\hat{S}, \infty)} L(y)\}.$$

Observe que $\inf_{y \in [\hat{S}, \infty)} L(y)$ existe, ya que L es continua y acotada inferiormente. Así, el conjunto $\{y \in (\hat{S}, \infty) : \ L(y) \leqslant \inf_{x \in [\hat{S}, \infty)} L(x)\}$ es compacto, y este conjunto es igual a $\{y \in (\hat{S}, \infty) : \ L(y) = \inf_{y \in [\hat{S}, \infty)}\}$.
Note que S^* es finito, ya que el conjunto sobre el que se define es compacto.

Sea

$$S_* := \sup\{y \in (S_\alpha, \hat{S}) : \ L(y) = \inf_{x \in [\hat{S}, \infty)} L(x)\}.$$

Ahora, tome $\{z_n\}$ y $\{z_n'\}$ sucesiones en X tales que $z_n \uparrow S_*$ y $z_n' \downarrow S_*$. Así

$$V^*(z_n) = \min_{a \in A(z_n)} \{L(z_n + a) - bz_n\}. \tag{2.3}$$

Como $z_n \leqslant S_* < \hat{S}$ y L es creciente en $(S_\alpha, \hat{S})$, se tiene que la acción que minimiza es $a = 0$. Sustituyendo a en (2.3), se tiene

$$\begin{aligned} V^*(z_n) &= L(z_n) - bz_n \\ &= H(z_n) + bz_n + \alpha E\Big[V^*\Big(\max\{z_n - \xi, 0\}\Big)\Big] - bz_n \\ &= H(z_n) + \alpha E\Big[V^*\Big(\max\{z_n - \xi, 0\}\Big)\Big]. \end{aligned}$$

Entonces,

$$\lim_{z_n\to\infty} V^*(z_n) = H(S_*) + \alpha E\Big[V^*\Big((S_* - \xi)^+\Big)\Big]. \tag{2.4}$$

De manera similar, para $\{z'_n\}$,

$$V^*(z'_n) = \min_{a\in A(z'_n)} \{L(z'_n + a) - bz'_n\}.$$

Note que $z'_n + a \geqslant S_*$, además se tiene que en (S_*, ∞), el minimizador es S^*, pues no hay minimizadores de L en (S_*, S^*), por definicion de éstos, así $z'_n+a = S^*$, es decir, $a = S^* - z'_n$. Entonces

$$\begin{aligned} V^*(z'_n) &= L(z'_n + S^* - z'_n) - bz'_n \\ &= L(S^*) - bz'_n \\ &= H(S^*) + b(S^* - z'_n) + \alpha E\Big[V^*\Big((S^* - \xi)^+\Big)\Big]. \end{aligned}$$

Aplicando límite, tenemos

$$\lim_{n\to\infty} V^*(z'_n) = H(S^*) + b(S^* - S_*) + \alpha E\Big[V^*\Big((S^* - \xi)^+\Big)\Big]. \tag{2.5}$$

Como $S_* < S^*$, tenemos que (2.4) y (2.5) son distintas y ésto contradice la continuidad de V^*. Así, se concluye que L es no decreciente en (S_α, ∞). □

2.2.4. Caso 3. Espacios de Estados y Acciones no Acotados

En este caso consideramos $X = A = A(x) = [0, \infty)$, $x \in X$.

Hipótesis 2.6. Se cumple:

- La función H es continua en X.
- La función L definida como $L(y) := H(y) + by + \alpha E[V^*(y - \xi)^+]$, es inf-compacta.
- La distribución F de nuestra variable ξ tiene función de densidad continua y acotada ρ.

Teorema 2.7. Considere el modelo y la función de costos descritos en 2.2.1, los espacios de estados y acciónes definidos arriba y las Hipótesis 2.6. Entonces la función de valor V^* es continua y existe una política umbral óptima f_S con umbral $S \in X$.
DEMOSTRACIÓN.
Se tiene que la multifunción $\psi : X \to A$, con $\psi(x) = A(x) = [0, \infty)$ es cerrado-valuada y continua, ya que $A(x)$ es cerrado y constante para toda $x \in X$.

La función de costo en un paso es inf-compacta en $\mathbb{K}$ pues, $\forall\, r \in \mathbb{R}$, se tiene que $\{a \in A : \ C(x, a) \leqslant r\}$ es cerrado, debido a que es una bola cerrada en $\mathbb{R}$, y acotado, de ésto que sea compacto. También se tiene que el kérnel estocástico es fuertemente continuo (*ver A.4, Apéndice A*).

Observe que $L(y) = H(y)+by+\alpha E\Big[V^*((y-\xi)^+)\Big]$ es continua, pues $E\Big[V^*((y-\xi)^+)\Big]$ lo es por la Proposición C.2 (*ver Apéndice C*) y la continuidad de H se tiene por hipótesis.

Note que L tiene ínfimo en X, ya que la función es continua y acotada inferiormente, pues es no negativa. Ahora, definimos $\beta = \text{ínf}_{y \in X} L(y) \in \mathbb{R}$; por la inf-compacidad de L, se tiene que el conjunto $A = \{y \in X : \ L(y) \leqslant \text{ínf}_{x \in X} L(x)\}$ es compacto, observe que $A = \{y \in X : \ L(y) = \text{ínf}_{x \in X} L(x)\}$, ya que no existe $y \in X$ tal que $L(y) < \text{ínf}_{x \in X} L(x)$, este conjunto es no vacío por la Proposición C.6.
Así, podemos definir

$$S_\alpha := \text{máx}\{y \in X : \ L(y) = \min_{z \in X} L(z)\}.$$

Definimos ahora,

$$\hat{S} := \sup\{S > S_\alpha : \ L \text{ es creciente en } (S_\alpha, S)\}.$$

Si $\hat{S} = \infty$, entonces L es creciente en $[S_\alpha, \infty)$ y tenemos el resultado deseado.

Si $\hat{S} < \infty$, suponemos que L es no creciente y podemos definir

$$S^* := \text{ínf}\{y \in (\hat{S}, \infty) : \ L(y) = \inf_{x \in [\hat{S}, \infty)} L(x)\}.$$

Dado que L es continua y acotada inferiormente, existe el ínfimo de la función. Así que el conjunto sobre el cual se define S^* es compacto, por la inf-compacidad de L y dada la compacidad del conjunto se tiene que S^* es finito.

