



Universidad Juárez Autónoma de Tabasco

División Académica de Ciencias Básicas

CONTROL DE LÍNEAS DE ESPERA CON COSTO PROMEDIO

TESIS

Que para obtener el grado de

**Maestra en Ciencias en
Matemáticas Aplicadas**

P R E S E N T A:

Lic. Mat. Mirelda Dionicio Arévalo

DIRECTOR DE TESIS:

Dr. Heliodoro Daniel Cruz Suárez

Cunduacán, Tab. México

Septiembre 2014

CARTA DE AUTORIZACIÓN

El que suscribe, autoriza por medio del presente escrito a la Universidad Juárez Autónoma de Tabasco que utilice tanto física como digitalmente la tesis de grado denominada "**CONTROL DE LÍNEAS DE ESPERA CON COSTO PROMEDIO**", de la cual soy autor y titular de los Derechos de Autor.

La finalidad del uso por parte de la Universidad Juárez Autónoma de Tabasco de la tesis antes mencionada, será única y exclusivamente para difusión, educación y sin fines de lucro; autorización que se hace de manera enunciativa más no limitativa para subirla a la Red Abierta de Bibliotecas Digitales (RABID) y a cualquier otra red académica con la que la Universidad tenga relación institucional.

Por lo antes manifestado, libero a la Universidad Juárez Autónoma de Tabasco de cualquier reclamación legal que pudiera ejercer respecto al uso y manipulación de la tesis mencionada y para los fines estipulados en este documento.

Se firma la presente autorización en la ciudad de Comalcalco, Tabasco a los 27 días del mes de Septiembre del año 2014.

AUTORIZO



MIRELDA DIONICIO ARÉVALO

EL TESISISTA



UNIVERSIDAD JUÁREZ
AUTÓNOMA DE TABASCO

"ESTUDIO EN LA DUDA. ACCIÓN EN LA FE"

DIVISIÓN ACADÉMICA DE CIENCIAS BÁSICAS

Dirección



"2014, Conmemoración del 150 Aniversario de
la Gesta Heroica del 27 de febrero de 1864"

10 de septiembre de 2014

Lic. en Mat. Mirelda Dionicio Arévalo
Pasante del Maestría en Ciencias en Matemáticas Aplicadas
Presente.

Por medio del presente y de la manera más cordial, me dirijo a Usted para hacer de su conocimiento que proceda a la impresión del trabajo titulado "**CONTROL DE LINEAS DE ESPERA CON COSTO PROMEDIO**", en virtud de que reúne los requisitos para el EXAMEN PROFESIONAL para obtener el grado de Maestro en Ciencias en Matemáticas Aplicadas

Sin otro particular, reciba un cordial saludo.

Atentamente,

Dr. Gerardo Delgadillo Piñón
Director



DIVISIÓN ACADÉMICA DE
CIENCIAS BÁSICAS

DIRECCIÓN.

C.c.p.- Archivo
Dr'GDP/Dr'GBG/emt

Miembro CUMEX desde 2008

Consortio de
Universidades
Mexicanas
UNA ALIANZA DE CALIDAD POR LA EDUCACIÓN SUPERIOR

Km. 1, Carretera Cunduacán-Jalpa de Méndez, A.P. 24, C.P. 86690, Cunduacán, Tabasco
Tel/Fax (914)3360928, (993)3581500 Ext. 6702

DEDICATORIAS

A DIOS

Por haberme regalado la vida y la salud, por su constante amor; sus bendiciones para poder concluir con este trabajo.

A MIS PADRES

Mirelda Arévalo Alcocer y Juan Dionicio de los Santos, como muestra de mi amor, cariño, respeto y admiración porque durante años de incesante dedicación, apoyo incondicional, genuina comprensión y sus sabios consejos han logrado guiar mi andar.

A DOS PERSONAS ESPECIALES

José Armando Dionicio Arévalo y Miguel Domínguez Martínez mis amores, gracias por formar parte de mi vida.

AGRADECIMIENTOS

Agradezco de todo corazón e infinitamente a Dios, ya que Él ha sido el centro de mi vida y porque me regalo la paciencia a lo largo de mi carrera y en el transcurso de la realización de este trabajo.

A mis padres por estar siempre conmigo brindándome su apoyo incondicional en todos los aspectos de mi vida.

En especial, quiero agradecer sinceramente y profundamente al Dr. Heliodoro Daniel Cruz Suárez quien fue el director de este trabajo de tesis; por la comprensión y gran paciencia mostrada durante su realización y por sus valiosas sugerencias y correcciones. Igualmente por sus consejos apoyo y motivación no sólo del trabajo sino también de mi formación matemática, que fueron factores que contribuyeron significativamente a mi desarrollo intelectual y con los cuales aprendí un paso más en el que hacer matemático.

A mis familiares, amigos y a todas aquellas personas que me brindaron su apoyo, amistad y confianza sin condición alguna a todos ellos GRACIAS...

Índice general

1. Líneas de Espera	7
1.1. Introducción a Líneas de Espera	7
1.2. Rendimiento de un Sistema	11
1.3. Modelos de Colas	12
1.4. Descripción de un Sistema de Espera	13
1.5. Cadenas de Markov	19
2. Programación Dinámica y Colas	31
2.1. Cadenas de Decisión de Markov a Tiempo Discreto	32
2.1.1. Políticas	34
2.1.2. Distribución Condicional de Costos	36
2.1.3. Criterio de Optimización	38
2.2. Optimización a Tiempo Discreto con Espacio Nu- merable de Estados	40
2.2.1. Método de Aproximaciones Sucesivas	42
2.3. Optimización del Caso a Tiempo Discreto con Es- pacio Finito de Estados	46
2.4. Optimización del Caso a Tiempo Continuo	63
2.4.1. Método de Uniformización	67
3. Control de las Tasas de Servicio en una Línea de Espera	75

ÍNDICE GENERAL

2

3.1. Ejemplo 76

Bibliografía 103

Universidad Juárez Autónoma de Tabasco.
México.

Introducción

En este trabajo se resuelve un problema de control de líneas de espera, con el criterio de costo promedio y con espacio numerable de estados a tiempo continuo; para lograr este objetivo se propone transformar dicho problema de control a un problema de control con espacio finito de estados a tiempo discreto, el cual se resolverá utilizando programación dinámica.

Una línea de espera o cola se forma en sistemas que ofrecen servicios con cierta capacidad de atención, al llegar los clientes si el servidor no está disponible, y el cliente decide esperar, se forma dicha cola.

La teoría de líneas de espera proporciona modelos para problemas interesantes, estos modelos se utilizan para la toma de decisiones que influyen en los siguientes aspectos: tiempo de espera total del sistema o en la cola, cantidad de clientes esperando en el sistema o en las colas y tiempo ocioso de los servidores. Así, el interés no sólo es modelar la cola sino controlarla. Para el control de colas se cuenta con la teoría de procesos de control estocástico, cuyo interés más importante es encontrar la función de valores óptimos y la correspondiente política óptima. Este tipo de problemas se estudian en el contexto de la teoría de Procesos de Decisión de Markov (PDMs), donde los PDMs modelan un

problema de decisión secuencial en el cual un sistema evoluciona en el tiempo influido por un controlador o decisor. Los PDMs consisten de un espacio de estados, de un espacio de acciones, de una familia de conjuntos de acciones admisibles, de una ley de transición y de una función de costo, junto con un índice de funcionamiento (también llamado criterio de optimalidad) mediante el cual se mide la respuesta del sistema a los controles aplicados. Este índice por lo general representa costos (o ganancias), así, el objetivo de un problema de decisión de Markov o problema de control, es encontrar políticas que minimicen el costo esperado (o maximicen la ganancia esperada) en el tiempo.

La metodología para resolver el problema de interés es programación dinámica. Los métodos más utilizados en programación dinámica para PDMs son:

- (a) iteración de valores,
- (b) iteración de políticas.

Para este trabajo se utilizará el método de iteración de valores.

Para resolver el problema de control en PDMs a tiempo continuo se establece una equivalencia entre una Cadena de Decisión de Markov a Tiempo Discreto (CDMTD) y una Cadena de Decisión de Markov a Tiempo Continuo (CDMTC) a través del método de uniformización, el cual fue propuesto en [15] para resolver problemas de PDMs a tiempo continuo. La uniformización consiste en introducir una probabilidad, distinta de cero, de ir de un estado a el mismo; ya que en una CDMTC esta probabilidad es cero, así el problema se puede observar en periodos determinados de tiempo, es decir, a tiempo discreto.

El problema de control de PDMs con espacio numerable de estados y a tiempo discreto se resuelve con el método de aproximaciones sucesivas, es decir, por una sucesión de CDMTD, pero cada una con espacio finito de estados.

Se utiliza el método de iteración de valores para calcular la función de valor y las políticas óptimas, cuando el espacio de estados es finito y a tiempo discreto.

La tesis consiste de tres capítulos. El primer capítulo proporciona la información necesaria sobre teoría de colas o líneas de espera. Además, menciona los modelos de colas simples, describe un sistema de espera y concluye aportando la teoría mínima de cadenas y procesos de Markov. El capítulo dos, llamado programación dinámica y colas, está conformado de la siguiente manera: la sección 2.1 proporciona la definición de un modelo de control estocástico con espacio numerable de estados a tiempo discreto, mencionando cada uno de los elementos que lo conforman, la sección 2.2 presenta el problema de control estocástico con costo promedio, con espacio numerable de estados y a tiempo discreto, tal problema de control se resuelve por el método de aproximaciones sucesivas. En la sección 2.3 se considera un problema de control estocástico con espacio finito de estados y para resolverlo se utiliza iteración de valores, en la sección 2.4 se establece el resultado para resolver un problema de control estocástico con costo promedio a tiempo continuo utilizando uniformización. El capítulo tres presenta un ejemplo de la teoría de procesos de control estocástico con costo promedio en un sistema de líneas de espera con un servidor. Finalmente se proporciona la conclusión y las referencias.

⁶
Universidad Juárez Autónoma de Tabasco.
México.

Capítulo 1

Líneas de Espera

En este capítulo se abordará la teoría básica para establecer el problema de interés. En la primera sección se describe la teoría de colas. En la segunda sección se dará el criterio de rendimiento. En la tercera sección se clasifican los distintos sistemas de líneas de espera. La cuarta sección describe los elementos de una línea de espera. Finalmente se proporciona la teoría mínima de procesos estocásticos para modelar una cola.

1.1. Introducción a Líneas de Espera

La espera en situaciones importantes de nuestra vida no es nada agradable. Se tiene que hacer cola casi todos los días para recibir ciertos servicios. Esto suele ocurrir cuando la demanda real de un servicio es superior a la capacidad que existe para dar dicho servicio (ver [6]).

Las líneas de espera son parte de la vida diaria. Ejemplos reales de esta situación son: hacer un depósito en el banco, comprar un boleto para el cine, pagar en el supermercado, enviar un paquete por correo, aparatos que deben ser reparados por un ser-

vicio técnico, etc. Todavía más frecuentes son las situaciones de espera en el contexto de la informática, las telecomunicaciones y en general las nuevas tecnologías. Así, por ejemplo, los procesos enviados a un servidor para ejecución se forman en la cola mientras no son atendidos; la información solicitada a través de internet puede recibirse con demora debido a congestión en la red, la señal de líneas ocupadas si la central de la que depende un teléfono móvil está colapsada en ese momento, etc.

La teoría de colas es el estudio matemático del comportamiento de líneas de espera, a partir de los datos que suministra el estudio de la teoría de colas se puede obtener la información necesaria, por ejemplo, para definir el número de asientos necesarios en una sala de espera, o la estructura de etapas de un proceso de atención al cliente. En cualquier caso, para poder tomar decisiones hacen falta datos que la teoría de colas puede facilitar en algunos de los siguientes tres aspectos:

1. Tiempo de espera (en el total del sistema o en la cola).
2. Cantidad de clientes esperando (en el sistema o en las colas).
3. Tiempo ocioso de los servidores (total o particular de cada servicio).

El estudio de las líneas de espera es importante, porque proporciona tanto una base teórica del tipo de servicio que se puede esperar de un determinado sistema, como la forma en la cual dicho sistema puede ser diseñado para proporcionar un determinado grado de servicio a sus clientes. Cuando los clientes llegan a un lugar demandando un servicio a un servidor (el cual tiene una

cierta capacidad de atención) y si el servidor no está disponible y el cliente decide esperar entonces se forma la línea de espera.

El problema en un sistema de espera es determinar que capacidad o tasa de servicio proporciona el balance correcto. Esto no es sencillo, ya que un cliente no llega en un horario fijo, es decir, no se sabe con exactitud en que momento llegarán los clientes, asimismo el tiempo de servicio no está determinado a priori.

La teoría de colas en si no resuelve directamente el problema, pero contribuye con la información fundamental que se requiere para tomar las decisiones concernientes, prediciendo algunas características sobre la línea de espera tales como la probabilidad de que se formen, el tiempo de espera promedio, etc.

En la teoría de colas generalmente se llama sistema a un grupo de unidades físicas, integradas de tal modo que pueden operar de acuerdo con una serie de operaciones determinadas. La teoría de colas busca una solución al problema de la espera prediciendo primero el comportamiento del sistema. Una solución al problema de la espera no sólo consiste en minimizar el tiempo que los clientes pasan en el sistema, sino también en minimizar los costos totales de aquellos que solicitan el servicio y de quienes lo prestan, este tipo de problema lo resuelve la teoría de control estocástico.

Los objetivos de controlar colas consisten en:

- Identificar el nivel óptimo de capacidad del sistema que minimiza el costo de funcionamiento.

- Evaluar el impacto que las posibles alternativas de modificación de la capacidad del sistema tendrían en el costo de funcionamiento.
- Establecer un balance equilibrado (óptimo) entre las consideraciones cuantitativas de costos y las cualitativas de servicio.
- ² Prestar atención al tiempo de permanencia en el sistema o en la cola; la paciencia de los clientes depende del tipo de servicio específico considerado y eso puede propiciar que un cliente abandone el sistema.

Proporcionar demasiada capacidad de servicio para operar el sistema implica costos excesivos; pero al no contar con suficiente capacidad de servicio la espera aumenta con todas sus desafortunadas consecuencias (ver [13]).

La teoría de colas es una colección de modelos matemáticos que describen sistemas de líneas de espera particulares o sistemas de colas. Los sistemas de colas son modelos de sistemas que proporcionan servicio. Los modelos se utilizan para encontrar un buen funcionamiento entre costos del sistema y los tiempos promedio de la línea de espera para un sistema dado; como ⁶ modelo pueden representar cualquier sistema, donde los clientes llegan buscando un servicio de algún tipo y salen después de que dicho servicio haya sido atendido, además, se pueden modelar los sistemas de este tipo tanto como líneas de espera sencillas o como un sistema de colas interconectadas formando una red de colas.

Las fórmulas para cada modelo indican cuál debe ser el desempeño del sistema correspondiente y señalan la cantidad promedio de espera que ocurrirá en una serie de circunstancias. Por lo tanto, estos modelos de líneas de espera son muy útiles para determinar como operar un sistema de colas de la manera más efectiva.

1.2. Rendimiento de un Sistema

Existen medidas de rendimiento que se utilizan para evaluar un sistema de colas. Para diseñar y poner en operación un sistema de colas, por lo general los administradores se preocupan por el nivel de servicio que recibe un cliente, así como el uso apropiado de las instalaciones de servicio de la empresa.