Ahora, tenemos que $L(S_\alpha) < L(S^*) < L(\hat{S})$ y L es continua, así por el teorema del valor intermedio, se tiene que el conjunto $\{y \in (S_\alpha, \hat{S}) : \ L(y) = \inf_{x\in[\hat{S},\infty)} L(x)\}$ es no vacío. Así, podemos definir

$$S_* := \sup\{y \in (S_\alpha, \hat{S}) : \ L(y) = \inf_{x\in[\hat{S},\infty)} L(x)\}.$$

Sean $\{z_n\}$ y $\{z'_n\}$ sucesiones en X tales que $z_n \uparrow S_*$ y $z'_n \downarrow S_*$. Para $\{z_n\}$, tenemos

$$V^*(z_n) = \min_{a\in A(x)} \{L(z_n + a) - bz_n\}.$$

Se tiene que L es creciente en $(S_\alpha, \hat{S})$ así la acción que minimiza es $a = 0$, entonces

$$\begin{aligned} V^*(z_n) &= L(z_n) - bz_n \\ &= H(z_n) + \alpha E\Big[V^*((z_n - \xi)^+)\Big]. \end{aligned}$$

Aplicando límite,

$$\lim_{n\to\infty} V^*(z_n) = H(S_*) + \alpha E\Big[V^*\Big((S_* - \xi)^+\Big)\Big]. \tag{2.6}$$

Análogamente,

$$V^*(z'_n) = \min_{a\in A(x)} \{L(z'_n + a) - bz'_n\}.$$

Como $z_n + a \geqslant S_*$ y el minimizador en (S_*, ∞) es S^*, se tiene que $a = S^* - z_n$, así

$$\begin{aligned} V^*(z'_n) &= L(S^*) - bz'_n \\ &= H(S^*) + b(S^* - z'_n) + \alpha E\Big[V^*\Big((S^* - \xi)^+\Big)\Big]. \end{aligned}$$

Entonces,

$$\lim_{n\to\infty} V^*(z'_n) = H(S^*) + b(S^* - S_*) + \alpha E\Big[V^*\Big((S^* - \xi)^+\Big)\Big]. \tag{2.7}$$

Por definición $S_* < S^*$, entonces las ecuaciones (2.6) y (2.7) son distintas, es decir, los límites de V^* para las sucesiones $\{z_n\}$ y $\{z'_n\}$ son distintos aunque éstas convergen al mismo punto S_*, ésto contradice la unicidad del límite, dada la continuidad de V^*, de aquí que la función L sea creciente.

$\square$

Capítulo 3

APLICACIONES

Se mencionarán dos aplicaciones de las políticas umbral, en el contexto de inventarios y líneas de espera. A continuación se darán breves descripciones de ambas.

3.1. Inventarios

Un inventario es un conjunto de mercancias o artículos acumulados en un almacén en espera de ser vendidos o utilizados en un proceso de producción. En este caso estamos interesados en la modelación del flujo de mercancía en el inventario observándolo como un sistema dinámico estocástico.

Los conceptos básicos en los sistemas de inventarios, son:

- Demanda. Cantidad de bienes o servicios que se ofrecen,
- Tiempo de espera. El tiempo que transcurre desde que se hace el pedido hasta que la empresa recibe el producto,
- Tamaño del pedido. Número de artículos que conforman la orden de pedido,

- Nivel de inventario. Número de artículos que se encuentran conformando el inventario,
- Punto de reorden. Nivel de inventario en el que la empresa define que es momento de hacer un nuevo pedido (*ver* [2]).

En teoría de inventarios, podemos ver el umbral como la cantidad de producto en *stock* que se desea tener en cada instante de tiempo. Así, lo que nos proporciona el umbral es un equilibrio en los costos del sistema pues, tener exceso de inventario, aumenta costos de almacenaje y tener poco producto en almacen genera costos por demanda no cumplida.

3.2. Líneas de Espera

Una cola o línea de espera es el efecto resultante en un sistema cuando la demanda de un servicio supera la capacidad de proporcionarlo.
El proceso básico de los modelos de colas es el siguiente. Los clientes que requieren un servicio se generan en el tiempo en una fuente de entrada. Luego entran al sistema y se unen a una cola. En determinado momento se selecciona un cliente de la cola para proporcionarle el servicio mediante alguna regla conocida como disciplina de servicio. Se lleva a cabo el servicio que el cliente requiere mediante un mecanismo de servicio y después el cliente sale del sistema de colas.

La estructura básica de un modelo de colas consta de:

- Fuente de entrada: Clientes potenciales, población que puede requerir del servicio en algún momento,
- Cola: Sistema físico donde los clientes esperan antes de recibir servicio. Una cola se caracteriza por el número máximo permisible de clientes que pueden admitir,
- Disciplina de servicio: Orden en el que se selecciona el cliente que recibirá el servicio,
- Mecanismo de servicio: Consiste en las instalaciones de servicio y el número de servidores en cada una de éstas.

La notación que se dará a continuación fue establecida por Kendall y es llamada notación Kendal extendida, la cual facilita el manejo de las características del modelo de colas que se utiliza.
Considere la etiqueta

$$A/B/C/D/E$$

donde

- A es la distribución de los tiempos de arribos,
- B es la distribución de los tiempos de servicio,
- C da el número de servidores,
- D denota la capacidad del sistema,
- E da la disciplina de atención en el sistema (*ver* [5]).

Trabajaremos modelos con disciplina de servicio FIFO, que son las siglas de *first in first out*, lo cual denota el tipo de servicio en el cual el primero que llega es el primero en ser atendido.

3.3. Ejemplo

El ejemplo que se dará a continuación es útil tanto en inventarios como en líneas de espera. Primero será descrito en el contexto de inventarios y seguido de ésto se darán las respectivas especificaciones para líneas de espera.

3.3.1. Descripción para Inventarios

Considere un modelo de inventarios con un único producto cuya dinámica está dada por

$$x_{t+1} = (x_t + a_t - \xi_t)^+, \tag{3.1}$$

con $t \in \mathbb{N}_0$ y $x_0 = x$, donde

- x_t representa la cantidad de cierto producto al tiempo t,

- a_t denota la cantidad de producto ordenada que se proporciona al principio del periodo t,
- ξ_t representa la demanda del producto durante el periodo t,

y con función de costo

$$C(x, a) := H(x + a) + ba, \tag{3.2}$$

donde $H(x + a) = pE[\text{máx}\{0, \xi - (x + a)\}] + hE[\text{máx}\{0, x + a - \xi\}]$ y

- b es el costo de producción por unidad,
- h es el costo de almacenaje por unidad,
- p es el costo por demanda no cumplida.

con $b,\ p,\ h > 0$ y $p > b$, de ésto que C sea no negativa.

Las variables aleatorias $\{\xi_t\}$ son independientes e idénticamente distribuidas, con F y ρ su función de distribución de probabilidad y su densidad, respectivamente, y con $E[\xi_t] < \infty$.

Los espacios de estados, acciones y acciones admisibles son los mecionados en el Capítulo 2, es decir, espacios de estados y acciones acotados, espacio de estados no acotado y espacio de acciones acotado y espacios de estados y acciones no acotados.