La tarea de un analista de colas es de dos tipos:

- Establecer mecanismos para medir la efectividad del sistema.
- Diseñar un sistema óptimo (de acuerdo a algún criterio).

El diseño eficiente ⁹ consiste básicamente en definir un sistema cuyos costos (de diseño y de operación) se justifique por el servicio que proporciona. Dicho servicio se puede evaluar mediante el costo de no proporcionarlo. De este modo el fin es minimizar los costos totales.

El cálculo ² de las medidas de rendimiento depende de los procesos de llegadas y del servicio del sistema de colas. Estos procesos son descritos matemáticamente mediante distribuciones de

llegada y de servicio.

1.3. Modelos de Colas

Los modelos de colas más frecuentes son los siguientes:

1. Un servidor y una línea de espera.
 2. Servidores en serie.
 3. Servidores en paralelo.
 4. Un solo servidor y múltiples líneas de espera.
 5. Múltiples servidores y una línea de espera.
 6. Red de servidores.
1. *Un servidor y una línea de espera.* Aquí, el cliente que entra es atendido por un servidor. Si el servidor está ocupado el cliente se forma en la cola (única), esperando que el servidor se desocupe y pueda ser atendido. Después de esperar cierto tiempo en la cola, el cliente entra a la unidad de servicio para ser atendido; al completar el servicio el cliente sale del sistema.
 2. *Servidores en series.* En este sistema se tienen varios servidores en serie. Aquí, el cliente entra en el sistema a la cola del servidor 1, recibe el servicio y después pasa a la cola del servidor 2, recibe servicio, y así sucesivamente hasta que complete todos los servicios y finalmente abandone el sistema.

3. *Servidores en paralelo.* Un sistema con servidores en paralelo se caracteriza porque hay más de un servidor que ejecuta la misma función con la misma eficiencia, es decir, hay un número de servidores con sus respectivas colas. El cliente que llega es enviado a una de las colas, una vez que le asignaron a que cola ir debe permanecer en ella hasta que reciba el servicio y abandone el sistema.
4. *Un solo servidor y múltiples líneas de espera.* Aquí, solamente un servidor es responsable para servir las múltiples colas. El servidor puede considerar el mecanismo de decidir cual cola servir en una época de decisión.
5. *Múltiples servidores y una línea de espera.* En este sistema las razones de servicios son fijas, pero pueden variar de servidor a servidor. Como máximo un cliente puede recibir servicio de cualquier servidor en cualquier tiempo. Si hay un cliente esperando servicio y por lo menos un servidor está desocupado, entonces las opciones de control consisten en enviar al cliente a un servidor libre.
6. *Red de servidores.* Aquí, un número de servidores proporcionan servicios diferentes a clientes y cada servidor tiene su propia cola. Los clientes que llegan se dirigen a cada cola. Cuando un cliente termina el servicio en un servidor puede dirigirse a otro servidor o salir del sistema.

1.4. Descripción de un Sistema de Espera

La descripción completa de tales sistemas requiere de técnicas matemáticas, esencialmente probabilísticas.

Un sistema de espera consiste de una unidad de servicio, un proceso de arribo de clientes que deben ser atendidos por la unidad y un proceso de servicio.

En un sistema de espera se tiene una **fente** o **población** de clientes que entrarán al sistema a recibir servicio. La **unidad de servicio** puede tener uno o más **servidores**. Si todos los servidores están ocupados cuando un cliente llega al sistema, entonces debe unirse a la **cola** y esperar hasta que un servidor esté desocupado y lo pueda atender. Después de esperar un cierto tiempo en la cola, el cliente entra a la **unidad de servicio** y es atendido por un servidor; al completar el servicio, el cliente sale del sistema (ver [6]).

Definición 1.4.1. Sea t_n el tiempo aleatorio en el que llega el n -ésimo cliente, $n = 1, 2, \dots$; $t_0 \equiv 0$. Suponiendo que $0 = t_0 < t_1 < t_2 < \dots$, los tiempos entre llegadas se definen como

$$\tau_n = t_n - t_{n-1}, \quad n = 1, 2, \dots \quad (1.1)$$

Una hipótesis importante que se tendrá en lo que sigue es que los tiempos entre arribos $\tau_1, \tau_2, \dots$, son variables aleatorias (v.'s a.'s) independientes e idénticamente distribuidas (i.i.d). Suponga que los tiempos entre llegadas tienen un valor medio $E(\tau)$ finito y positivo, es decir, $0 < E(\tau) < \infty$, donde τ denota un tiempo entre llegadas arbitrario. Si

$$E(\tau) = \frac{1}{\lambda},$$

el parámetro λ se puede interpretar como la intensidad promedio de llegadas al sistema por unidad de tiempo.

Suponga que los servidores son idénticos, denotando por s la v.a. que describe el tiempo de servicio, cuyo valor medio $E(s)$ es finito y positivo, es decir, $0 < E(s) < \infty$. Si

$$E(s) = \frac{1}{\mu},$$

se puede interpretar a μ como la intensidad promedio de servicio por servidor.

Para distinguir entre el número de clientes en el sistema y el número de clientes en la cola se puede ver que,

$$N(t) = N_q(t) + N_s(t), \quad t \geq 0, \quad (1.2)$$

donde

$N(t)$ = Número de clientes en el sistema en el tiempo t ,

$N_q(t)$ = Número de clientes en la cola en el tiempo t , y

$N_s(t)$ = Número de clientes recibiendo servicio en el tiempo t .

Como indicadores del funcionamiento de un sistema de espera se pueden mencionar dos parámetros importantes, uno de ellos es la **intensidad de tráfico** (I):

$$\begin{aligned} I &= \frac{E(s)}{E(\tau)} \\ &= \lambda E(s) \\ &= \frac{\lambda}{\mu}, \end{aligned}$$

el valor de I determina (en un sentido probabilístico) si la cola crecerá sin límite o si, eventualmente, alcanzará un estado estacionario; el otro es la **ocupación por servidor** (u):

$$\begin{aligned}
 u &= \frac{I}{\varrho} \\
 &= \frac{\lambda}{\varrho\mu},
 \end{aligned}$$

donde ϱ es el número de servidores. Este puede interpretarse como la probabilidad de que un servidor particular esté ocupado o como el número promedio de personas que van a ser atendidos por cada servidor.

Resumiendo, las características para describir un sistema de espera son:

- **Fuente de entrada.** Es el conjunto de individuos (no necesariamente seres humanos) que pueden llegar a solicitar el servicio. La fuente puede ser finita o infinita. Siempre que la fuente es grande conviene suponer que es infinita, ya que permite resolver de forma más sencilla muchas situaciones del desarrollo matemático. Dicha suposición de infinitud no resulta restrictiva, ya que el número de individuos que están solicitando el citado servicio, no afecta a la frecuencia con la que la población genera nuevas peticiones de servicio.
- **Mecanismo de servicio.** Es el procedimiento por el cual se proporciona servicio a los clientes que lo solicitan. Para determinar totalmente el mecanismo de servicio de una línea de espera se debe conocer el número de servidores (ϱ) en el sistema y la distribución de probabilidad del tiempo que cada servidor tarda en satisfacer un servicio. Si los servidores tienen distinta destreza para dar el servicio, se debe especificar la distribución del tiempo de servicio para cada uno.

- **Cliente.** Es todo individuo de la fuente de entrada que solicita servicio. Dependiendo del sistema que se trate, los clientes pueden ser personas, máquinas, paquetes, etc.
- **La cola (q).** Es el conjunto de clientes que esperan servicio, es decir, los clientes que ya han solicitado el servicio pero que aún no han pasado al mecanismo de servicio.
- **Proceso de arribos o llegadas.** Esto se refiere a la frecuencia media de llegada de los clientes. En situaciones de colas más usuales la llegada es estocástica, es decir, la llegada depende de una cierta variable aleatoria, en este caso es necesario conocer la distribución probabilística entre dos llegadas sucesivas. Además, habría que tener en cuenta si los clientes arriban de manera independiente o en grupo. En este segundo caso habría que definir la distribución probabilística de éstos.

También, es posible que los clientes sean impacientes, es decir, que lleguen a la cola y si es demasiado larga abandonen el sistema, o que después de esperar cierto tiempo en la cola decidan abandonarla. Por último, es posible que el patrón de llegada varíe con el tiempo. Si se mantiene constante se le llama estacionario, y si varía es no-estacionario. Este proceso está determinado por la función de distribución $F(t) = P\{\tau \leq t\}$ de los tiempos entre llegadas definido en (1.1).

- **Distribución del tiempo de servicio.** Conociendo las distribuciones del tiempo de servicio se pueden contestar las siguientes preguntas: ¿cuándo está disponible el servicio?, ¿cuántos clientes pueden atenderse a la vez?, ¿se pueden

atender en grupos o de modo individual? y ¿cuánto tiempo toma el proporcionar el servicio? El tiempo de servicio varía con el número de clientes en la cola. El tiempo de servicio está determinado por la función de distribución de probabilidad:

$$S(t) = P\{s \leq t\}.$$

- **Número de servidores** (ϱ). En los sistemas de espera más simples se tiene sólo un servidor, $\varrho = 1$, pero en muchas situaciones prácticas se tiene un número $\varrho > 1$ de servidores. También, se pueden considerar sistemas con un número infinito de servidores, esto nos indica que en el sistema no se forman colas porque siempre se disponen de servidores libres para atender a los nuevos clientes que llegan.
- **Capacidad máxima del sistema** (K). Es el máximo número de clientes que puede contener el sistema y en el caso de que haya ϱ servidores se tiene que

$$K = C_q + \varrho,$$

en donde C_q es la capacidad de la cola, es decir, el número máximo de clientes permitidos en la cola.

- **Disciplina de servicio**. La disciplina de servicio es la manera en que los clientes se ordenan en el momento de ser servidos. Cuando se piensa en colas se admite que la disciplina de cola normal es FIFO (atender primero a quien llegó primero). Sin embargo, en muchas colas es natural usar la disciplina LIFO (atender primero al último). También es posible encontrar reglas de secuencia con prioridades, por ejemplo, secuenciar primero las tareas con menor duración o según tipos de clientes.

La notación de Kendall Lee es una forma abreviada de describir un sistema de espera con seis de los componentes antes mencionados. En esta notación se escribirá

$$A/S/\varrho/K/F/d,$$

en donde cada una de estas letras describe:

A = distribución de los tiempos entre arribos,

S = distribución del tiempo de servicio,

ϱ = número de servidores,

K = capacidad del sistema,

F = número de individuos en la fuente,

d = disciplina de servicio.

1.5. Cadenas de Markov

La teoría de esta sección está basada en [2], [6], [10] y [12].

Definición 1.5.1. ¹ *Un proceso estocástico es una familia $X = \{X(t), t \in T\}$ de variables aleatorias. Es decir, para cada $t \in T$, en donde T es el conjunto de parámetros o conjunto de índices, $X(t)$ es una variable aleatoria (v.a.) definida sobre algún espacio de probabilidad (Ω, F, P) ,*

$$X(t) : \Omega \rightarrow \mathbb{R}.$$

Para enfatizar que el proceso depende de los eventos elementales $w \in \Omega$, se escribe

$$X(t) = X(t, w).$$

Así, el proceso se puede ver como una colección de funciones reales definidas sobre $T \times \Omega$, es decir,

$$X(\cdot, \cdot) : T \times \Omega \rightarrow \mathbb{R},$$

tal que para cada $t \in T$ fijo y para cada $w \in \Omega$ fijo se tiene

1. $X(t, \cdot) : \Omega \rightarrow \mathbb{R}$ es una v.a.,
2. $X(\cdot, w) : T \rightarrow \mathbb{R}$ es una función de $t \in T$, llamada una realización del proceso X .

El **espacio de estados** del proceso es el conjunto X de todos los valores posibles que pueden tomar las variables $X(t)$, para cada $t \in T$.

Cada punto en X se llama **estado** del proceso. Si X es finito o numerable se dice que el proceso es discreto; si X es un intervalo se dice que el proceso es continuo.

El proceso también se clasifica de acuerdo a la naturaleza del conjunto de parámetro T . Si T es numerable, por ejemplo, $T = \{0, 1, 2, \dots\}$, $T = \{0, \pm 1, \pm 2, \dots\}$, etc., se dice que X es un proceso con parámetro (o en tiempo) discreto. Si T es un intervalo, es decir, $T = [a, b]$, $T = [0, \infty]$, $T = (-\infty, \infty)$, etc., X es un proceso con parámetro (o en tiempo) continuo.

Si $X = \{X(t), t \in T\}$, es un proceso a tiempo discreto, es decir, $T = \{0, 1, \dots\}$, entonces denotando $X(t) = x_t$, se tiene que X es una sucesión de v.'s a.'s $X = \{x_0, x_1, \dots\}$.

EJEMPLO 1.5.2. a) Para cada $t \geq 0$, sea $N(t)$ el número de clientes en un cierto sistema de espera. Entonces $\{N(t), t \geq 0\}$,

es un proceso en tiempo continuo, $T = [0, \infty)$, con espacio de estados discretos $X = \{0, 1, 2, \dots\}$.

b) Una sucesión de v.'s a.'s $\{x_1, x_2, \dots\}$, es un proceso a tiempo discreto. El espacio de estados puede ser continuo o discreto, dependiendo de si las x_i son v.'s a.'s continuas o discretas.

Mediante un proceso estocástico se pretende modelar la evolución aleatoria a lo largo del tiempo de una variable o sistema de interés. Por ejemplo:

- a) X_t = Número de clientes en espera de ser atendidos al tiempo t .
- b) X_n = Estado operacional (1) o no operacional (0) de una máquina al inicio del n -ésimo día.
- c) X_t = Valor de la divisa al tiempo t .
- d) X_t = Magnitud del movimiento sísmológico de un cierto lugar al tiempo t .
- e) X_t = Número de artículos en un almacén al tiempo t .

Los distintos tipos de procesos se obtienen al considerar distintos espacios parametrales T , distintos espacios de estados (posibles valores de las variables aleatorias), y sobre todo al establecer la dependencia o independencia entre las distintas variables aleatorias que constituyen el proceso estocástico.

Algunos tipos de procesos

- a) Procesos con incrementos independientes. Si $t_1 < t_2 < \dots < t_n$, entonces las variables $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ son independientes.

- b) Procesos estacionarios u homogéneos. Para $t_1 < t_2 < \dots < t_n$ y para cualquier $h > 0$, la distribución de probabilidad del vector $(X_{t_1}, \dots, X_{t_n})$ es idéntica a la del vector $(X_{t_1+h}, \dots, X_{t_n+h})$, es decir,

$$P(X_{t_{n+1}} \leq x_{n+1} | X_{t_n} = x_n) = P(X_{t_{n+1}-t_n} \leq x_{n+1} | X_0 = x_n).$$

- c) Procesos de Markov. ¹ Un proceso estocástico $X = \{X_t, t \in T\}$, con $T \subset \mathbb{R}$, es un proceso de Markov si para cualquier $x \in \mathbb{R}$ y $s, t \in T$, con $s \leq t$, se satisface que $P(X_t \leq x | X_r, r \leq s) = P(X_t \leq x | X_s)$. Esta ecuación se conoce como la propiedad de Markov, debido a esta propiedad se dice también que X es una v.a. sin memoria.

El proceso estocástico utilizado en la modelación de una línea de espera tiene la propiedad de Markov, ya que la probabilidad condicional de llegar a un estado futuro depende exclusivamente del estado actual en el que se encuentre el sistema, sin importar el estado anterior de dicho sistema.

Definición 1.5.3. Una cadena de Markov es un proceso de Markov $\{X_t : t = 0, 1, \dots\}$ a tiempo discreto y con espacio de estados finito o numerable, típicamente $\{0, 1, \dots, n\}$ o bien $\mathbb{N} \cup \{0\}$. La propiedad de Markov se escribe

$$P(X_t = x_t | X_0 = x_0, \dots, X_{t-1} = x_{t-1}) = P(X_t = x_t | X_{t-1} = x_{t-1}). \quad (1.3)$$

El conjunto de probabilidades condicionales se conoce como *probabilidades de transición*, para cualesquiera $x, y \in X$ se denota la probabilidad de transición $p_{xy}^{t,t+1} = P(x_{t+1} = y | x_t = x)$. Si se supone estacionariedad esta probabilidad no depende de t ,

es decir, $p_{xy}^{t,t+1} = p_{xy}$, donde $p_{xy} = P(X_1 = y | X_0 = x)$.

Haciendo variar los estados x y y , se obtiene la matriz de probabilidad de transición en un paso:

$$\mathbb{P} = (p_{xy}) = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & \cdots \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \end{matrix} & \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{matrix}.$$

También se pueden definir probabilidades de transición en n pasos, o sea, $p_{xy}^n = P(X_n = y | X_0 = x)$. Haciendo variar x y y , se obtiene la matriz de probabilidad de transición en n pasos:

$$\mathbb{P}^n = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & \cdots \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \end{matrix} & \begin{pmatrix} p_{00}^n & p_{01}^n & p_{02}^n & \cdots \\ p_{10}^n & p_{11}^n & p_{12}^n & \cdots \\ p_{20}^n & p_{21}^n & p_{22}^n & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{matrix}.$$

La matriz de probabilidad de transición $\mathbb{P} = (p_{xy})$ satisface que

1. $p_{xy} \geq 0$.
2. $\sum_{y=0}^{\infty} p_{xy} = 1$.

En el caso donde el espacio de estados es finito $X = \{0, 1, 2, \dots, n\}$, haciendo variar los estados x y y , se obtiene la matriz de probabilidad de transición $\mathbb{P}$ con dimensión en $(n+1) \times (n+1)$

en un paso, esto es,

$$\mathbb{P} = (p_{xy}) = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & \cdots & n \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ n \end{matrix} & \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots & p_{0n} \\ p_{10} & p_{11} & p_{12} & \cdots & p_{1n} \\ p_{20} & p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n0} & p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \end{matrix}.$$

Se pueden definir también probabilidades de transición en n pasos, es decir, $p_{xy}^n = P(X_n = y | X_0 = x)$. Haciendo variar x y y , se obtiene la matriz de probabilidad de transición en n pasos:

$$\mathbb{P}^n = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & \cdots & n \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ n \end{matrix} & \begin{pmatrix} p_{00}^n & p_{01}^n & p_{02}^n & \cdots & p_{0n}^n \\ p_{10}^n & p_{11}^n & p_{12}^n & \cdots & p_{1n}^n \\ p_{20}^n & p_{21}^n & p_{22}^n & \cdots & p_{2n}^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n0}^n & p_{n1}^n & p_{n2}^n & \cdots & p_{nn}^n \end{pmatrix} \end{matrix}.$$

La matriz de probabilidad de transición $\mathbb{P} = (p_{xy})$ satisface que

1. $p_{xy} \geq 0$.
2. $\sum_{y=0}^n p_{xy} = 1$.

Toda matriz cuadrada que cumpla estas condiciones se llama matriz estocástica. Si también la suma por columnas es uno, la matriz se llama doblemente estocástica.

Dada una cadena de Markov esta se determina por sus probabilidades de transición $\mathbb{P} = (p_{xy})$ y una distribución de probabilidad inicial $P(X_0 = x)$, con $x = 0, 1, 2, \dots$.

Ecuación de Chapman-Kolmogorov

Sea $\mathbb{P} = (p_{xy})$ la matriz de probabilidad de transición de una cadena de Markov. Para cualesquiera tiempos n y m , las potencias de $\mathbb{P}$ se definen por

$$\mathbb{P}^{n+m} = \mathbb{P}^n \mathbb{P}^m.$$

Es decir, las probabilidades de transición son simplemente las entradas de la $n + m$ -ésima potencia de $\mathbb{P}$, escribiendo en términos de elementos está dada por

$$p_{xy}^{n+m} = \sum_z p_{xz}^n p_{zy}^m. \quad (1.4)$$

En particular, para

$$\mathbb{P}^2 = \mathbb{P}\mathbb{P},$$

con el (x, y) -ésimo elemento dado por

$$p_{xy}^2 = \sum_k p_{xk} p_{ky},$$

para

$$\mathbb{P}^3 = \mathbb{P}\mathbb{P}^2,$$

con el (x, y) -ésimo elemento dado por

$$p_{xy}^3 = \sum_{k,l} p_{xk} p_{kl} p_{ly}.$$

Estados Comunicantes

Un estado y es accesible desde el estado x si existe $n \geq 0$ tal que $p_{xy}^n > 0$ y se escribe $x \rightarrow y$. Los estados x y y están comunicados si $x \rightarrow y$ y $y \rightarrow x$. Se define

$$p_{xy}^0 = \delta_{xy} = \begin{cases} 1 & \text{si } x = y, \\ 0 & \text{si } x \neq y. \end{cases}$$

La comunicación es una relación de equivalencia. Ello induce una partición del espacio de estados en clases de comunicación. En particular una cadena de Markov es irreducible si todos los estados se comunican entre sí.

Estados recurrentes y transitorios.

Un estado x es recurrente si se cumple $P(X_n = x \text{ para algún } n \geq 1 | X_0 = x) = 1$. Un estado que no es recurrente es llamado transitorio. Estos conceptos pueden también definirse usando la siguiente notación:

Para cualesquiera estados x y y , y para cualquier $n \geq 1$, se define la probabilidad de visitar el estado y por primera vez en exactamente n pasos, partiendo de x como:

$$\hat{p}_{xy}^n = P(X_n = y, X_k \neq y \text{ para } k = 1, \dots, n-1 | X_0 = x).$$

También, se define la probabilidad de una eventual visita al estado y partiendo de x como:

$$\hat{p}_{xy} = \sum_{n=0}^{\infty} \hat{p}_{xy}^n.$$

Se tiene que $\hat{p}_{xy}^0 \equiv 0$. Los conceptos de recurrencia y transitoriedad pueden establecerse como sigue:

- a) x es recurrente $\Leftrightarrow \hat{p}_{xx} = 1$.
 b) x es transitorio $\Leftrightarrow \hat{p}_{xx} = 0$.

Existe además, el siguiente criterio para determinar la recurrencia y la transitoriedad.

PROPOSICIÓN 1.5.4.

- a) x es recurrente $\Leftrightarrow \sum_{n=0}^{\infty} p_{xx}^n = \infty$.
 b) x es transitorio $\Leftrightarrow \sum_{n=0}^{\infty} p_{xx}^n < \infty$.

La recurrencia y la transitoriedad son propiedades de clase, es decir,

- a) x es recurrente, $x \longleftrightarrow y \Rightarrow y$ es recurrente.
 b) x es transitorio, $x \longleftrightarrow y \Rightarrow y$ es transitorio.

Procesos de Markov a Tiempo Continuo.

La definición de las Cadenas de Markov a Tiempo Continuo (CMTC) hace referencia a la definición de los procesos generales de Markov pero ahora el parámetro es continuo y el espacio de estados numerable (ver [2], [5] y [12]).

Definición 1.5.5. Dado un proceso estocástico $\{X_t : t \in T\}$, define una CMTC si para $t_i \in \mathbf{R}_0^+$, $i = 0, \dots, n+1$, con $0 = t_0 < t_1 < \dots < t_n < t_{n+1}$, $\forall n \in \mathbb{N}$, y $\forall x_i \in S$, la probabilidad condicional cumple la siguiente relación:

$$\begin{aligned} & P(X_{t_{n+1}} = x_{n+1} | X_{t_n} = x_n, X_{t_{n-1}} = x_{n-1}, \dots, X_{t_0} = x_0) \\ &= P(X_{t_{n+1}} = x_{n+1} | X_{t_n} = x_n). \end{aligned} \quad (1.5)$$

La ecuación (1.5) expresa la propiedad de Markov para cadenas de Markov a tiempo continuo, como en la ecuación (1.3) para el caso de Cadenas de Markov a Tiempo Discreto (CMTD). Si además se supone la propiedad de homogeneidad, entonces debido a la distribución exponencial, la cual es la única distribución a tiempo continuo que ofrece la propiedad de pérdida de memoria, los tiempos de permanencia en los estados de una CMTC están distribuidos exponencialmente.

El lado derecho de la ecuación (1.5) se conoce como la probabilidad transición $p_{ij}(u, v)$ de la CMTC de ir del estado i al estado j durante el período de tiempo $[u, v)$, con $u, v \in T$ y $u < v$:

$$p_{ij}(u, v) = P(X_v = j | X_u = i). \quad (1.6)$$

Para $u = v$ se define

$$p_{ij}(u, v) = \begin{cases} 1, & i = j, \\ 0 & \text{otro caso.} \end{cases} \quad (1.7)$$

Si las probabilidades de transición $p_{ij}(u, v)$ sólo dependen de la diferencia de tiempo $t = v - u$, y no sobre los valores reales de u y v , las probabilidades de transición simplificadas para la CMTC de tiempo homogéneo resultan:

$$\begin{aligned} p_{ij}(t) &= p_{ij}(0, t) \\ &= P(X_{u+t} = j | X_u = i) \\ &= P(X_t = j | X_0 = i), \end{aligned} \quad (1.8)$$

para toda $u \in T$. Dadas las probabilidades de transición $p_{ij}(u, v)$ de la CMTC al tiempo u y $\gamma_i(u)$ la probabilidad de estar en el estado i al tiempo u , la probabilidad del estado no condicionadas

$\gamma_j(v), j \in S$ del proceso al tiempo v se puede obtener de la siguiente manera:

$$\gamma_j(v) = \sum_{i \in S} p_{ij}(u, v) \gamma_i(u), \quad (1.9)$$

para cualesquiera $u, v \in T$ ($u < v$). Con $\mathbf{P}(u, v) = [p_{ij}(u, v)]$ como la matriz de las probabilidades de transición, para cualquier par de estados i, j y cualquier intervalo de tiempo $[u, v], u, v \in T$, desde el dominio de parámetros y el vector $\gamma(u) = (\gamma_0(u), \gamma_1(u), \gamma_2(u), \dots)$ de las probabilidades de estado en cualquier instante de tiempo u , la ecuación (1.9) se puede dar en forma matricial:

$$\gamma(v) = \gamma(u) \mathbf{P}(u, v), \quad (1.10)$$

para cualesquiera $u, v \in T$ ($u < v$). Nótese que para todo $u \in T, \mathbf{P}(u, u) = \mathbf{I}$ es la matriz identidad.

En el caso de tiempo homogéneos, la ecuación (1.9) se reduce a:

$$\gamma_j(t) = \sum_{i \in S} p_{ij}(t) \gamma_i(0) = \sum_{i \in S} p_{ij}(0, t) \gamma_i(0), \quad (1.11)$$

o en notación matricial:

$$\gamma(t) = \gamma(0) \mathbf{P}(t) = \gamma(0) \mathbf{P}(0, t). \quad (1.12)$$

Como en el caso a tiempo discreto (ver (1.4)), la ecuación de Chapman-Kolmogorov (1.13) para las probabilidades de transición de una CMTC puede derivarse de la ecuación (1.5) aplicando el teorema de probabilidad total:

$$p_{ij}(u, v) = \sum_{k \in S} p_{ik}(u, w) p_{kj}(w, v), \quad (1.13)$$

$$0 < u < w < v.$$

Estos conceptos básicos se emplearán en el desarrollo de la teoría de procesos de decisión de Markov en el siguiente capítulo.

Capítulo 2

Programación Dinámica y Colas

Este capítulo trata con Procesos de Control Estocástico, también conocidos como Procesos de Decisión de Markov (PDMs) o Procesos de Control de Markov (PCMs), con horizonte infinito, tiempo discreto, tiempo continuo y con costo promedio. Para los procesos de control considerados se supondrá que los espacios de estados y acciones son espacios discretos y que la dinámica está determinada por probabilidades de transición. El problema principal en procesos de control óptimo es determinar la función de valor óptimo y la política óptima de control que minimiza el criterio de rendimiento, el cual evalúa la dinámica del sistema.

En este capítulo se muestran los métodos que son utilizados para resolver problemas de Procesos de Decisión de Markov a tiempo continuo con espacio numerable de estados. Primeramente el problema de PDMs a tiempo continuo se transforma a un problema a tiempo discreto utilizando el método de uniformización. El problema a tiempo discreto con espacio numerable de estados se resuelve con el método de aproximaciones

sucesivas el cual obtiene una sucesión de Cadenas Decisión de Markov (CDMs) con espacio finito de estados. Y por último se hace uso de programación dinámica para resolver el problema a tiempo discreto con espacio finito de estados.

Este capítulo se organiza de la siguiente forma: en la primera sección se hace una revisión de los Modelos de Control Estocástico a tiempo discreto y continuo. En la segunda sección se presenta la solución al problema de optimalidad a tiempo discreto con espacio numerable de estados, posteriormente se proporciona la solución en el caso con espacio finito de estados. Finalmente se presentan los resultados principales para este trabajo.

2.1. Cadenas de Decisión de Markov a Tiempo Discreto

Esta sección se basa en [3]. Si en un Proceso de Decisión de Markov (PDM) el espacio de estados es discreto el problema de control se conoce como Cadena de Decisión de Markov (CDM), esta estructura matemática permite analizar sistemas de control en tiempo discreto, involucrando aspectos aleatorios. El tiempo se particiona en longitudes iguales, llamadas periodos. A continuación, se proporciona la definición de Cadena de Decisión de Markov.

Definición 2.1.1. *Una cadena de decisión de Markov Δ consiste de un espacio de estados, un espacio de acciones, probabilidades de transición y una función de costo. Se denota por*

$$(S, A, \{A(i)|i \in S\}, P, C).$$

1. El espacio de estados es el conjunto de todos los estados del sistema, el cual se denota por S y puede ser infinito numerable o finito. El estado del sistema es la información relevante necesaria para describir las condiciones actuales del sistema.
2. Cuando el sistema está en el estado $i \in S$, el controlador tiene varias acciones disponibles. Estas acciones comprenden un conjunto finito $A(i)$ (no vacío). El espacio de acciones es el conjunto de la unión de todas las $A(i)$, $i = 1, 2, \dots, K$, el cual se denota por A .
3. Si el sistema está en el estado i y una acción a se elige, entonces el estado al comienzo del siguiente periodo es j con probabilidad $p_{ij}(a)$, donde $\sum_{j \in S} p_{ij}(a) = 1$. $p_{ij}(a)$ se llama la probabilidad de transición del estado i al estado j , bajo la acción a .
4. Si el sistema se encuentra actualmente en el estado i y se elige una acción $a \in A(i)$, entonces se incurre en un costo (finito) no negativo $C(i, a)$.