Se quiere minimizar el costo esperado descontado, dado por

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

Veremos que la función de costo es continua, acotada inferiormente e inf-compacta. Es fácil notar que es acotada pues toma valores no negativos, de esto que $C(x, a) \geqslant 0$ para toda pareja $(x, a) \in \mathbb{K}$.

Recordemos que

$$\text{máx}\{x,y\} = \frac{x+y+|x-y|}{2}, \quad x,y \in \mathbb{R}.$$

Así,

$$\begin{aligned} H(x+a) &= hE\Big[\frac{x+a-\xi+|x+a-x|}{2}\Big] + pE\Big[\frac{\xi-x-a+|x+a-\xi|}{2}\Big] \\ &= \frac{h-p}{2}(x+a) + \frac{p-h}{2}E[\xi] + \frac{h+p}{2}E[|x+a-\xi|], \quad (x,a) \in \mathbb{K}. \end{aligned}$$

Definimos

$$g(x,a) = E[|x+a-\xi|] = \int |x+a+s|\ \rho(s)\ ds.$$

Lema 3.1. g es continua en $\mathbb{K}$.
Demostración.
Considere $\{x_n\}$ y $\{a_n\}$ sucesiones convergentes en X y A, con límites x y a, respectivamente. Ahora, definimos h_n y h como:

$$\begin{aligned} h_n(s) &= |x_n + a_n - s|\ \rho(s), \\ h(s) &= |x+a-s|\ \rho(s), \end{aligned}$$

note que $h_n(s) \to h(s)$, para $s \in [0,\infty)$, pues ρ es no negativa.

Por otra parte

$$\begin{aligned} h_n(s) &\leqslant (|x_n| + |a_n| + s)\ \rho(s) \\ &\leqslant (M+s)\ \rho(s). \end{aligned}$$

Esta desigualdad se debe a que $\{x_n\}$ y $\{a_n\}$ son convergentes, de ésto que sean acotadas, así

$$\int h_n(s)\, ds \leqslant \int (M+s)\ \rho(s)\ ds = M + E[\xi] < \infty.$$

Ahora, por el Teorema de Convergencia Dominada (*ver* [3]), tenemos

$$\begin{aligned}
\lim_{n\to\infty} g(x_n, a_n) &= \lim_{n\to\infty} \int |x_n + a_n - s|\, \rho(s)\, ds \\
&= \int \lim_{n\to\infty} |x_n + a_n|\, \rho(s)\, ds \\
&= \int |x + a - s|\, \rho(s)\, ds \\
&= g(x, a).
\end{aligned}$$

Así, g es continua en $\mathbb{K}$.

□

Del Lema 3.1 se puede garantizar que H es continua.

Lema 3.2. La función de costo C es inf-compacta, es decir, el conjunto $A_\lambda(x) = \{a \in A(x) : \ C(x, a) \leqslant \lambda\}$ es compacto, para todo $x \in X$ y $\lambda \in \mathbb{R}$.

DEMOSTRACIÓN.
Por definición de C,

$$\lim_{a\to\infty} C(x, a) = \infty.$$

Sean $x \in X$ y $\lambda \in \mathbb{R}$, afirmamos que $A_\lambda(x)$ es acotado, en caso contrario existiría $\{a_n\} \subset A_\lambda(x)$, con $a_n \to \infty$, y

$$\lim_{n\to\infty} C(x, a_n) = \infty,$$

lo que implica que $\lambda \geqslant \infty$, lo cual es una contradicción. Ahora, sea $\{a_n\}$ una sucesión en $A_\lambda(x)$, convergente a algún $a' \in A$, tenemos que $0 \leqslant C(x, a_n) \leqslant \lambda$ y C es continua, entonces $0 \leqslant C(x, a') \leqslant \lambda$, de esto que $a' \in A_\lambda(x)$ y así $A_\lambda(x)$ es cerrado. Por el teorema de Heine-Borel, se tiene que $A_\lambda(x)$ es compacto, y ésto sucede para cualquier $x \in X$ y $\lambda \in \mathbb{R}$, se concluye que C es inf-compacta en $\mathbb{K}$.

□

De manera análoga se prueba que $E\Big[V^*(\max\{0, y - \xi\})\Big]$ es inf-compacta en $\mathbb{K}$.

Con esto, tenemos que el ejemplo cumple la Hipótesis 2.6, así, se puede aplicar alguno de los Teoremas 2.3, 2.5 o 2.7, dependiendo de las especificaciones de los espacios de estados y acciones que consideremos.

Ahora, usaremos iteración de valores para encontrar la forma general de V_n y la política umbral. Considere $V_0(x) \equiv 0$, así

$$\begin{aligned} V_1(x) &= \min_{a \in A(x)} \{C(x,a) + \alpha E[V_0(x+a-\xi)^+]\} && (3.3)\\ &= \min_{a \in A(x)} \{C(x,a)\} && (3.4)\\ &= \min_{y \in B(x)} \{H(y) + b(y-x)\}. && (3.5) \end{aligned}$$

Definición 3.3. Sea $y := x + a$, así tenemos que $B(x)$ se define como

- $B(x) = [x, K]$ para el caso de espacios de estados y acciones $X = A = [0, K]$ y acciones admisibles $A(x) = [0, K - x]$,
- $B(x) = [x, K + x]$ para el caso de espacio de estados $X = [0, \infty)$ y acciones admisibles $A = A(x) = [0, K]$,
- $B(x) = [x, \infty)$ para el caso de espacios de estados acciones y acciones admisibles no acotados, es decir $X = A = A(x) = [0, \infty)$.

Por el criterio de la segunda derivada, tenemos que el mínimo se alcanza en

$$y = F^{-1}\Big(\frac{p-b}{h+p}\Big) \ \Rightarrow \ a_1 = F^{-1}\Big(\frac{p-b}{h+p}\Big) - x$$

y

$$V_1(x) = H\Big(F^{-1}\Big(\frac{p-b}{p+h}\Big)\Big) + b\Big(F^{-1}\Big(\frac{p-b}{p+h}\Big) - x\Big).$$

Ahora,

$$\begin{aligned} V_2(x) &= \min_{a \in A(x)} \{C(x,a) + \alpha E[V_1(x+a-\xi)^+]\}\\ &= \min_{y \in B(x)} \{H(y) + b(y-x) + \alpha E[V_1(y-\xi)^+]\}\\ &= \min_{y \in B(x)} \Big\{H(y) + b(y-x) + \alpha\Big\{H\Big(F^{-1}\Big(\frac{p-b}{h+p}\Big)\Big)\\ &+ b\Big(F^{-1}\Big(\frac{p-b}{p+h}\Big) - E[(y+\xi)^+]\Big)\Big\}\Big\}. \end{aligned}$$