Historia del proceso

La historia al tiempo $t = 0$ está dada por $h_0 = (i)$, $i \in S$. Se elige una acción inicial de $A(i)$ de acuerdo a la distribución $\pi(\cdot|h_0)$. Esta es la distribución de probabilidad de la acción $a \in A(i)$. Suponga que $a_0 \in A(i)$ se selecciona. Entonces el estado del sistema en $t = 1$ está determinado por la distribución de probabilidad de transición asociada con i y a_0 . Suponga que ese estado es j . La historia al tiempo $t = 1$ está dada por $h_1 = (i, a_0, j)$. La acción al tiempo $t = 1$ se elige de $A(j)$ de

acuerdo a la distribución $\pi(\cdot|i, a_0, j) = \pi(\cdot|h_1)$. Una vez que esta acción se elige (por ejemplo, a_1), entonces el estado del sistema en $t = 2$ está determinado por la distribución asociada con j en a . Suponga que este estado es k . La historia al tiempo $t = 2$ está dada por $h_2 = (i, a_0, j, a_1, k)$. El proceso continúa de esta forma. Suponga que el proceso estuvo operando en los periodos $t = 0, 1, 2, \dots, n - 1$ y que el estado al tiempo $t = n$ está determinado. La historia al tiempo $t = n$ es $h_n = (i, a_0, i_1, a_1, \dots, i_{n-1}, a_{n-1}, i_n)$ de los estados pasados y acciones en el periodo actual. Entonces la acción en n se elige de acuerdo a la distribución de probabilidad $\pi(\cdot|h_n)$ en el conjunto de acciones asociado con i_n . Una vez que esta acción se elige, el estado en el tiempo $n + 1$ puede ser determinado. El proceso continúa de manera infinita. En general, H_n es el conjunto de historias bajo $n \in \mathbb{N}$.

2.1.1. Políticas

Una política es una regla de operación de una CDM que especifica como se elige la decisión (acción) de cada período, proporciona una prescripción para seleccionar una acción bajo cualquier posible estado del sistema futuro, dada una historia. Sea $t = 0$ el periodo inicial. La CDM puede operar en dos modos. En el modo de horizonte infinito, el sistema opera en periodos $t = 0, 1, 2, \dots$. En el modo de horizonte finito, el sistema opera en N periodos $t = 0, 1, 2, \dots, N - 1$, donde N es un entero positivo fijo.

Definición 2.1.2. Una política estacionaria, denotada por f , es una función $f : S \rightarrow A(i)$ tal que asocia con cada estado i una acción $f(i) \in A(i)$. Si en cualquier tiempo el controlador en-

cuentra el sistema en el estado i , entonces el controlador siempre elige la acción $f(i)$.

Así, una política estacionaria depende de la historia del proceso a través del estado actual. Para implementar una política estacionaria, el controlador necesita solo conocer el estado actual del sistema. Los estados pasados y acciones pasadas son irrelevantes.

Definición 2.1.3. Una política estacionaria aleatorizada, denotada por δ y asociada con cada estado i , es una distribución de probabilidad $\delta(i)$ sobre $A(i)$. Si en cualquier tiempo el controlador encuentra el sistema en el estado i , entonces siempre elige la acción a con probabilidad $\delta(i)(a)$.

Definición 2.1.4. Una política de Markov determinista es una sucesión $\pi = (f_0, f_1, f_2, \dots)$ de políticas estacionarias. La cual opera de la siguiente manera: si el proceso está en el estado i al tiempo $t = n$, entonces el controlador elige la acción $f_n(i)$. Así, una política de Markov determinista depende de la historia del proceso a través del estado actual y del índice de tiempo.

Definición 2.1.5. Una política de Markov aleatorizada es una sucesión $\pi = (\delta_0, \delta_1, \delta_2, \dots)$ de políticas estacionarias aleatorizadas. Y opera como sigue. Si el proceso está en el estado i al tiempo $t = n$, entonces el controlador elige la acción $a \in A(i)$ con probabilidad $\delta_n(i)(a)$. Así, una política de Markov aleatorizada depende de la historia del proceso a través del estado actual y del índice de tiempo.

Además, una política de horizonte infinito se entendera como una regla que se aplica a lo largo del tiempo.

2.1.2. Distribución Condicional de Costos

Suponga que i es el estado inicial y que el proceso opera bajo una política $\pi \in \Pi$, donde Π es el conjunto de todas las políticas. El estado al tiempo t es una variable aleatoria denotada por x_t ; similarmente la acción elegida al tiempo t es una variable aleatoria, denotada por a_t . La distribución de probabilidad conjunta de (x_t, a_t) está dada por: $P_\pi(x_t = j, a_t = a | x_0 = i)$, donde se entiende que $a \in A(j)$. En este caso, la distribución de probabilidad está bien definida. A continuación, se muestra el cálculo para $t = 0, 1, 2$.

Sea $P_\pi(a_0 = a | x_0 = i) = \pi(a|i)$. Para $t = 1$, se tiene:

$$P_\pi(x_1 = j, a_1 = a | x_0 = i) = \sum_{b \in A(i)} \pi(b|i) p_{ij}(b) \pi(a|i, b, j).$$

El último término es la probabilidad de elegir la acción b , transitando a j , y eligiendo la acción a ; los términos se suman sobre las acciones $b \in A(i)$.

Para $t = 2$, se tiene:

$$P_\pi(x_2 = j, a_2 = a | x_0 = i) = \sum_{b \in A(i)} \pi(b|i) \left(\sum_k p_{ik}(b) \sum_{d \in A(k)} \pi(d|i, b, k) p_{kj}(d) \pi(a|i, b, k, d, j) \right).$$

Se asocia con el par aleatorio (x_t, a_t) el costo $C(x_t, a_t)$, es decir, es el costo incurrido al tiempo t cuando un controlador opera bajo π . Dado que $C(x_t, a_t)$ es una variable aleatoria, para

determinarlo se emplea la esperanza dada por:

$$E_{\pi}[C(x_t, a_t)|x_0 = i] = \sum_j \sum_{a \in A(j)} C(j, a) P_{\pi}(x_t = j, a_t = a | x_0 = i), \quad (2.1)$$

1 representa el costo esperado al tiempo t . Dado que los costos son no negativos, el valor esperado (2.1) está bien definido. Puede suceder que este valor sea $+\infty$.

Si π es una política estacionaria f , se emplea una notación especial. El costo asociado en el estado i está denotado por $C(i, f)$, donde $f = f(i)$. Similarmente, la probabilidad de transición se denotará por $p_{ij}(f)$ (ver [3] y [14]). Entonces $P_f(x_t = j, a_t = a | x_0 = i)$ es cero al menos que $a = f(j)$, y se tiene que:

$$\begin{aligned} P_f(x_t = j, a_t = f(j) | x_0 = i) &= P_f(x_t = j | x_0 = i) \\ &= \sum_{k_1 \in S} p_{ik_1}(f) \sum_{k_2 \in S} p_{k_1 k_2}(f) \cdots \sum_{k_{t-1} \in S} p_{k_{t-1} j}(f) \\ &:= p_{ij}^{(t)}(f). \end{aligned}$$

Aplicando (2.1) cuando se fija a_t y por la ecuación anterior se obtiene:

$$\begin{aligned} E_f[C(x_t, a_t) | x_0 = i] &= \sum_j C(j, f) P_f(x_t = j | x_0 = i) \\ &= \sum_j C(j, f) p_{ij}^{(t)}(f). \end{aligned}$$

Este es el costo esperado al tiempo t , bajo la política estacionaria f .

2.1.3. Criterio de Optimización

El criterio de optimización de interés en este trabajo es el costo promedio.

Suponga que el proceso opera bajo un horizonte finito n y que existe un costo terminal no negativo $F(k)$ que se incurre cuando el proceso está en el estado k . Sean i un estado inicial, n el horizonte y $\pi \in \Pi$ una política dada.

Definición 2.1.6. El costo total esperado de horizonte n bajo la política π se define como:

$$V_{\pi,n}(i) = E_{\pi}^i \left[\sum_{t=0}^{n-1} C(x_t, a_t) + F(x_n) \right].$$

Observe que

$$\begin{aligned} V_{\pi,n}(i) &= E_{\pi}^i \left[\sum_{t=0}^{n-1} C(x_t, a_t) + F(x_n) \right] \\ &= E_{\pi} \left[\sum_{t=0}^{n-1} [C(x_t, a_t) + F(x_n) | x_0 = i] \right] \\ &= \sum_{t=0}^{n-1} E_{\pi} [C(x_t, a_t) | x_0 = i] + E_{\pi} [F(x_n) | x_0 = i]. \end{aligned}$$

La última igualdad se sigue por la linealidad del valor esperado.

Definición 2.1.7. La función de valor óptimo de horizonte finito n , se define como:

$$V_n(i) = \inf_{\pi \in \Pi} V_{\pi,n}(i). \quad (2.2)$$

Definición 2.1.8. Una política π es óptima para el criterio de costo total esperado de horizonte n si $V_{\pi,n}(i) = V_n(i)$ para $i \in S$.

Definición 2.1.9. El índice de funcionamiento está definido por

$$V_{\pi,\alpha}(i) = E_{\pi}^i \left[\sum_{t=0}^{\infty} \alpha^t C(x_t, a_t) \right], \quad (2.3)$$

$\pi \in \Pi$, $i \in S$, llamado el costo descontado total esperado, donde $0 < \alpha < 1$ es el factor de descuento.

Definición 2.1.10. Una política $\pi^* \in \Pi$ que satisfaga

$$V_{\alpha}(i) =: V_{\pi^*,\alpha}(i) = \min_{\pi \in \Pi} V_{\pi,\alpha}(i), \quad (2.4)$$

$i \in S$, se llama política óptima y $V_{\alpha}(i)$ función de valor óptima descontada.

Definición 2.1.11. El costo promedio se define como:

$$J_{\pi}(i) = \limsup_{n \rightarrow \infty} \frac{1}{n} E_{\pi}^i \left[\sum_{t=0}^{n-1} C(x_t, a_t) \right] \quad (2.5)$$

$$= \limsup_{n \rightarrow \infty} \frac{v_{\pi,n}(i)}{n}, \quad (2.6)$$

$i \in S$, π es una política de horizonte infinito y

$$v_{\pi,n}(i) = E_{\pi}^i \left[\sum_{t=0}^{n-1} C(x_t, a_t) \right]. \quad (2.7)$$

Definición 2.1.12. La función de valores óptimos para el costo promedio se define como:

$$J(i) = \inf_{\pi} J_{\pi}(i), \quad (2.8)$$

$i \in S$, donde el ínfimo se toma sobre todas las políticas de horizonte infinito.

Definición 2.1.13. Sea π una política para el horizonte infinito. Entonces π es óptima para el costo promedio si $J_\pi(i) = J(i)$ para $i \in S$.

En el caso a tiempo continuo se considera una Cadena de Decisión de Markov a Tiempo Continuo (CDMTC) que se usa para modelar el control de ciertos sistemas que ocurren a tiempo continuo.

La CDMTC denotada por Ψ tiene espacio de estados S , que es un conjunto numerable. Asociado con cada $i \in S$ existe un conjunto finito no vacío $A(i)$ de acciones disponible en i . Suponer que la acción $a \in A(i)$ se elige. Entonces se incurre en un costo $C(i, a)$ el cual consiste de un costo instantáneo $G(i, a)$ incurrido inmediatamente y un parámetro de costo $g(i, a)$, el cual ocurre hasta la siguiente transición.

El tiempo hasta la siguiente transición está distribuido exponencialmente (ver [14]) con parámetro $\nu(i, a)$. El nuevo estado se elige de acuerdo a la distribución de probabilidad $(p_{ij}(a))_{j \in S}$. La teoría puede ser desarrollada asignando $p_{ii}(a) > 0$.

Sin embargo, para muchos sistemas se tiene que $p_{ii}(a) = 0$, para este trabajo se supondrá esto último.

2.2. Optimización a Tiempo Discreto con Espacio Numerable de Estados

A continuación se analiza el criterio de costo promedio cuando la CDM tiene espacio numerable de estados S . Además, se proporcionan algunos resultados para establecer la ecuación de

optimalidad de costo promedio.

Para las siguientes demostraciones se supondrá que el estado inicial es $X_0 = i$ y no se mencionará explícitamente.

LEMA 2.2.1. Sea e una política estacionaria. Sea S un espacio numerable de estados. Suponga que existe una constante J y una función h acotada inferiormente en i tal que

$$J + h(i) \geq C(i, e) + \sum_j p_{ij}(e)h(j), \quad (2.9)$$

$i \in S$. Entonces $J_e(i) \leq J$ para $i \in S$.

Demostración. Por hipótesis, existe una constante no negativa L tal que $h(i) \geq -L$ para $i \in S$. Suponga que $X_0 = i$ y $X_1, X_2, \dots$ una sucesión de valores del proceso operando bajo una política e . Entonces de (2.9) para un estado arbitrario X_t se obtiene

$$J + h(X_t) \geq C(X_t, e) + E_e[h(X_{t+1})|X_t], \quad (2.10)$$

$t \geq 0$. Observe que $E_e[h(X_t)] < \infty$, esto se establece por

$$E_e[h(X_t)] \leq tJ + h(i),$$

lo cual se demuestra por inducción a continuación.

Para $t = 0$, se tiene que $E_e[h(X_0)] \leq h(i)$. Suponga que es cierto para t , es decir, $E_e[h(X_t)] \leq tJ + h(i)$. Por demostrar que vale para $t + 1$. Entonces de (2.10) se sigue que

$$E_e[h(X_{t+1})|X_t] \leq J + h(X_t).$$

Tomando el valor esperado en ambos lados y usando una propiedad del valor esperado ($E[E[X|Y]] = E[X]$), se obtiene

$$\begin{aligned} E_e[E_e[h(X_{t+1})|X_t]] &\leq E_e[J + h(X_t)] \\ E_e[h(X_{t+1})] &\leq J + E_e[h(X_t)] \\ &\leq J + tJ + h(i) = (t + 1)J + h(i). \end{aligned}$$

La tercera desigualdad se sigue de la hipótesis de inducción. Esto completa la inducción.

Ahora, aplicando el valor esperado en ambos lados de (2.10) se obtiene

$$\begin{aligned} E_e[C(X_t, e)] + E_e[E_e[h(X_{t+1})|X_t]] &\leq E_e[J] + E_e[h(X_t)] \\ E_e[C(X_t, e)] + E_e[h(X_{t+1})] &\leq J + E_e[h(X_t)] \\ E_e[C(X_t, e)] &\leq J + E_e[h(X_t)] - E_e[h(X_{t+1})], \end{aligned} \quad (2.11)$$

$t \geq 0$. Observe que no se tiene una indeterminación de la forma $\infty - \infty$. Sumando en la ecuación (2.11) de $t = 0$ hasta $n - 1$, y dividiendo por n se sigue que

$$\begin{aligned} \sum_{t=0}^{n-1} E_e[C(X_t, e)] &\leq \sum_{t=0}^{n-1} J + \sum_{t=0}^{n-1} E_e[h(X_t)] - \sum_{t=0}^{n-1} E_e[h(X_{t+1})] \\ \sum_{t=0}^{n-1} E_e[C(X_t, e)] &\leq nJ + h(i) - E_e[h(X_n)] \leq nJ + h(i) + L \\ \frac{\sum_{t=0}^{n-1} E_e[C(X_t, e)]}{n} &\leq J + \frac{h(i) + L}{n} \\ \nu_{e,n}(i) &\leq J + \frac{h(i) + L}{n}. \end{aligned} \quad (2.12)$$

Tomando el límite superior en ambos lados de (2.12) se tiene

$$J_e(i) \leq J,$$

que es el resultado que se quería demostrar. $\square$

2.2.1. Método de Aproximaciones Sucesivas

El método de aproximaciones sucesivas (MAS) se utiliza para calcular la función de valor y las políticas óptimas, cuando el espacio de estados es infinito numerable (ver [14]).