Usando nuevamente el criterio de la segunda derivada, llegamos a

$$y = F^{-1}\Big(\frac{p-b}{p+h-\alpha b}\Big) \Rightarrow a_2 = F^{-1}\Big(\frac{p-b}{p+h-\alpha b}\Big) - x,$$

entonces

$$\begin{aligned} V_2(x) &= H\Big(F^{-1}\Big(\frac{p-b}{p+h-\alpha b}\Big)\Big) + b\Big(F^{-1}\Big(\frac{p-b}{p+h-\alpha b}\Big) - x\Big) \\ &\quad +\alpha H\Big(F^{-1}\Big(\frac{p-b}{p+h}\Big)\Big) + \alpha b F^{-1}\Big(\frac{p-b}{p+h}\Big) \\ &\quad +\alpha E\Big[\Big(F^{-1}\Big(\frac{p-b}{p+h-\alpha b} - \xi\Big)^{+}\Big]. \end{aligned}$$

Si se sigue iterando, llegamos a que la forma general para la política está dada por

$$a_1 = f_1(x) = \begin{cases} S_1 - x, & \text{si } x \leqslant S_1, \\ 0, & \text{si } x \geqslant S_1. \end{cases} \tag{3.6}$$

donde el umbral es $S_1 = F^{-1}\Big(\frac{p-b}{p+h}\Big)$ y para $n \geqslant 2$

$$a_n = f_n(x) = \begin{cases} S_2 - x, & \text{si } x \leqslant S_2, \\ 0, & \text{si } x \geqslant S_2. \end{cases} \tag{3.7}$$

con umbral $S_2 = F^{-1}\Big(\frac{p-b}{p+h-\alpha b}\Big)$, y V_n es de la forma

$$V_n(x) = \begin{cases} 0, & \text{si } n = 0, \\ H\Big(F^{-1}\Big(\frac{p-b}{p+h}\Big)\Big) + b(F^{-1}\Big(\frac{p-b}{p+h}\Big) - x\Big), & \text{si } n = 1, \\ \\ \sum_{i=0}^{n-2} \Big\{\alpha^i \Big[H\Big(F^{-1}\Big(\frac{p-b}{p+h}\Big)\Big) + bF^{-1}\Big(\frac{p-b}{p+h}\Big)\Big] & \\ +\alpha^{i+1} b\Big[\int_{-\infty}^{F^{-1}(\frac{p-b}{p+h-\alpha b})} s\rho(s)ds - \Big(\frac{p-b}{p+h-\alpha b}\Big)F^{-1}\Big(\frac{p-b}{p+h-\alpha b}\Big)\Big]\Big\} & \\ +\alpha^n\Big\{H\Big(F^{-1}\Big(\frac{p-b}{p+h}\Big)\Big) + bF^{-1}\Big(\frac{p-b}{p+h}\Big)\Big\}, & \text{si } n \geqslant 2. \end{cases}$$

3.3.2. Descripción para Líneas de Espera

Para líneas de espera, las variables de la dinámica descrita en la ecuación (3.1), denotan:

- x_t, número de clientes en el sistema al tiempo t,
- a_t, número de clientes a los cuales se les permite el acceso al sistema,
- ξ_t, número de servicios concluidos al tiempo t.

Y para la función de costos dada en (3.2), tenemos que

- p es el costo de rechazo de clientes,
- h es el costo de servicio,
- b es el costo por arribo de clientes.

En líneas de espera, el umbral es el número de clientes que se desean tener en el sistema, para minimizar costos. Es importante estudiar los casos de capacidad de cola finita e infinita de sistemas de espera, en el caso de capacidad finita es importante controlar la entrada de clientes ya que un sobrecupo puede provocar inestabilidad en el sistema. Por otro lado, en el caso de capacidad infinita podemos pensar que admitir a todos los clientes que buscan servicio es la acción correcta pero, en la función de costos se contempla el costo por clientes en el sistema, así tener demasiados clientes en la cola, a largo plazo no resulta rentable.

Capítulo 4

SIMULACIONES

En este capítulo realizaremos simulaciones del Ejemplo planteado en el capítulo anterior para mostrar numéricamente que los sistemas de líneas de espera y de inventarios controlados mediante estrategias umbral generan menores costos comparados con sistemas no controlados.

4.1. Modelo de Línea de Espera Controlado con Capacidad del Sistema Finita

Considere una línea de espera con tiempos de servicios distribuidos exponencialmente, con un único servidor, capacidad del sistema finita y disciplina de servicio FIFO. Note que este sistema tiene espacio de estados finito, es decir, admite un número finito K de clientes en el servidor.
Sea $K = 10$, así el espacio de estados, está dado por $X = [0, 10]$, de esto que el número de clientes que pueden ingresar al sistema en un tiempo t depende de cuantos haya en él al tiempo t, así $A(x) = [0, 10 - x]$, el espacio de acciones está dado por $A = [0, 10]$ pues no pueden ingresar más clientes que la capacidad del sistema. Sea $\lambda = 1$ el parámetro de la variable exponencial que determina los servicios concluidos, es decir, en promedio se concluye 1 servicio por cada tiempo en el que se observa el sistema.

Definimos la función de costo como

$$C(x,a) = 35E[(\xi - (x+a))^+] + 0.9E[(x+a-\xi)^+] + 7a.$$

Se tiene que la función de valor, satisface la ecuación de optimalidad, es decir,

$$V^*(x) = \min_{a \in A(x)} \{C(x,a) + \alpha E[V^*((x+a-\xi)^+)]\}.$$

Tomando como factor de descuento $\alpha = 0.9$, se tiene

$$V^*(x) = \min_{a \in [0,10-x]} \{C(x,a) + 0.9E[V^*((x+a-\xi)^+)]\}.$$

Ahora, por (3.7), se tiene que el umbral para este ejemplo está dado por

$$S = F^{-1}\Big(\frac{p-b}{p+h-\alpha b}\Big) = F^{-1}(0.9756) = 3.71 \approx 4.$$

Así, se tiene que la política umbral óptima está dada por

$$f(x) = \begin{cases} 4 - x, & \text{si} \quad x \leqslant 4, \\ \\ 0, & \text{si } x > 4. \end{cases}$$

A continuación mostraremos en una tabla para algunos estados iniciales del sistema, los costos que se generan al usar políticas umbral y sin usarlas, después de 90 iteraciones.

Estado	*Costo sin control*	*Costo con control*
1	70875	62070
3	74559	69660
5	74570	72900
8	73431	71025
10	76683	70195

donde la primera columna representa el estado inicial del sistema, es decir cuantos clientes llegan inicialmente a la cola; la segunda columna da costos generados si usar la política umbral y la tercera, da los costos obtenidos usando la estrategia umbral obtenida.