Suponga que se proporciona la Cadena de Decisión de Markov (CDM) Δ como en la definición 2.1.1, con el espacio de estados S infinito numerable.

Sea (Δ_N) una sucesión de CDM que converge a Δ . El espacio de estados de cada Δ_N es finito, en este caso se determina una política óptima para cada Δ_N . Entonces, bajo ciertas condiciones (ver pag. 132 de [14]) la política óptima en cada Δ_N converge a una política óptima en Δ .

Definición 2.2.2. Sea N_0 un entero no negativo. La sucesión $(\Delta_N)_{N \geq N_0}$ es una sucesión de aproximación para Δ si existe una sucesión creciente $(S_N)_{N \geq N_0}$ de subconjuntos finitos no vacíos de S tal que $\bigcup S_N = S$. Cada Δ_N es una CDM con espacio de estados S_N que satisface dos condiciones:

1. Para $i \in S_N$, el costo de tomar la acción $a \in A(i)$ es $C(i, a)$.
2. Para cada $i \in S_N$ y $a \in A(i)$, $p_i(a; N)$ (llamada distribución de aproximación) es una distribución de probabilidad en S_N tal que

$$\lim_{N \rightarrow \infty} p_{ij}(a; N) = p_{ij}(a), \quad (2.13)$$

$j \in S$.

Si se considera el caso de horizonte finito con un costo terminal $c(x_n)$ entonces, este mismo costo terminal se aplica para Δ_N . El entero N indica el nivel de aproximación. La CDM Δ_N tiene como espacio de estados un subconjunto finito de S . En este subconjunto finito el conjunto de acciones y los costos para Δ_N son exactamente los mismos como para Δ . Únicamente las probabilidades de transición son diferentes. Éstos forman las distribuciones sobre el conjunto finito que converge a un punto sobre la

distribución original en Δ . Recordar que sólo dos elementos son requeridos para especificar una sucesión de aproximaciones: el subconjunto finito y las distribuciones de aproximación.

Definición 2.2.3. *La sucesión de aproximación (Δ_N) es una sucesión de aproximación del tipo de aumentación, si la distribución de aproximación es como sigue: dados un estado $i \in S_N$ y $a \in A_{(i)}$, para cada $r \notin S_N$ existe una distribución de probabilidad $(q_j(i, a, r, N))_{j \in S_N}$, llamada distribución de aumento asociada con (i, a, r, N) , tal que*

$$p_{ij}(a; N) = p_{ij}(a) + \sum_{r \in S - S_N} p_{ir}(a) q_j(i, a, r, N), \quad (2.14)$$

$j \in S_N$.

PROPOSICIÓN 2.2.4. *La ecuación (2.14) define una distribución de probabilidad sobre S_N .*

Demostración. Para demostrar que (2.14) define una aproximación de una distribución de probabilidad, primero se demuestra que

$$\sum_{j \in S_N} p_{ij}(a; N) = 1,$$

$$\begin{aligned} \sum_{j \in S_N} p_{ij}(a; N) &= \sum_{j \in S_N} p_{ij}(a) + \sum_{j \in S_N} \sum_{r \notin S_N} p_{ir}(a) q_j(i, a, r, N) \\ &= \sum_{j \in S_N} p_{ij}(a) + \sum_{r \notin S_N} p_{ir}(a) \left(\sum_{j \in S_N} q_j(i, a, r, N) \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{j \in S_N} p_{ij}(a) + \sum_{r \notin S_N} p_{ir}(a) \\
&= \sum_{j \in S} p_{ij}(a) \\
&= 1.
\end{aligned}$$

La primera igualdad es por sustituir (2.14), la segunda es válida ya que los términos son no negativos, la tercera igualdad se sigue debido a que las probabilidades en una distribución de aumentación suman 1.

Ahora, se demuestra que la distribución en (2.14) satisface (2.13). Observando primero que

$$1 = \sum_{j \in S_N} p_{ij}(a) + \sum_{r \notin S_N} p_{ir}(a). \quad (2.15)$$

Ya que el conjunto finito S_N crece a S , entonces el primer término del lado derecho de (2.15) se aproxima a 1 cuando $N \rightarrow \infty$. Por lo que se tiene que

$$\lim_{N \rightarrow \infty} \sum_{r \notin S_N} p_{ir}(a) = 0. \quad (2.16)$$

Fijando $j \in S$, suponga ¹ que N es suficientemente grande tal que $j \in S_N$. Como los términos en (2.14) son no negativos y $q_j \leq 1$, se sigue que

$$p_{ij}(a) \leq p_{ij}(a; N) \leq p_{ij}(a) + \sum_{r \notin S_N} p_{ir}(a). \quad (2.17)$$

Tomando el límite cuando $N \rightarrow \infty$ en la ecuación (2.17) y por (2.16), se concluye que (2.14) denota una distribución de aproximación. $\square$

2.3. Optimización del Caso a Tiempo Discreto con Espacio Finito de Estados

A continuación se analiza el criterio de costo promedio para la CDM con espacio finito de estados S . Además, se proporcionan algunos resultados para calcular la ecuación de optimalidad de costo promedio.

NOTACIÓN 2.3.1. Sea S un espacio finito de estados se denota por:

$$\begin{aligned} W_\alpha(i) &= \sum_k p_k(i) V_\alpha(z_k), \\ w_\alpha(i) &= V_\alpha(i) - W_\alpha(i), \\ w(i) &= \lim_{\alpha \rightarrow 1^-} w_\alpha(i), \\ w^*(i) &= w(i) - \sum_k p_k(i) \left(\sum_{s \in R_k} \pi_s(f) w(s) \right), \\ p_k(i) &= \sum_j p_{ij}(f) p_k(j), \end{aligned}$$

donde $\alpha \in (0, 1)$ es el factor de descuento, z_k es un estado fijo de la clase recurrente R_k llamado estado distinguido, $\pi_s(f)$ es la probabilidad del estado estacionario asociado con $s \in R_k$ y $p_k(i)$ es la probabilidad de que la clase recurrente R_k sea alcanzada.

Además,

$$V_\alpha(i) = \min_{a \in A(i)} \left\{ C(i, a) + \alpha \sum_{j \in S} p_{ij}(a) V_\alpha(j) \right\}, \quad (2.18)$$

$i \in S$, es la ecuación de optimalidad descontada y cualquier política estacionaria f que alcanza el mínimo en (2.18) es óptima (ver pag. 68 de [3] y pag. 36 de [14]).

Teorema 1. Sea S un espacio finito de estados, entonces la ecuación de optimalidad de costo promedio es

$$J(i) + w(i) = C(i, f) + \sum_j p_{ij}(a)w(j), \quad (2.19)$$

para $i \in S$.

Demostración. Condicionando sobre el primer estado visitado,

$$p_k(i) = \sum_j p_{ij}(f)p_k(j),$$

esto es válido sólo si i es una clase recurrente positiva. Multiplicando ambos lados de la ecuación anterior por $V_\alpha(z_k)$ y sumando sobre k ,

$$\begin{aligned} p_k(i)V_\alpha(z_k) &= \sum_j p_{ij}(f)p_k(j)V_\alpha(z_k) \\ \sum_k p_k(i)V_\alpha(z_k) &= \sum_k \sum_j p_{ij}(f)p_k(j)V_\alpha(z_k) \\ W_\alpha(i) &= \sum_j p_{ij}(f)W_\alpha(j). \end{aligned} \quad (2.20)$$

Para $\alpha > \alpha_0$ y f la política óptima, la ecuación de optimalidad descontada (2.18) se puede escribir como

$$V_\alpha(i) = C(i, f) + \alpha \sum_{j \in S} p_{ij}(f)V_\alpha(j),$$

restando y sumando $W_\alpha(i)$ en el lado derecho y restando $\alpha W_\alpha(i)$ en ambos lados

$$\begin{aligned} V_\alpha(i) - W_\alpha(i) &+ W_\alpha(i) - \alpha W_\alpha(i) \\ &= C(i, f) + \alpha \sum_{j \in S} p_{ij}(f)V_\alpha(j) - \alpha W_\alpha(i), \end{aligned}$$

entonces

$$\begin{aligned}
 w_\alpha(i) + (1 - \alpha)W_\alpha(i) &= C(i, f) + \alpha \sum_{j \in S} p_{ij}(f) V_\alpha(j) \\
 &\quad - \alpha \sum_{j \in S} p_{ij}(f) W_\alpha(j) \\
 &= C(i, f) + \alpha \sum_{j \in S} p_{ij}(f) [V_\alpha(j) - W_\alpha(j)] \\
 &= C(i, f) + \alpha \sum_{j \in S} p_{ij}(f) w_\alpha(j). \quad (2.21)
 \end{aligned}$$

Aplicando el límite cuando $\alpha \rightarrow 1^-$ en ambos lados de (2.21)

$$\lim_{\alpha \rightarrow 1^-} [w_\alpha(i) + (1 - \alpha)W_\alpha(i)] = \lim_{\alpha \rightarrow 1^-} \left[C(i, f) + \alpha \sum_{j \in S} p_{ij}(f) w_\alpha(j) \right]$$

$$\begin{aligned}
 \lim_{\alpha \rightarrow 1^-} w_\alpha(i) + \lim_{\alpha \rightarrow 1^-} (1 - \alpha)W_\alpha(i) &= \lim_{\alpha \rightarrow 1^-} C(i, f) \\
 &\quad + \lim_{\alpha \rightarrow 1^-} \left[\alpha \sum_{j \in S} p_{ij}(f) w_\alpha(j) \right]
 \end{aligned}$$

$$\begin{aligned}
 w(i) + \lim_{\alpha \rightarrow 1^-} (1 - \alpha)W_\alpha(i) &= C(i, f) \\
 &\quad + \lim_{\alpha \rightarrow 1^-} [\alpha] \lim_{\alpha \rightarrow 1^-} \left[\sum_{j \in S} p_{ij}(f) w_\alpha(j) \right] \\
 w(i) + \lim_{\alpha \rightarrow 1^-} (1 - \alpha)W_\alpha(i) &= C(i, f) + \sum_{j \in S} p_{ij}(f) w(j), \quad (2.22)
 \end{aligned}$$

donde

$$\begin{aligned}
 \lim_{\alpha \rightarrow 1^-} (1 - \alpha) W_\alpha(i) &= \lim_{\alpha \rightarrow 1^-} (1 - \alpha) \sum_k p_k(i) V_\alpha(z_k) \\
 &= \sum_k p_k(i) \lim_{\alpha \rightarrow 1^-} (1 - \alpha) V_\alpha(z_k) \\
 &= \sum_k p_k(i) J(z_k) \\
 &= \sum_k p_k(i) J_k \\
 &= J(i),
 \end{aligned}$$

la primera igualdad es por definición de $W_\alpha(i)$, como todos los términos son finitos y convergen se obtiene la segunda igualdad, para la tercera igualdad se tiene que

$$J(i) = \lim_{\alpha \rightarrow 1^-} (1 - \alpha) V_\alpha(i), \quad (2.23)$$

para $i \in S$ (ver pag. 100 de [14]) y para las dos últimas igualdades se tiene que el costo promedio bajo una política f es constante sobre la clase recurrente, denotado por J_k , por otra parte $J(i)$ es una combinación del costo promedio sobre la clase recurrente positiva.

Por lo anterior, la ecuación (2.22) resulta

$$w(i) + J(i) = C(i, f) + \sum_{j \in S} p_{ij}(f) w(j),$$

que es lo que se quería demostrar. $\square$

Teorema 2. *Sea S un espacio finito de estados. Entonces*

$$V_\alpha(i) = \frac{J(i)}{1 - \alpha} + w^*(i) + \epsilon_\alpha(i), \quad (2.24)$$

para $i \in S$ y $\alpha \in (0, 1)$, donde $\epsilon_\alpha(i)$ es una función que se aproxima a 0 cuando $\alpha \rightarrow 1^-$.

Demostración. Observe que el término en la Notación 2.3.1, $\sum_{s \in R_k} \pi_s(f)w(s)$ es el valor esperado de w sobre R_k . Por lo tanto, se considera a w^* como una versión normalizada de w .

De la ecuación (2.19) de optimalidad de costo promedio,

$$J(i) + w(i) = C(i, f) + \sum_j p_{ij}(f)w(j),$$

restándole $\sum_k p_k(i) \left(\sum_{s \in R_k} \pi_s(f)w(s) \right)$ se obtiene:

$$\begin{aligned} J(i) + w(i) - \sum_k p_k(i) \left(\sum_{s \in R_k} \pi_s(f)w(s) \right) \\ = C(i, f) + \sum_j p_{ij}(f)w(j) - \sum_k p_k(i) \left(\sum_{s \in R_k} \pi_s(f)w(s) \right), \end{aligned}$$

sustituyendo $p_k(i) = \sum_j p_{ij}(f)p_k(j)$ en el segundo lado de la igualdad

$$\begin{aligned} J(i) + w^*(i) &= C(i, f) + \sum_j p_{ij}(f)w(j) \\ &\quad - \sum_k \left(\sum_j p_{ij}(f)p_k(j) \right) \left(\sum_{s \in R_k} \pi_s(f)w(s) \right) \\ &= C(i, f) + \sum_j p_{ij}(f)w(j) \\ &\quad - \sum_j p_{ij}(f) \sum_k p_k(j) \left(\sum_{s \in R_k} \pi_s(f)w(s) \right) \end{aligned}$$

$$\begin{aligned}
&= C(i, f) + \sum_j p_{ij}(f) \cdot \\
&\quad \left[w(j) - \sum_k p_k(j) \left(\sum_{s \in R_k} \pi_s(f) w(s) \right) \right] \\
&= C(i, f) + \sum_j p_{ij}(f) w^*(j), \tag{2.25}
\end{aligned}$$

aplicando el operador esperanza en ambos lados y por una propiedad de esperanza ($E[E[X]] = E[X]$) se obtiene:

$$\begin{aligned}
E_f[J(i) + w^*(i)] &= E_f \left[C(i, f) + \sum_j p_{ij}(f) w^*(j) \right] \\
J(i) + E_f[w^*(X_t)] &= E_f[C(X_t, f)] + E_f[w^*(X_{t+1})] \\
J(i) + E_f[w^*(X_t)] &= E_f[C(X_t, f)] + E_f[w^*(X_{t+1})] \\
J(i) + E_f[w^*(X_t)] &- E_f[w^*(X_{t+1})] = E_f[C(X_t, f)]. \tag{2.26}
\end{aligned}$$