4.2. Modelo de Línea de Espera Controlado con Capacidad del Sistema Infinita

Considere una línea de espera con tiempos de servicios distribuidos exponencialmente, con un único servidor y capacidad del sistema infinita.
Así, el espacio de estados está dado por $X = [0, \infty]$, tenemos que el sistema es controlado así, se tiene que $A = A(x) = [0, K]$. Sea $\lambda = 1$ la tasa de servicios concluidos por periodo de observación.Considere $K = 12$.
Definimos la función de costo como

$$C(x, a) = 28E[(\xi - (x + a))^+] + 0.2E[(x + a - \xi)^+ + 0.95a].$$

Sabemos que la función de valor satisface la ecuación de optimalidad, de ésto que la podamos reescribir como

$$V^*(x) = \min_{a \in A(x)} \{C(x, a) + \alpha E[V^*((x + a - \xi)^+)]\}.$$

Tomando el factor de descuento $\alpha = 0.9$.
Se tiene que

$$V^*(x) = \min_{a \in A(x)} \{C(x, a) + 0.9E[V^*((x + a - \xi)^+)]\}.$$

Por (3.7),tenemos que el umbral para este ejemplo está dado por

$$S = F^{-1}\Big(\frac{p - b}{p + h - \alpha b}\Big) = F^{-1}(0.9892) = 4.529 \approx 5.$$

Así, se tiene que la política umbral óptima está dada por

$$f(x) = \begin{cases} 5 - x & \text{si} \quad x \leqslant 5 \\ 0 & \text{si } x > 5. \end{cases}$$

En la siguiente tabla, daremos algunos estados iniciales del sistema y los costos que se generan al controlar el sistema con la estrategia umbral encontrada y sin hacer uso de esta, después de 100 iteraciones.

Estado	*Costo sin control*	*Costo con control*
1	157011	147177
5	171597	154644
8	198224	165364
12	142619	114895.8
20	142179	129101

donde la primera columna da el número de clientes que llegan al sistema inicialmente; la segunda, el costo que se genera al observar el sistema sin aplicar una estrategia umbral y la tercera, ds los costos generados al usar la política umbral.

4.3. Modelo de Inventarios con Capacidad Infinita

Considere un sistema de inventarios con capacidad y producción infinita, es decir, $X = A = A(x) = [0, \infty)$.
Y con función de costo

$$C(x, a) = 30E[\text{máx}\{0, \xi - (x + a)\}] + 8E[\text{máx}\{x + a - \xi, 0\}] + 2a.$$

La demanda ξ sigue una distribución normal con parámetros $\mu = 5$ y $\sigma = 1$.

Tenemos que la función de valor satisface la ecuación de optimalidad, es decir,

$$V^*(x) = \min_{a \in A(x)} \{C(x, a) + 0.9E[V^*((x + a - \xi)^+)]\},$$

por (3.7), tenemos que el umbral S, está dado por

$$S = F^{-1}\Big(\frac{p - b}{p + h - \alpha b}\Big) = F^{-1}(0.9055) = 1.32.$$

Pero esa es la preimagen de nuestra variable estandarizada, de aquí que el umbral sea $S^* = 1.32\sigma + \mu = 7.64 \approx 8$.

Así, nuestra política umbral óptima está dada por

$$f(x) = \begin{cases} 8 - x, & \text{si} \quad x \leqslant 8, \\ \\ 0, & \text{si } x > 8. \end{cases}$$

Ahora, mostraremos una tabla que dado el nivel inicial de inventario, da los costos generados al aplicar políticas umbral y sin usarlas, después de 100 iteraciones.

Estado	*Costo sin control*	*Costo con control*
1	178285	175237
7	180711	171045
12	178225	128635
20	175902	151910
35	189032	136427

La primera columna representa el nivel de inventario del sistema al iniciar la observación, la segunda, los costos que se generan sin aplicar la estrategia umbral que se obtuvo y la tercera columna, muestra los costos que se generaron usando la estrategia obtenida.

CONCLUSIONES

El trabajo de tesis se centró en el estudio de la existencia de políticas umbral en el contexto de Procesos de Decisión de Markov. Se presentan condiciones en el modelo de control de Markov las cuales garantizan que la estrategia óptima presenta la forma umbral. El umbral en este caso se entiende como un valor en el espacio de estados, el cual define el cambio de regla entre una acción fija y otra. Las estrategias umbral son importantes desde el punto de vista práctico, ya que cuentan con una caracterización propia y de este modo su implementación es más sencilla.

El estudio de dichas políticas se ejemplificó en líneas de espera y teoría de inventarios. En la literatura se pueden encontrar textos específicos donde es natural considerar esta clase de estrategia (*ver* [9] *y* [14]). Sin embrago, en la literatura existente no se encuentran condiciones sobre el modelo las cuales garanticen a priori que la búsqueda de la estrategia es de tipo umbral. En este sentido el trabajo de tesis busca dar una respuesta inicial para caracterizar en base a las componentes del modelo de control la estrategia umbral. Como se menciona en la introducción, el trabajo en que se basó la tesis se encuentra sustentado en el reporte técnico *On the optimality of base stock policies for a class of inventory systems with no set up costs and no backorders*, de Óscar Vega-Amaya y Joaquín López-Borbón (*ver* [16]), en específico el Capítulo 2. Es importante señalar que las hipótesis que se trabajan en ese reporte fueron modificadas, considerando condiciones más simples de verificar en la práctica y que son más conocidas en el contexto de Procesos de Decisión de Markov. Al realizar dichas modificaciones las conclusiones del resultado principal siguen siendo válidas.

Finalmente el trabajo de tesis presenta un capítulo de aplicaciones donde se verificaron las hipótesis del teorema principal del trabajo de tesis en las áreas de teoría de inventarios y líneas de espera. Porteriormente se presentaron algunos ejemplos numéricos resueltos usando simulación y elaborados en R.

Trabajos futuros en esta línea contemplan la implementación de políticas umbral en procesos de crecimiento económico (*ver* [6]) y procesos de asignación de recursos (*ver* [7]). Desde el punto de vista teórico se puede estudiar la generalización de condiciones en la función de costo, considerando funciones de costo acotadas en norma ponderada y permitiendo cambio de signo, es decir, en etapas de observación del sistema que permiten obtener utilidades.

Apéndice A

Kérneles Estocásticos

Definición A.1. Un kérnel estocástico sobre X dado Y es una función $Q(\cdot \mid \cdot)$ tal que:

(a) $Q(\cdot|y)$ es una medida de probabilidad sobre X para cada $y \in Y$,

(b) $Q(B|\cdot)$ es una función medible en Y para cada $B \in \mathbb{B}(X)$.

El conjunto de todos los kérneles estocásticos sobre X dado Y es denotado por $\mathcal{P}(X|Y)$

Definición A.2. Un kérnel estocástico P se dice

1. Débilmente continuo si la función

$$y \mapsto \int \nu(x)P(dx|y),$$

es continua y acotada, para cada función ν continua y acotada en X.

2. Fuertemente continua si la función

$$y \mapsto \int \nu(x)P(dx|y),$$

es continua y acotada, para cada ν función acotada en X.

Definición A.3. Φ denota el conjunto de todos los kérneles estocásticos $\varphi \in \mathcal{P}(A|X)$ tales que $\varphi(A(x)|x) = 1$.

Observación A.4. Dada la función de distribución F con densidad ρ, continua y acotada, se tiene que la función

$$(x, a) \to \int_X \nu(y) Q(dy|x, a), \tag{A.1}$$

es continua y acotada, para cualquier función medible y acotada ν, en X. Pues,

$$\begin{aligned}
\int_X \nu(y)\, Q(dy|x, a) &= E_0\Big[\nu\Big((x + a - \xi)^+\Big)\Big] \\
&= \int_{-\infty}^{\infty} \nu\Big((x + a - \xi)^+\Big)\, \rho(\xi)\, d\xi \\
&= \int_0^{x+a} \nu(z)\, \rho(x + a - z)\, dz \;+\; \int_{x+a}^{\infty} \nu(0)\, \rho(x + a - z)\, dz \\
&= \int_0^{x+a} \nu(z)\, \rho(x + a - z)\, dz \;+\; \nu(0)\, [1 - F(x + a)].
\end{aligned}$$

Y se tiene que la función de densidad ρ es continua y acotada, de esto que la función (A.1) sea continua y acotada.

Apéndice B

Multifunciones y Selectores

Sean X y A espacios de Borel no vacíos.

Definición B.1. Una multifunción (también llamada correspondencia o mapeo conjunto valuado) $\psi : X \to A$, es una función tal que $\psi(x)$ es un subconjunto no vacío de A, para cada $x \in X$. El grafo de la multifunción ψ es el subconjunto de $X \times A$ dado por

$$\mathbb{K} := \{(x, a)|\ x \in X \quad \text{y} \quad a \in A\}.$$

Para cada subconjunto B de A, sea $\psi^{-1}[B] := \{x \in X|\ \psi(x) \cap B \neq \emptyset\}$.

Una multifunción $\psi : X \to A$ se dice compacto valuada si para cada $x \in X$, $\psi(x)$ es un subconjunto compacto de A.

Definición B.2. Una correspondencia $\psi : X \to Y$ se dice semicontinua inferior ($s.c.i$) en x, si $\psi(x)$ es no vacío y, para cada $y \in \psi(x)$ y cada sucesión $x_n \to x$, existen $N \geqslant 1$ y $\{y_n\}_{n\in\mathbb{N}}$ tal que $y_n \to y$ y $y_n \in \psi(x_n)$ para todo $n \geqslant \mathbb{N}$.

Definición B.3. Una multifunción compacto valuada, $\psi : X \to A$ es semicontinua superior ($s.c.s.$) en x si $\psi(x)$ es no vacío y, para cada sucesión $x_n \to x$ y cualquier sucesión $\{y_n\}_{n\in\mathbb{N}}$ tal que $y_n \in \psi(x_n)$, para todo $n \in \mathbb{N}$, existe una subsucesión convergente de $\{y_n\}$ cuyo punto límite y está en $\psi(x)$.

Definición B.4. Una multifunción se dice continua si es $s.c.s$ y $s.c.i.$.

Proposición B.5. Para $X = A = [0, K]$ y $A(x) = [0, k - x]$, la multifunción $x \to A(x)$ es continua.

Demostración.

Sea $y \in A(x)$ y $x_n \to x$, se tiene que $0 \leqslant y \leqslant K - x$. Definimos $y_n := x - x_n + y$, observe que $y_n \to y$, pues $x_n \to x$. Además

$$\begin{aligned} 0 \leqslant y \leqslant K - x &\Leftrightarrow 0 \leqslant y + x \leqslant K \\ &\Leftrightarrow 0 \leqslant x_n - x_n + y - x \leqslant K \\ &\Leftrightarrow 0 \leqslant x_n + y_n \leqslant K \\ &\Leftrightarrow 0 \leqslant y_n \leqslant K - x_n \\ &\Leftrightarrow y_n \in A(x_n) \end{aligned}$$

y esto sucede para todo $n \in \mathbb{N}$, de esto que la multifunción sea semicontinua inferior.

Ahora, considere $x_n \to x$ y $\{y_n\}_{n\in\mathbb{N}}$ tal que $y_n \in A(x_n)$ para todo $n \in \mathbb{N}$, es decir $0 \leqslant y_n \leqslant K + x_n \quad \forall n \in \mathbb{N}$. Se tiene que $\{y_n + x_n\} \subset [0, K]$ y este conjunto es compacto, de esto que exista $\{y_{n_i} + x_{n_i}\}$ subsucesión de $\{y_n + x_n\}$ tal que $\{y_{n_i} + x_{n_i}\} \to J$, para algún $J \in [0, K]$, también se tiene que $x_{n_i} \to x$, entonces $y_{n_i} \to J - x$, renombramos $y = J - x$ y así tenemos que $y_{n_i} \to y$, donde $0 \leqslant y \leqslant K - x$, de esto que la multifunción sea semicontinua superior. Así, por la definición B.4 se concluye lo deseado.

$\square$

Definición B.6. $\mathbb{F}$ denota el conjunto de todas las funciones medibles $f : X \to A$ que satisfacen $f(x) \in A(x)$, para toda $x \in X$. Las funciones en $\mathbb{F}$ son llamadas selectores de la multifunción $x \mapsto A(x)$.

Apéndice C

Otros Resultados

Proposición C.1. La función de valor óptimo es finita y acotada en conjuntos acotados.
DEMOSTRACIÓN.
Sea f_S una política umbral con $S > 0$, si $x_0 = x \in [0, S]$, entonces

$$\begin{aligned} x_{n+1} &= x_n + a_n - \omega_n \\ &= x_n + (S - x_n) - \omega_n \\ &= (S - \omega_n)^+, \end{aligned}$$

pues x_n no puede tomar valores negativos y esto sucede para toda $n \in \mathbb{N}$. Ahora, note que

$$\begin{aligned} V_\alpha(f_S, x) = E\Big[\sum_{t=0}^{\infty} \alpha^t \, C(x_t, a_t)\Big] \\ &= E\Big[C(x_0, a_0) + \sum_{t=1}^{\infty} \alpha^t \, C(x_t, a_t)\Big], \quad \text{por (2.1) podemos sustituir } C \\ &= E\Big[H(x_0 + a_0) + ba_0 \; + \; \sum_{t=1}^{\infty} \alpha^t \, \{H(x_t + a_t) + ba_t\}\Big] \end{aligned}$$

$$
\begin{aligned}
&= E\Big[H(x+S-x)+b(S-x) \\
&\quad + \sum_{t=0}^{\infty}\alpha^t \Big\{H\Big((S-\omega_{t-1})^+ + S-(S-\omega_{t-1})^+ + b(S-(S-\omega_t)^+)\Big)\Big\}\Big] \\
&= H(S)+bS-bx+E\Big[\sum_{t=1}^{\infty}\alpha^t\ \{H(S)+bS-b(S-\omega_t)^+\}\Big] \\
&= H(S)+bS-bx+\frac{\alpha}{1-\alpha}H(S)+\frac{\alpha}{1-\alpha}bS-bE\Big[\sum_{t=1}^{\infty}\alpha^t(S-\omega_t)^+\Big] \\
&= \frac{1}{1-\alpha}H(S)+\frac{1}{1-\alpha}bS-bx+\frac{\alpha}{1-\alpha}bE[\text{mín}\{0,\omega_0-S\}].
\end{aligned}
$$