Multiplicando ambos lados de (2.26) por α^t y sumando sobre t :

$$\begin{aligned}
E_f[C(X_t, f)] &= J(i) + E_f[w^*(X_t)] - E_f[w^*(X_{t+1})] \\
\alpha^t E_f[C(X_t, f)] &= \alpha^t J(i) + \alpha^t E_f[w^*(X_t)] - \alpha^t E_f[w^*(X_{t+1})] \\
\sum_{t=0}^{\infty} E_f[\alpha^t C(X_t, f)] &= J(i) \sum_{t=0}^{\infty} \alpha^t + \sum_{t=0}^{\infty} \alpha^t E_f[w^*(X_t)] \\
&\quad - \sum_{t=0}^{\infty} \alpha^t E_f[w^*(X_{t+1})],
\end{aligned}$$

por (2.3),

$$\begin{aligned}
V_{f,\alpha}(i) &= J(i) \frac{1}{1-\alpha} + \alpha^0 E_f[w^*(X_0 = i)] + \sum_{t=1}^{\infty} \alpha^t E_f[w^*(X_t)] \\
&\quad - \sum_{t=0}^{\infty} \alpha^t E_f[w^*(X_{t+1})] \\
&= \frac{J(i)}{1-\alpha} + w^*(i) + \sum_{t=0}^{\infty} \alpha^{t+1} E_f[w^*(X_{t+1})] \\
&\quad - \sum_{t=0}^{\infty} \alpha^t E_f[w^*(X_{t+1})] \\
&= \frac{J(i)}{1-\alpha} + w^*(i) + \alpha \sum_{t=0}^{\infty} \alpha^t E_f[w^*(X_{t+1})] \\
&\quad - \sum_{t=0}^{\infty} \alpha^t E_f[w^*(X_{t+1})] \\
&= \frac{J(i)}{1-\alpha} + w^*(i) - (1-\alpha) \sum_{t=0}^{\infty} \alpha^t E_f[w^*(X_{t+1})]. \quad (2.27)
\end{aligned}$$

Como f es óptima descontada para $\alpha \in (\alpha_0, 1)$, (2.27) implica que:

$$V_{\alpha}(i) = \frac{J(i)}{1-\alpha} + w^*(i) - (1-\alpha) \sum_{t=0}^{\infty} \alpha^t E_f[w^*(X_{t+1})]. \quad (2.28)$$

Para $\alpha \in (0, \alpha_0]$ la función $\epsilon_{\alpha}(i)$ se puede definir para establecer la igualdad en (2.24). Para $\alpha \in (\alpha_0, 1)$, $\epsilon_{\alpha}(i)$ se define como el último término en (2.28), la cual se puede reescribir como sigue:

$$\epsilon_{\alpha}(i) = \frac{(1-\alpha)w^*(i)}{\alpha} - \frac{1-\alpha}{\alpha} \sum_{t=0}^{\infty} \alpha^t E_f[w^*(X_t)],$$

de lo anterior se observa que el primer término es 0 cuando $\alpha \rightarrow 1^-$. Del segundo término ignorando el α que está en el denominador, considerando el costo promedio w^* se obtiene como un límite igual a $\sum_k p_k(i) (\sum_{s \in R_k} \pi_s(f) w^*(s))$ y por la definición de w^* , esto es 0. Utilizando el teorema A.4.2 (ver pag. 282 de [14]) se obtiene que $\epsilon_\alpha(i)$ es una función que aproxima a 0 cuando $\alpha \rightarrow 1^-$. $\square$

PROPOSICIÓN 2.3.2. Sean $J(i) \equiv J$ para $i \in S$, z un estado distinguido fijo, $h_\alpha(i) = V_\alpha(i) - V_\alpha(z)$ y L una cota para $|h_\alpha|$. Entonces

a) Para $i \in S$ se tiene

$$\lim_{\alpha \rightarrow 1^-} h_\alpha(i) = h(i) = w^*(i) - w^*(z).$$

b) La ecuación de optimalidad de costo promedio es

$$\begin{aligned} J + h(i) &= \min_a \left\{ C(i, a) + \sum_j p_{ij}(a) h(j) \right\} \\ &= C(i, f) + \sum_j p_{ij}(f) h(j), \end{aligned} \quad (2.29)$$

donde $i \in S$ y f es la política estacionaria óptima.

Demostración. Para demostrar (a), por hipótesis y por el teorema 2 se obtiene:

$$\begin{aligned} h_\alpha(i) &= V_\alpha(i) - V_\alpha(z) \\ &= \frac{J(i)}{1 - \alpha} + w^*(i) + \epsilon_\alpha(i) - \frac{J(z)}{1 - \alpha} - w^*(z) - \epsilon_\alpha(z) \end{aligned}$$

$$\begin{aligned}
&= \frac{J(i) - J(z)}{1 - \alpha} + w^*(i) + \epsilon_\alpha(i) - w^*(z) - \epsilon_\alpha(z) \\
&= \frac{J - J}{1 - \alpha} + w^*(i) + \epsilon_\alpha(i) - w^*(z) - \epsilon_\alpha(z) \\
&= w^*(i) + \epsilon_\alpha(i) - w^*(z) - \epsilon_\alpha(z).
\end{aligned}$$

Aplicando el límite cuando $\alpha \rightarrow 1^-$ se alcanza

$$\begin{aligned}
\lim_{\alpha \rightarrow 1^-} h_\alpha(i) &= \lim_{\alpha \rightarrow 1^-} [w^*(i) + \epsilon_\alpha(i) - w^*(z) - \epsilon_\alpha(z)] \\
h(i) &= w^*(i) - w^*(z) + \lim_{\alpha \rightarrow 1^-} [\epsilon_\alpha(i) - \epsilon_\alpha(z)] \\
h(i) &= w^*(i) - w^*(z).
\end{aligned}$$

Que es lo que se quería demostrar en el inciso (a).

Para demostrar (b), observe que la ecuación de optimalidad descontada (2.18),

$$V_\alpha(i) = \min_a \left\{ C(i, a) + \alpha \sum_j p_{ij}(a) V_\alpha(j) \right\}, \quad (2.30)$$

se obtiene

$$\begin{aligned}
V_\alpha(i) &= V_\alpha(z) + V_\alpha(z) - \alpha V_\alpha(z) \\
&= \min_a \left\{ C(i, a) + \alpha \sum_j p_{ij}(a) V_\alpha(j) \right\} - \alpha V_\alpha(z) \\
h_\alpha(i) &= (1 - \alpha) V_\alpha(z) \\
&= \min_a \left\{ C(i, a) + \alpha \sum_j p_{ij}(a) V_\alpha(j) - \alpha V_\alpha(z) \right\}
\end{aligned}$$

$$\begin{aligned}
 &= \min_a \left\{ C(i, a) + \alpha \sum_j p_{ij}(a) V_\alpha(j) - \alpha \sum_j p_{ij}(a) V_\alpha(z) \right\} \\
 &= \min_a \left\{ C(i, a) + \alpha \sum_j p_{ij}(a) [V_\alpha(j) - V_\alpha(z)] \right\} \\
 &= \min_a \left\{ C(i, a) + \alpha \sum_j p_{ij}(a) h_\alpha(j) \right\}, \tag{2.31}
 \end{aligned}$$

la segunda igualdad es por hipótesis, en la tercera igualdad se tiene que $\sum_j p_{ij}(a) = 1$ y la última igualdad es por hipótesis. Además f alcanza el mínimo para $\alpha > \alpha_0$. Tomando el límite cuando $\alpha \rightarrow 1^-$ en (2.31) resulta

$$\begin{aligned}
 &\lim_{\alpha \rightarrow 1^-} [h_\alpha(i) + (1 - \alpha)V_\alpha(z)] \\
 &= \lim_{\alpha \rightarrow 1^-} \min_a \left\{ C(i, a) + \alpha \sum_j p_{ij}(a) h_\alpha(j) \right\} \\
 &\lim_{\alpha \rightarrow 1^-} h_\alpha(i) + \lim_{\alpha \rightarrow 1^-} (1 - \alpha)V_\alpha(z) \\
 &= \min_a \left\{ \lim_{\alpha \rightarrow 1^-} C(i, a) + \lim_{\alpha \rightarrow 1^-} \alpha \sum_j p_{ij}(a) h_\alpha(j) \right\} \\
 h(i) + J(z) &= \min_a \left\{ C(i, a) + \sum_j p_{ij}(a) h(j) \right\} \\
 h(i) + J &= \min_a \left\{ C(i, a) + \sum_j p_{ij}(a) h(j) \right\}, \tag{2.32}
 \end{aligned}$$

la tercera igualdad es por (2.23) y la última es por la hipótesis $J(i) \equiv J$. Por lo tanto, (2.32) es lo que se quería demostrar en el inciso (b). $\square$

Teorema 3. Sea Δ una CDM con espacio finito de estados S .

1. Suponga que se tiene una constante F y una función r tal que

$$F + r(i) \geq \min_a \left\{ C(i, a) + \sum_j p_{ij}(a)r(j) \right\}, \quad (2.33)$$

$i \in S$. Si e es una política estacionaria que optimiza (2.33), entonces $J_e(i) \leq F$ para $i \in S$. Por lo tanto, si F es una cota inferior para el costo promedio, entonces el costo promedio mínimo es igual a la constante F y e es óptima de costo promedio.

2. Suponga que se tiene una constante F y una función r tal que

$$F + r(i) = \min_a \left\{ C(i, a) + \sum_j p_{ij}(a)r(j) \right\}, \quad (2.34)$$

$i \in S$. Entonces el costo promedio mínimo es igual a la constante F y cualquier política estacionaria que optimiza (2.34) es óptima de costo promedio. Si h es como en la proposición 2.3.2, entonces $r(i) = h(i) + r(z_k) - h(z_k)$ para $i \in R_k$. (Observe que éste último es una clase recurrente bajo f con espacio de estados distinguidos z_k). Para i transitorio bajo f se cumple que $r(i) \leq h(i) + \sum_k P_k(i)[r(z_k) - h(z_k)]$.

Demostración. Para demostrar (1), suponga que (2.33) es válida y que e es una política estacionaria que alcanza el mínimo, es decir,

$$F + r(i) \geq C(i, e) + \sum_j p_{ij}(e)r(j). \quad (2.35)$$

Suprimiendo el estado inicial $X_0 = i$ y para un estado arbitrario X_t se tiene de la ecuación anterior que

$$F + r(X_t) \geq C(X_t, e) + E_e[r(X_{t+1})|X_t], \quad (2.36)$$

$t \geq 0$, tomando el operador esperanza en ambos lados de (2.36) y por una propiedad de esperanza ($E[E[X|Y]] = E[X]$):

$$\begin{aligned} E_e[F + r(X_t)] &\geq E_e[C(X_t, e) + E_e[r(X_{t+1})|X_t]] \\ E_e[F] + E_e[r(X_t)] &\geq E_e[C(X_t, e)] + E_e[E_e[r(X_{t+1})|X_t]] \\ F + E_e[r(X_t)] &\geq E_e[C(X_t, e)] + E_e[r(X_{t+1})], \end{aligned}$$

si los valores esperados son finitos entonces

$$F + E_e[r(X_t)] - E_e[r(X_{t+1})] \geq E_e[C(X_t, e)],$$

sumando de $t = 0$ hasta $n-1$ y dividiendo entre n en la ecuación anterior, se sigue que

$$\begin{aligned} \sum_{t=0}^{n-1} F + \sum_{t=0}^{n-1} E_e[r(X_t)] - \sum_{t=0}^{n-1} E_e[r(X_{t+1})] &\geq \sum_{t=0}^{n-1} E_e[C(X_t, e)] \\ nF + E_e[r(X_0)] - E_e[r(X_n)] &\geq \sum_{t=0}^{n-1} E_e[C(X_t, e)] \\ nF + r(i) - E_e[r(X_n)] &\geq \sum_{t=0}^{n-1} E_e[C(X_t, e)] \\ F + \frac{r(i) - E_e[r(X_n)]}{n} &\geq \frac{\sum_{t=0}^{n-1} E_e[C(X_t, e)]}{n} \\ F + \frac{r(i) + M}{n} &\geq v_{e,n}(i), \end{aligned}$$

la última desigualdad es porque $-M$ es una cota inferior para r

y por (2.7). Tomando el límite en ambos lados se obtiene

$$\lim_{n \rightarrow \infty} \left(F + \frac{r(i) + M}{n} \right) \geq \lim_{n \rightarrow \infty} \left(\frac{v_{e,n}(i)}{n} \right)$$

$$F \geq J_e(i).$$

Si F es una cota inferior para el costo promedio, entonces

$$J(i) \leq J_e(i) \leq F \leq J(i).$$

Por lo tanto se tiene que

$$J(i) = J_e(i) \equiv F.$$

Para demostrar (2), si (2.34) es válida entonces se sigue que

$$F + r(i) = \min_a \left\{ C(i, a) + \sum_j p_{ij}(a)r(j) \right\}$$

$$\leq C(i, f) + \sum_j p_{ij}(f)r(j), \quad (2.37)$$

suprimiendo el estado inicial $X_0 = i$ y para un estado arbitrario X_t se tiene de (2.37) que

$$F + r(X_t) \leq C(X_t, f) + E_f[r(X_{t+1})|X_t], \quad (2.38)$$

$t \geq 0$, tomando el operador esperanza en ambos lados de (2.38) y por una propiedad de esperanza ($E[E[Y|X]] = E[Y]$),

$$E_f[F + r(X_t)] \leq E_f[C(X_t, f) + E_f[r(X_{t+1})|X_t]]$$

$$E_f[F] + E_f[r(X_t)] \leq E_f[C(X_t, f)] + E_f[E_f[r(X_{t+1})|X_t]]$$

$$F + E_f[r(X_t)] \leq E_f[C(X_t, f)] + E_f[r(X_{t+1})],$$

si las esperanzas son finitas entonces

$$F + E_f[r(X_t)] - E_f[r(X_{t+1})] \leq E_f[C(X_t, f)],$$

sumando de $t = 0$ hasta $n - 1$ y dividiendo entre n en la ecuación anterior, se sigue que

$$\begin{aligned}
 \sum_{t=0}^{n-1} F + \sum_{t=0}^{n-1} E_e[r(X_t)] - \sum_{t=0}^{n-1} E_e[r(X_{t+1})] &\leq \sum_{t=0}^{n-1} E_e[C(X_t, e)] \\
 nF + E_e[r(X_0)] - E_e[r(X_n)] &\leq \sum_{t=0}^{n-1} E_e[C(X_t, e)] \\
 nF + r(i) - E_e[r(X_n)] &\leq \sum_{t=0}^{n-1} E_e[C(X_t, e)] \\
 F + \frac{r(i) - E_e[r(X_n)]}{n} &\leq \frac{\sum_{t=0}^{n-1} E_e[C(X_t, e)]}{n} \\
 F + \frac{r(i) + M}{n} &\leq v_{e,n}(i),
 \end{aligned}$$

la última desigualdad es porque $-M$ es una cota inferior para r y por (2.7). Tomando el límite en ambos lados y usando la optimalidad de f se obtiene

$$\begin{aligned}
 \lim_{n \rightarrow \infty} \left[F + \frac{r(i) + M}{n} \right] &\leq \lim_{n \rightarrow \infty} \left[\frac{v_{e,n}(i)}{n} \right] \\
 F &\leq J_f(i) = J(i),
 \end{aligned}$$

de (1) se tiene que el costo promedio es igual a la constante F y f es óptima de costo promedio, es decir,

$$J(i) = J_f(i) = F.$$

Entonces

$$F + r(i) \leq C(i, f) + \sum_j p_{ij}(f)r(j), \quad (2.39)$$

$$F + h(i) = C(i, f) + \sum_j p_{ij}(f)h(j), \quad (2.40)$$

la segunda igualdad se sigue de la proposición (2.3.2). Sea $b = r - h$. Restando (2.39) y (2.40):

$$\begin{aligned} r(i) - h(i) &\leq \sum_j p_{ij}(f)[r(j) - h(j)] \\ b(i) &\leq \sum_j b(j)p_{ij}(f), \end{aligned} \quad (2.41)$$

iterando (2.41), sumando los términos de $t = 1$ hasta n y dividiendo por n se obtiene:

$$\begin{aligned} b(i) &\leq \sum_j b(j)p_{ij}^{(t)}(f) \\ \sum_{t=1}^n b(i) &\leq \sum_{t=1}^n \sum_j b(j)p_{ij}^{(t)}(f) \\ nb(i) &\leq \sum_j b(j) \sum_{t=1}^n p_{ij}^{(t)}(f) \\ b(i) &\leq \frac{1}{n} \sum_j b(j) \sum_{t=1}^n p_{ij}^{(t)}(f) \\ b(i) &\leq \sum_j b(j) \left(\frac{1}{n} \sum_{t=1}^n p_{ij}^{(t)}(f) \right) \\ \lim_{n \rightarrow \infty} b(i) &\leq \lim_{n \rightarrow \infty} \sum_j b(j) \left(\frac{1}{n} \sum_{t=1}^n p_{ij}^{(t)}(f) \right) \\ b(i) &\leq \sum_j b(j) \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n p_{ij}^{(t)}(f) \right). \end{aligned} \quad (2.42)$$