Definimos

$$T(S) \ := \ H(S) \ + \ bS \ + \ \alpha bE[\text{mín}\{0,\omega_0-S\}].$$

Así,

$$
\begin{aligned}
V_\alpha(f_S,x) \quad &= \quad \frac{1}{1-\alpha}\,T(S) \ - \ bx \\
&\leqslant \quad \frac{1}{1-\alpha}\,T(S) \\
&= \quad V_\alpha(f_S,0).
\end{aligned}
$$

De esto que se cumpla la proposición.

$\square$

Además, si la política umbral es óptima, entonces

$$0 \leqslant \ V^*(0) \ = \ \left(\frac{1}{1-\alpha}\right)T(S) \ = \ \left(\frac{1}{1-\alpha}\right)\ \inf_{y\in X} T(y).$$

Esto implica que $T(S) = \text{ínf}_{y\in X}\,T(y)$, donde S es el umbral.
Esta ecuación caracteriza los puntos umbral y nos permite encontrar de manera explícita la política umbral.

Proposición C.2. $M(y) = E\Big[V^*\Big(\text{máx}\{y-\xi, 0\}\Big)\Big]$ es continua.

Demostración.

Observe que, haciendo $U = \text{máx}\{0, y-\xi\}$

$$E\Big[V^*\Big(\text{máx}\{y-\xi, 0\}\Big)\Big] = \int_{-\infty}^{\infty} V^*(U) f_U(U) dU. \tag{C.1}$$

Por otro lado,

$$\begin{aligned} P(U < u) &= P(\text{máx}\{0, y-\xi\} < u) \\ &= P(y - \xi < u) \\ &= P(y - u < \xi) \\ &= 1 - P(\xi \leqslant y - u) \\ &= 1 - F_\xi(y-u). \end{aligned}$$

Así, $f_U(u) = f_\xi(y-u)$, de esto que

$$E\Big[V^*\Big(\text{máx}\{y-\xi, 0\}\Big)\Big] = \int_{-\infty}^{\infty} V^*(U) f_\xi(y-U) dU.$$

Ahora, sea $\{y_n\}_{n\in\mathbb{N}} \subset X$ una sucesión convergente a un punto $y_0 \in X$ y sea S una cota superior del conjunto $\{y_n \,:\, n \in \mathbb{N}_0\}$, por la Proposición C.1 se tiene que V^* es acotada en $[0, S]$, así las funciones $V^*(\text{máx}\{0, y_n - \xi\})$, $n \in \mathbb{N}_0$, son uniformemente acotadas. Es decir, $\forall\, n \in \mathbb{N}_0$, $\Big|V^*(\text{máx}\{0, y_n - \xi\})\Big| \leqslant N$, para algún $N > 0$. Así, por el teorema de convergencia dominada, se tiene

$$\begin{aligned} \lim_{n\to\infty} \int_{-\infty}^{\infty} V^*(U)\ f_\xi(y_n - U)\ dU &= \int_{-\infty}^{\infty} \lim_{n\to\infty} V^*(U)\ f_\xi(y_n - U)\ dU \\ &= \int_{-\infty}^{\infty} V^*(U)\ f_\xi(y_0 - U)\ dU. \end{aligned}$$

De esto que M sea continua.

$\square$

Proposición C.3. Sea X una variable aleatoria P-integrable sobre $(\Omega, \mathscr{F}, P)$ y sean $\mathscr{A}$ y $\mathscr{B}$ σ-álgebras contenidas en $\mathscr{F}$. Si $\mathscr{A} \subset \mathscr{B}$ entonces,

$$E[X|\mathscr{A}] = E\Big[E[X|\mathscr{A}]\Big|\mathscr{B}\Big] = E\Big[E[X|\mathscr{B}]\Big|\mathscr{A}\Big].$$

Proposición C.4. Sea ν una variable aleatoria P-integrable en $(\Omega, \mathscr{F}, P)$ y $\mathscr{G}$ sub σ-álgebra de $\mathscr{F}$ entonces, $E[E[\nu|\,\mathscr{G}]] = E[\nu]$.

Proposición C.5. Sean ν y ν' variables aleatorias P-integrables en $(\Omega, \mathscr{F}, P)$ y $\mathscr{G}$ y $\mathscr{G}'$ sub σ-álgebras de $\mathscr{F}$. Si ν es $\mathscr{G}$-medible entonces,

$$E[\nu\nu'|\,\mathscr{G}] = \nu E[\nu'|\,\mathscr{G}],$$

particularmente

$$E[\nu|\,\mathscr{G}] = \nu.$$

Proposición C.6. Suponga que $\mathbb{K}$ es un subconjunto de Borel de $x \times A$, ν es s.c.i., acotada inferiormente e inf-compacta en $\mathbb{K}$. Entonces

1. Existe un selector $f^* \in \mathbb{F}$ tal que

$$\nu(x, f^*(x)) = v^*(x) = \min_{A(x)} \nu(x, a).$$

2. Si la multifunción $x \mapsto A(x) := \{a \in A(x)|\ \nu^*(x) = \nu(x, a)\}$ es s.c.i., entonces ν^* es s.c.i.. Más aún, si ν es continua entonces ν^* también lo es.

Apéndice D

Continuidad de V^*

D.1. Caso 1. Espacios de Estados y Acciones Acotados Caso 2. Espacio de Estados no Acotado y Espacio de Acciones Acotado

Hipótesis D.1. Sea

i) Para cada $x \in X$, el conjunto de acciones admisibles, $A(x)$, es un subconjunto compacto y no vacío de A.

ii) Para alguna constante R, $|C(x,a)| \leqslant R, \quad \forall\ (x,a) \in \mathbb{K}$, más aún, para cada $x \in X$, la función $C(x,a)$ es continua en $a \in A(x)$.

iii) $\int \nu(x) Q(dy|x,a)$ es una función continua, para cada función $\nu \in B(X)$.