Si j es transitorio, entonces $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n p_{ij}^{(t)}(f) = 0 \Rightarrow b(i) \leq 0$. Pero si $j \in R_k$, entonces

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n p_{ij}^{(t)}(f) = p_k(i) \pi_j(f).$$

Primero suponga que $i \in R_k$ en (2.42) y se obtiene

$$b(i) \leq \sum_j b(j) \pi_j(f). \quad (2.43)$$

Esto es posible si y sólo si la función b es una constante B_k . Esto implica que $B_k = b(z_k) = r(z_k) - h(z_k)$. Por lo tanto,

$$\begin{aligned} r(i) &= h(i) + B_k \\ &= h(i) + r(z_k) - h(z_k). \end{aligned}$$

Ahora, suponga que i es transitorio bajo f . Tomando el límite cuando $n \rightarrow \infty$ en (2.42) y usando el hecho que $b \equiv B_k$ sobre la clase recurrente R_k se obtiene

$$\begin{aligned} \lim_{n \rightarrow \infty} b(i) &\leq \lim_{n \rightarrow \infty} \sum_j b(j) \left(\frac{1}{n} \sum_{t=1}^n p_{ij}^{(t)}(f) \right) \\ b(i) &\leq \sum_j b(j) \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n p_{ij}^{(t)}(f) \right) \\ b(i) &\leq \sum_k p_k(i) B_k \\ b(i) &\leq \sum_k p_k(i) [r(z_k) - h(z_k)]. \end{aligned}$$

□

El algoritmo de iteración de valores es de gran utilidad para obtener la solución a la ecuación de optimalidad (2.34).

Algoritmo de Iteración de Valores

Sea Δ una CDM con espacio finito de estados. Suponga que el costo promedio mínimo es una constante y que e es una política estacionaria óptima, entonces cualquier clase recurrente positiva en la Cadena de Markov inducida por e es aperiódica. Sea x un estado distinguido y ε un número pequeño positivo.

Se presentan el Algoritmo de Iteración de Valores (AIV).

1. Sea $n = 0$ y $u_0 \equiv 0$.

2. Sea

$$\omega_n(i) = \min_{a \in A(x)} \left\{ C(i, a) + \sum_{j \in S} p_{ij}(a) u_n(j) \right\}. \quad (2.44)$$

3. Si $n = 0$, sea $\delta = 1$.

Si $n \geq 1$, entonces sea $\delta = \max_{i \in S} |\omega_n(i) - \omega_{n-1}(i)|$.

Si $\delta < \varepsilon$ se continua con el paso 6.

4. Sea $u_{n+1}(i) = \omega_n(i) - \omega_n(x)$.

5. Ir al paso 2 y reemplazar n por $n + 1$.

6. Parar en una etapa n y escribir la función de valor óptimo $w_n(x)$ y la política óptima que alcanza el mínimo en la ecuación (2.44).

2.4. Optimización del Caso a Tiempo Continuo

En esta sección se proporciona un conjunto de suposiciones que permiten calcular una política estacionaria de costo promedio óptima para la Cadena de Decisión de Markov a Tiempo continuo (CDMTC) Ψ (ver Sección 2.2). Existen dos definiciones comunes del costo promedio bajo una política arbitraria. La primera definición considera el costo esperado que se incurre bajo la política durante el intervalo $[0, t)$, dividiendo sobre t y tomando el límite superior de esta cantidad cuando $t \rightarrow \infty$. Para este trabajo se empleará la segunda definición, la cual considera el costo esperado incurrido durante n transiciones, dividido por el tiempo esperado de estas transiciones, tomando el límite superior cuando $n \rightarrow \infty$. Formalmente, sea $E_\pi[C_n]$ el costo total esperado incurrido bajo π durante los primeros n periodos de transición. Sea $E_\pi[T_n]$ el tiempo total esperado tomado bajo π para los n primeros periodos de transición. Entonces se define:

$$J_\pi^\Psi(i) = \limsup_{n \rightarrow \infty} \frac{E_\pi[C_n | x_0 = i]}{E_\pi[T_n | x_0 = i]}, \text{ y} \quad (2.45)$$

$$J^\Psi(i) = \inf_{\pi} J_\pi^\Psi(i), \quad (2.46)$$

$i \in S$.

La igualdad (2.45) está bien definida, ver la sección 2.1.2.

Notar ¹ que si el proceso está en el estado $i \in S$ y una acción $a \in A(i)$ se elige, entonces el tiempo esperado hasta un cambio del estado está dado por $\tau(i, a) = \frac{1}{v(i, a)}$.

HIPÓTESIS 2.4.1. *Suponga que el proceso se encuentra en el estado $i \in S$ y una acción $a \in A(i)$ se elige, entonces existen dos constantes b y B tales que:*

$$0 < b < \inf_{i,a} \tau(i, a) \leq \sup_{i,a} \tau(i, a) \leq B < \infty.$$

LEMA 2.4.2. *Sean Ψ una CDMTC que satisface la hipótesis 2.4.1 y e una política estacionaria para Ψ . Suponga que existe una constante Z y una función z que está acotada inferiormente en i tal que:*

$$Z\tau(i, e) + z(i) \geq G(i, e) + g(i, e)\tau(i, e) + \sum_j p_{ij}(e)z(j), \quad (2.47)$$

$i \in S$, $G(i, e)$ es un costo que se genera instantáneamente y $g(i, e)$ se genera por realizar una transición. Entonces $J_e^\Psi(i) \leq Z$ para $i \in S$.

Demostración. Por hipótesis se tiene que existe un L tal que $-L \leq z(i)$.

Sean $X_0 = i, X_1, X_2, \dots$ una sucesión de valores del proceso bajo la política e , considérese la ecuación (2.47) para un estado arbitrario X_k :

$$\begin{aligned} Z\tau(X_k, e) + z(X_k) &\geq G(X_k, e) + g(X_k, e)\tau(X_k, e) \\ &\quad + E_e[z(X_{k+1})|X_k]. \end{aligned} \quad (2.48)$$

Aplicando el operador esperanza en ambos lados de la ecuación

anterior:

$$\begin{aligned}
 E_e[Z\tau(X_k, e) + z(X_k)] &\geq E_e[G(X_k, e) + g(X_k, e)\tau(X_k, e) \\
 &\quad + E_e[z(X_{k+1})|X_k]] \\
 E_e[Z\tau(X_k, e)] + E_e[z(X_k)] &\geq E_e[G(X_k, e) + g(X_k, e)\tau(X_k, e)] \\
 &\quad + E_e[E_e[z(X_{k+1})|X_k]] \\
 ZE_e[\tau(X_k, e)] + E_e[z(X_k)] &\geq E_e[G(X_k, e) + g(X_k, e)\tau(X_k, e)] \\
 &\quad + E_e[z(X_{k+1})]. \tag{2.49}
 \end{aligned}$$

Ahora se demuestra que $E_e[z(X_{k+1})] < \infty$. Observe que

$$E_e[z(X_k)] \leq ZBk + z(i),$$

ya que para $k = 0$, $E_e[z(X_0 = i)] \leq ZB(0) + z(i) = z(i)$. Suponga que vale para k , es decir, $E_e[z(X_k)] \leq ZBk + z(i)$. Para demostrar que vale para $k + 1$, de la ecuación (2.48) se tiene que

$$\begin{aligned}
 Z\tau(X_k, e) + z(X_k) &\geq G(X_k, e) + g(X_k, e)\tau(X_k, e) \\
 &\quad + E_e[z(X_{k+1})|X_k] \\
 Z\tau(X_k, e) + z(X_k) &\geq E_e[z(X_{k+1})|X_k].
 \end{aligned}$$

Aplicando el operador esperanza en ambos lados:

$$\begin{aligned}
 E_e[Z\tau(X_k, e) + z(X_k)] &\geq E_e[E_e[z(X_{k+1})|X_k]] \\
 ZE_e[\tau(X_k, e)] + E_e[z(X_k)] &\geq E_e[z(X_{k+1})] \\
 ZB + ZBk + z(i) &\geq E_e[z(X_{k+1})] \\
 ZB(k + 1) + z(i) &\geq E_e[z(X_{k+1})].
 \end{aligned}$$

La tercera línea se sigue de la hipótesis 2.4.1, de la hipótesis para k y de la propiedad de la esperanza $E[E[X|Y]] = E[X]$. Por lo tanto, $E_e[z(X_{k+1})] < \infty$. Así, de la ecuación (2.49) se tiene:

$$\begin{aligned}
 ZE_e[\tau(X_k, e)] + E_e[z(X_k)] - E_e[z(X_{k+1})] \\
 \geq E_e[G(X_k, e) + g(X_k, e)\tau(X_k, e)],
 \end{aligned}$$

sumando de $k = 0$ hasta $n - 1$, en la ecuación anterior

$$\begin{aligned} Z \sum_{k=0}^{n-1} E_e[\tau(X_k, e)] + \sum_{k=0}^{n-1} E_e[z(X_k)] - \sum_{k=0}^{n-1} E_e[z(X_{k+1})] \\ \geq \sum_{k=0}^{n-1} E_e[G(X_k, e) + g(X_k, e)\tau(X_k, e)], \end{aligned}$$

se sigue que

$$\begin{aligned} Z \sum_{k=0}^{n-1} E_e[\tau(X_k, e)] + z(i) - E_e[z(X_n)] \\ \geq \sum_{k=0}^{n-1} E_e[G(X_k, e) + g(X_k, e)\tau(X_k, e)], \end{aligned}$$

dividiendo entre $\sum_{k=0}^{n-1} E_e[\tau(X_k, e)]$,

$$\begin{aligned} Z + \frac{z(i) - E_e[z(X_n)]}{\sum_{k=0}^{n-1} E_e[\tau(X_k, e)]} &\geq \frac{\sum_{k=0}^{n-1} E_e[G(X_k, e) + g(X_k, e)\tau(X_k, e)]}{\sum_{k=0}^{n-1} E_e[\tau(X_k, e)]} \\ Z + \frac{z(i) + L}{nb} &\geq \frac{\sum_{k=0}^{n-1} E_e[G(X_k, e) + g(X_k, e)\tau(X_k, e)]}{\sum_{k=0}^{n-1} E_e[\tau(X_k, e)]}, \end{aligned}$$

aplicando el límite superior cuando n tiende a infinito en ambos lados de la ecuación anterior:

$$\begin{aligned} \limsup_{n \rightarrow \infty} \left(Z + \frac{z(i) + L}{nb} \right) \\ \geq \limsup_{n \rightarrow \infty} \left(\frac{\sum_{k=0}^{n-1} E_e[G(X_k, e) + g(X_k, e)\tau(X_k, e)]}{\sum_{k=0}^{n-1} E_e[\tau(X_k, e)]} \right) \end{aligned}$$

$$Z \geq J_e^\Psi(i).$$

□

Ahora, se introduce una Cadena de Decisión de Markov auxiliar a tiempo discreto Δ que se relaciona estrechamente con el proceso a tiempo continuo Ψ . Entonces, se puede calcular una política estacionaria óptica de costo promedio para Δ . Bajo ciertas suposiciones, esta política también es óptima para Ψ (ver [15]), donde Ψ es el proceso original a tiempo continuo con espacio de estados infinito y Δ es el proceso auxiliar a tiempo discreto con espacio de estados infinito.

2.4.1. Método de Uniformización

Se muestra en esta sección la equivalencia entre una CDMTD y una CDMTC a través del método de uniformización el cual fue propuesto por Serfozo (ver [15]), para resolver problemas de PDMs a tiempo continuo.

Considerar una CDMTC, en la cual el tiempo de permanencia, $\nu(i, a)$, en un estado es el mismo para todos los estados (es decir, $\nu(i, a) = \lambda + a$). Las probabilidades de transición en el sistema se obtienen de la siguiente forma $p_{01} = 1$, $p_{ii+1}(a) = \frac{\lambda}{\lambda+a}$, $p_{ii-1}(a) = \frac{a}{\lambda+a}$, $i \geq 1$ donde la matriz de probabilidades de transición es la siguiente

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ \frac{a}{\lambda+a} & 0 & \frac{\lambda}{\lambda+a} & 0 & 0 & \dots \\ 0 & \frac{a}{\lambda+a} & 0 & \frac{\lambda}{\lambda+a} & 0 & \dots \\ 0 & 0 & \frac{a}{\lambda+a} & 0 & \frac{\lambda}{\lambda+a} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Ahora, se define la CDMTD Δ donde los estados y las acciones son los mismos que para la CDMTC Ψ . Para poder definir Δ se hace uso de la técnica de Uniformización, es decir, se uniformiza

el parámetro de tiempo en el cual una transición ocurre para cada estado, introduciendo transiciones de un estado a el mismo, es decir, este método asigna una probabilidad $p_{ii} \neq 0$. La probabilidad de transición que se propone es

$$p_{ij}^*(a) = \begin{cases} b\nu(i, a)p_{ij}(a), & j \neq i, \\ 1 - b\nu(i, a), & j = i, \end{cases} \quad (2.50)$$

donde b es como en la hipótesis 2.4.1. Observe de (2.50) que $b\nu(i, a) = b/\tau(i, a) \leq 1$. Si el proceso está en el estado i , entonces en cada período la probabilidad de transitar a $j \neq i$ es proporcional a la probabilidad en Ψ . Además, la probabilidad de permanecer en i no es cero. Esto puede ser contrastante a Ψ para el cual $p_{ii}(a) = 0$.

Por lo anterior, las nuevas probabilidades de transición para llegar a los estados $i \geq 0$, están dadas por: $p_{00}^*(a) = 1 - b\lambda$, $p_{ii+1}^*(a) = b\lambda$, $p_{ii}^*(a) = 1 - b(\lambda + a)$, $p_{i-1}^*(a) = b\lambda$, $i \geq 1$ donde la matriz de transición es:

$$P^* = \begin{pmatrix} 1 - b\lambda & b\lambda & 0 & \cdots \\ ba & 1 - b(\lambda + a) & b\lambda & \cdots \\ 0 & ba & 1 - b(\lambda + a) & \cdots \\ 0 & 0 & ba & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Con lo anterior, se han encontrado las probabilidades de transición de un estado a otro y también las probabilidades de transición hacia el mismo estado de la cadena de decisión de Markov a tiempo continuo.