Teorema D.2. Si se cumplen las hipótesis (D.1), se tiene que

(a) La función de valor óptimo ν^* es la única solución en $B(X)$ de la ecuación de programación dinámica

$$\nu^*(x) = \min_{a \in A(x)} \left\{ C(x,a) + \alpha \int \nu^*(y) Q(dy|x,a) \right\}, \quad \text{para } x \in X.$$

(b) Una política estacionaria $f^* \in \mathbb{F}$ es óptima, si y sólo sí, f^* minimiza el lado derecho de la ecuación de programación dinámica para todo $x \in X$, es decir,

$$\nu^*(x) = C\Big(x, f^*(x)\Big) + \alpha \int \nu^*(y) Q\Big(dy|x, f^*(x)\Big), \quad \text{para todo } x \in X.$$

Proposición D.3. Sea $\Psi : X \to C(A)$ una multifunción Borel medible, y sea $\nu(x, a)$ una función medible, real valuada, en $\mathbb{K}$; tal que $\nu(x, a)$ es semicontinua superiormente ($s.c.s.$) en $a \in \Psi(x)$ para cada $x \in X$. Entonces

a) Existe un selector $f : X \to A$ para Ψ tal que

$$\nu\Big(x, f(x)\Big) = \min_{a \in \Psi(x)} \nu(x, a), \quad \text{para cada } x \in X$$

y la función $\nu^*(x) := \min_{a \in \Psi(x)} \nu(x, a)$ es medible.

b) Si Ψ y ν son $s.c.s$ y acotadas, entonces ν^* es $s.c.s.$ y acotada.

c) Si Ψ es continua y ν es continua y acotada, entonces ν^* es continua y acotada.

d) Si Ψ es continua y A es compacto, entonces $\mathbb{K}$ es cerrado.

D.2. Caso 3. Espacios de Estados y Acciones no Acotados

Lema D.4. Sean $\varphi : X \to Y$ una correspondencia s.c.i. y $f : Gr\varphi \to \mathbb{R}$ una función s.c.i.. Definimos $m : X \to \mathbb{R}$ como

$$m(x) = \inf_{y \in \varphi(x)} f(x, y),$$

entonces la función m es continua.
Demostración.
Veremos que $\{x \in X : \ m(x) < \alpha\}$ es abierto para cualquier $\alpha \in \mathbb{R}$.
Suponga que $\exists x_0 \in X$ con $m(x_0) < \alpha$, entonces existe $a_0 \in \varphi(x_0)$ tal que

$f(x_0, a_0) < \alpha$, como f es s.c.i., el conjunto $W = \{(x, y) \in \mathbb{K} : \ f(x, y) < \alpha\}$ es una vecindad abierta de $(x_0, a_0) \in \mathbb{K}$. De aquí que existan U y V abiertos con $x_0 \in U$ y $a_0 \in V$, tales que $U \times V \cap \mathbb{K} \subset W$.

Note que $N = U \cap \varphi^{-1}(V)$ es una vecindad abierta de x_0. Ahora, para cada $x \in N$, existe $y \in \varphi(x) \cap V$ tal que, $(x, y) \in U \times V \cap \mathbb{K} \subset W$, entonces $f(x, y) < \alpha$ así, $m(x) < \alpha$ y esto sucede para toda $x \in N$. Así $N \subset \{x \in X : \ m(x) < \alpha\}$, de esto que $\{x \in X : \ m(x)\alpha\}$ sea abierto. Y se concluye que m es continua.

$\square$

Bibliografía

[1] Aliprantis, Charalambos D. y Border, Kim C., *Infinite Dimensional Analysis: A Hitchhiker's Guide*, tercera edición, Springer, Alemania, 2006.

[2] Ángel López, Liliana Karina y López Garcia, Karina, *Desarrollo de un sistema simulador y optimizador de inventarios* , Escuela de Ingeniería y Ciencias, Universidad de las Américas Puebla, México, 2006.

[3] Bartle, Robert G., *The Elements of Integration*, John Wiley & Sons,Inc, EUA, 1996.

[4] Bäuerle, Nicole y Rieder, Ulrich, *Markov Decision Processes with Applications to Finance*, Springer-Verlag, Alemania, 2011.

[5] Bolch, Gunter; Greiner, Stefan; de Meer, Hermann y Trivedi, Kishor S., *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*, John Wiley & Sons, Inc., EUA, 1998.

[6] Cruz-Suárez, Hugo; Montes-de-Oca, Raúl y Zacarías, Gabriel, *A Consumption-Invesment Problem Modelled as a Discounted Markov Decision Process* , Kybernetica, Vol. 47, No.6, pp.909-929, 2011.

[7] Dolgov, Dmitri y Durfee, Edmund, *Computationally-efficient combinatorial auctions for resourse allocation in weakly-coupled MDPs*, In Proceedings of the 4th International Joint Conference on Autonomous Agents and Multiagent Systems, 2005.

[8] GHIANI, GIANPAOLO, LAPORTE, GILBERT y MUSMANNO, ROBERTO, *Introduction to Logistic Systems Planning and Control* , John Wiley & Sons, Ltd , Inglaterra, 2004.

[9] HASSIN, RAFAEL y HAVIV, MOSHE, *To Queue Or Not To Queue: Equilibrium Behavior in Queueing Systems*, Kluwer Academic Publishers, EUA, 2003.

[10] HERNÁNDEZ-LERMA, ONÉSIMO y LASSERRE, JEAN BERNARD, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, Springer-Verlag, EUA, 1996 .

[11] HILLIER, FREDERICK S. y LIEBERMAN, GERALD J., *Introducción a la Investigación de Operaciones*, novena edición, McGraw Hill, México, 2010.

[12] KITAEV, MIKHAIL YU. y RYKOV, VLADIMIR V., *Controlled Queueing Systems*, primera edición, CRC Press, EUA, 1995.

[13] LINDLEY, D. V., *The theory of queues with a single server*, Proceedings of the Cambridge Philosophical Society, vol. 48, issue 02, p. 277, 1951.

[14] PORTEUS, EVAN, *Foundations of Stochastic Inventory Theory*, primera edición, Stanford Business Books, EUA, 2002.

[15] STOKEY, NANCY L. y LUCAS, JR., ROBERT E., *Recursive Methods in Economic Dynamics*, Harvard University Press, EUA, 1999.

[16] VEGA-AMAYA, ÓSCAR y LÓPEZ-BORBÓN, JOAQUÍN, *On the optimality of base-stock policies for a class of inventory with no set-up cost and no backorders*, Departamento de Matemáticas, Universidad de Sonora, México, 2013.

[17] ZACARÍAS ESPINOZA, GABRIEL, *Procesos de Decisión de Markov Descontados*, Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, México, 2007

Índice alfabético