Observe que los costos para la CDMTC y para la CDMTD

son los mismos, los cuales son

$$C(i, a) = G(i, a)\nu(i, a) + g(i, a). \quad (2.51)$$

El conjunto de políticas para la CDMTD Δ y para la CDMTC Ψ no son idénticas. Ya que, una política para Ψ solo puede seleccionar una nueva acción cuando una transición ocurre en el estado. Una política para Δ puede elegir una nueva acción en cada periodo de tiempo, incluso si permanece en el mismo estado. Además, el conjunto de políticas estacionarias es idéntico.

Por lo anterior, se demuestra el siguiente lema.

LEMA 2.4.3. *Sean Ψ una CDMTC la cual satisface la hipótesis 2.4.1, Δ una CDM auxiliar y e una política estacionaria.*

a) *Suponga que existe una constante Z y una función w tal que:*

$$Z + w(i) \geq C(i, e) + \sum_j p_{ij}^*(e)w(j), \quad (2.52)$$

$i \in S$. Entonces Z y $z = bw$ satisfacen (2.47).

b) *Suponga que existe una constante finita Z y $w = z/b$ que satisfacen (2.47). Entonces Z y $w = z/b$ satisfacen (2.52).*

Demostración. Sea Ψ una CDMTC que satisface la hipótesis 2.4.1, Δ una CDM auxiliar y e una política estacionaria.

a). Para la demostración de este inciso, se tiene que existen una constante Z y una función z , tales que

$$Z + w(i) \geq C(i, e) + \sum_j p_{ij}^*(e)w(j),$$

entonces sustituyendo $w(i) = z(i)/\tau(i, e)$, y multiplicando por $\tau(i, e)$ se tiene:

$$\begin{aligned} Z + \frac{z(i)}{\tau(i, e)} &\geq C(i, e) + \sum_j p_{ij}^*(e) \frac{z(j)}{\tau(j, e)} \\ Z\tau(i, e) + z(i) &\geq C(i, e)\tau(i, e) + \sum_j p_{ij}^*(e)z(j), \end{aligned}$$

sustituyendo (2.50) y (2.51) y además, recordando que $\tau(i, e) = 1/\nu(i, e)$ y que $b\nu(i, e) < 1$ se tiene

$$\begin{aligned} Z\tau(i, e) + z(i) &\geq [G(i, e)\nu(i, e) + g(i, e)]\tau(i, e) \\ &\quad + \sum_j b\nu(i, e)p_{ij}(e)z(j) \end{aligned}$$

$$Z\tau(i, e) + z(i) \geq G(i, e) + g(i, e)\tau(i, e) + \sum_j p_{ij}(e)z(j).$$

De la última ecuación se concluye que Z y $z = bw$ satisfacen (2.47).

- b). Para demostrar la segunda parte se tiene que existen una constante Z y una función z tales que

$$Z\tau(i, e) + z(i) \geq G(i, e) + g(i, e)\tau(i, e) + \sum_j p_{ij}(e)z(j),$$

es válida y además, sustituyendo $z(i) = w(i)\tau(i, e)$ en dicha ecuación y multiplicando por $\frac{1}{\tau(i, e)} = \nu(i, e)$ se tiene que:

$$\begin{aligned} Z\tau(i, e) + w(i)\tau(i, e) &\geq G(i, e) + g(i, e)\tau(i, e) \\ &\quad + \sum_j p_{ij}(e)w(j)\tau(j, e) \\ Z + w(i) &\geq G(i, e)\nu(i, e) + g(i, e) \\ &\quad + \sum_j b\nu(i, e)p_{ij}(e)w(j), \end{aligned}$$

la última desigualdad es debido a que $b\nu(i, e) < 1$. Por lo tanto, utilizando (2.50) y (2.51) se obtiene (2.52).

□

Ahora, se proporcionan dos hipótesis que relacionan al costo promedio mínimo en Ψ y en Δ . Suponga que se tiene una sucesión de aproximación $\{\Delta_N\}_{N \geq N_0}$ para Δ (ver Sección 2.2.1).

HIPÓTESIS 2.4.4. *Se supone que $J^\Delta(\cdot) \leq J^\Psi(\cdot)$, donde $J^\Delta(\cdot)$ es el costo promedio mínimo en Δ .*

HIPÓTESIS 2.4.5.

(AC1) *Suponga que existen una constante J^N y una función r^N sobre S_N tales que*

$$J^N + r^N(i) = \min_a \left\{ C(i, a) + \sum_{j \in S_N} p_{ij}^*(a; N) r^N(j) \right\}, \quad (2.53)$$

$$i \in S_N, N \geq N_0.$$

(AC2) *Se supone que $\limsup_{N \rightarrow \infty} r^N(i) < \infty$ para $i \in S$.*

(AC3) *Existe una constante no negativa Q tal que $-Q \leq \liminf_{N \rightarrow \infty} r^N(i) = w(i)$ para $i \in S$.*

(AC4) *Suponga que $\limsup_{N \rightarrow \infty} J^N =: J^* < \infty$ y $J^* \leq J(i)$ para $i \in S$.*

LEMA 2.4.6. *Suponga que la hipótesis (2.4.5) es válida. Entonces*

1. *La cantidad $J^* = \lim_{N \rightarrow \infty} J^N$ es el costo promedio mínimo en Δ .*

2. Cualquier punto límite e^* de una sucesión e^N de políticas estacionarias que alcanzan el mínimo en (2.53) es óptima de costo promedio para Δ .

El siguiente teorema proporciona el cálculo de una política estacionaria óptima de costo promedio para Ψ .

Teorema 4. Sean Ψ una CDMTC que satisface la hipótesis 2.4.1, Δ una CDM auxiliar tal que la hipótesis 2.4.4 es válida, y (Δ_N) una sucesión de aproximación para Δ que satisface las hipótesis 2.4.5. Se tiene

$$J^N + r^N(i) = \min_{a \in A(x)} \left\{ C(i, a) + \sum_{j \in S_N} p_{ij}^*(a; N) r^N(j) \right\}, \quad (2.54)$$

$i \in S_N, N \geq N_0$. Entonces:

i) La cantidad

$$J^* = \lim_{N \rightarrow \infty} J^N,$$

es el costo promedio mínimo en Δ y Ψ .

ii) Cualquier punto límite e^* de una sucesión e^N de políticas estacionarias alcanza el mínimo en (2.54), es óptima de costo promedio para Δ y Ψ .

Demostración. Ya que las hipótesis en (2.4.5) son válidas para Δ_N y Δ . Por el lema (2.4.6) se tiene que

$$J^* = \lim_{N \rightarrow \infty} J^N,$$

es el costo promedio mínimo en Δ y e^* es óptima de costo promedio en Δ .

Entonces de (2.54) se sigue que:

$$\begin{aligned} J^N + r^N(i) &= \min_{a \in A(x)} \left\{ c(i, a) + \sum_{j \in S_N} p_{ij}^*(a; N) r^N(j) \right\} \\ &= c(i, e^*) + \sum_{j \in S_N} P_{ij}^*(e^*; N) r^N(j), \end{aligned}$$

ya que e^* es óptima, tomando el límite inferior cuando n tiende a infinito se obtiene

$$\liminf_{N \rightarrow \infty} [J^N + r^N(i)] = \liminf_{N \rightarrow \infty} \left[c(i, e^*) + \sum_{j \in S_N} p_{ij}^*(e^*; N) r^N(j) \right]$$

$$\begin{aligned} \liminf_{N \rightarrow \infty} J^N + \liminf_{N \rightarrow \infty} r^N(i) &= \liminf_{N \rightarrow \infty} c(i, e^*) \\ &\quad + \liminf_{N \rightarrow \infty} \sum_{j \in S_N} p_{ij}^*(e^*; N) r^N(j) \\ J^* + w(i) &= c(i, e^*) + \sum_{j \in S} p_{ij}(e^*) w(j). \end{aligned}$$

Ahora, como Ψ es una CDMTC que satisface la hipótesis 2.4.1 y Δ una CDM auxiliar y e^* una política estacionaria óptima, por el lema 2.4.2 se tiene que

$$Z = J^* \quad \text{y} \quad z = \tau w,$$

satisfacen la ecuación (2.47), es decir,

$$J^* \tau(i, e^*) + z(i) \geq G(i, e^*) + g(i, e^*) \tau(i, e^*) + \sum_{j \in S} p_{ij}(e^*) z(j).$$

Por la hipótesis (AC3)

$$-Q \leq \liminf_{N \rightarrow \infty} r^N(i) = w(i),$$

es decir, que w esta acotada inferiormente por $-Q$, y como $z = \tau w$ entonces z también lo está.

Por el Lema 2.4.2 se tiene que

$$J_{e^*}^\Psi(\cdot) \leq J^*,$$

donde J^* es el costo promedio mínimo en Δ , es decir, $J^* \equiv J^\Delta(\cdot)$, entonces

$$J_{e^*}^\Psi(\cdot) \leq J^* \equiv J^\Delta(\cdot) \leq J^\Psi(\cdot) \leq J_{e^*}^\Delta(\cdot),$$

la segunda desigualdad es por la hipótesis (2.4.4) y la última se obtiene porque e^* es la óptima. Por lo tanto, todos estos términos son iguales,

$$J_{e^*}^\Psi(\cdot) = J^* = J^\Delta(\cdot) = J^\Phi(\cdot).$$

Esto prueba que e^* es óptima de costo promedio para Ψ con costo promedio constante J^* , es decir, que el costo de la solución para Δ es la misma que la de la original Ψ . □

PROPOSICIÓN 2.4.7. *Supóngase que la acción a satisface $\lambda < a$, sea $d(a)$ la política que siempre sirve con la razón a . Entonces, teniendo a $\rho_a = \lambda/a$, se tiene que*

$$J_{d(a)}^\Psi = \rho_a c(a) + \frac{H \rho_a}{1 - \rho_a} \quad (2.55)$$

y además,

$$J_d^\Psi = \min_{a > \lambda} J_{d(a)}^\Psi \quad (2.56)$$

Capítulo 3

Control de las Tasas de Servicio en una Línea de Espera

La teoría de líneas de espera proporciona buenos modelos para problemas interesantes de la vida real. Sin embargo, el interés no sólo es modelar la cola sino controlarla. Para este fin se utiliza la teoría de control estocástico en el caso de espacio de estados y acciones discretos, este contexto se conoce como cadenas de decisión de Markov o cadenas de control de Markov. Además, el método de programación dinámica proporciona la solución al problema de interés.

La teoría de los sistemas de control estocástico se utilizará para el estudio de sistemas de control de colas en tiempo discreto y continuo. En un sistema de líneas de espera los clientes solicitan el servicio y los servidores remueven estos clientes desde el sistema. **Un sistema de colas incluye servidores, clientes y usualmente líneas de espera o colas para que los clientes esperen un servicio. Los servidores pueden ser: transmisiones, computadoras, líneas de comunicación, estaciones o una línea de producción, etc. Los clientes pueden ser personas, mensajes, grupos de**

longitud fija de bits, conocidos como paquetes u objetos existentes en los procesos de manufactura, etc.

En este capítulo se plantea y se resuelve un problema de una línea de espera controlada con un solo servidor, con la teoría mostrada anteriormente.

3.1. Ejemplo

EJEMPLO 3.1.1. *Un pequeño taller mecánico solo tiene capacidad para mantener y atender a 2 autos, a estos se les presta el servicio de manera ordenada a como llegaron, por el mecánico encargado del taller el cual atiende con dos tasas de servicio una de 0.4 y otra de 0.7. Por utilizar cualquiera de las dos tasas de servicio se genera un costo de 1 y 2 unidades respectivamente. Además, con el solo hecho de aceptar algún auto se esta generando un costo de mantenimiento dentro del taller de una unidad por auto, se sabe también que un nuevo cliente arriba con probabilidad $p = 0.5$ en cualquier etapa. Además, si un cliente llega y encuentra el taller lleno, este se retira, o si éste está vacío puede entrar al servicio inmediatamente.*

Las probabilidades de transición están determinadas de la siguiente manera:

$$p_{01} = p_{21} = 1, p_{10} = \frac{a(1-p)}{p+a-2ap}, p_{12} = \frac{p(1-a)}{p+a-2ap}.$$

3 Con los datos que se proporcionaron anteriormente se puede modelar el problema de control en una cadena de decisión de Markov de la siguiente manera. El modelo de control Marko-

viano es

$$(S, A, \{A(i) | i \in S\}, P, C),$$

donde,

- $S = \{0, 1, 2\}$ es el espacio de estados.
- $A = A(i) = \{a_1, a_2\} = \{0.4, 0.7\}$ es el espacio de acciones (que en este ejemplo son las razones de servicio) donde $i \in S$.
- La matriz de probabilidades de transición con $p = 0.5$ es la siguiente:

$$P = [p_{ij}(a)] = \begin{pmatrix} 0 & 1 & 0 \\ \frac{a(1-p)}{p+a-2ap} & 0 & \frac{p(1-a)}{p+a-2ap} \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ a & 0 & 1-a \\ 0 & 1 & 0 \end{pmatrix}.$$

- Los costos que se generan por elegir cada una de las acciones se obtienen como sigue:

$$C(i, a) = i + c(a),$$

con $c(a_1) = 1$ y $c(a_2) = 2$. Entonces por tomar la acción a_1 los costos son:

$$C(0, a_1) = 0 + 1 = 1,$$

$$C(1, a_1) = 1 + 1 = 2,$$

$$C(2, a_1) = 2 + 1 = 3$$

y los costos que se generan por tomar la acción a_2 son los siguientes:

$$C(0, a_2) = 0 + 2 = 2,$$

$$C(1, a_2) = 1 + 2 = 3,$$

$$C(2, a_2) = 2 + 2 = 4.$$

Para resolver este ejemplo a tiempo continuo se aplica el método de Uniformización (ver Sección 2.5.1), es decir, el problema original se lleva a un problema a tiempo discreto. La probabilidad de transición propuesta en (2.50) es

$$p_{ij}^*(a) = \begin{cases} \nu(i, a)p_{ij}(a), & j \neq i, \\ 1 - \nu(i, a), & j = i, \end{cases}$$

con $b = 1$ y el tiempo de permanencia en cada uno de los estados está determinado por: $\nu(0, a) = \frac{1-a}{2}$, $\nu(1, a) = \frac{1}{2}$, $\nu(2, a) = \frac{a}{2}$.

Entonces, la nueva matriz de probabilidades de transición para llegar a los estados 0, 1 y 2 es

$$P^* = [p_{ij}^*(a)] = \begin{pmatrix} \frac{1+a}{2} & \frac{1-a}{2} & 0 \\ \frac{a}{2} & \frac{1}{2} & \frac{1-a}{2} \\ 0 & \frac{a}{2} & 1 - \frac{a}{2} \end{pmatrix}.$$

Entonces, las probabilidades de transición, eligiendo la acción $a_1 = 0.4$ (es decir, la razón de servicio más lenta), están dadas por:

$$P^* = [p_{ij}^*(a_1)] = \begin{pmatrix} 0.7 & 0.3 & 0 \\ 0.2 & 0.5 & 0.3 \\ 0 & 0.2 & 0.8 \end{pmatrix}$$

y eligiendo la acción $a_2 = 0.7$ (el servicio más rápido) están dadas por:

$$P^* = [p_{ij}^*(a_2)] = \begin{pmatrix} 0.85 & 0.15 & 0 \\ 0.35 & 0.5 & 0.15 \\ 0 & 0.35 & 0.65 \end{pmatrix}.$$

En este ejemplo se utiliza el criterio de costo promedio dado

por:

$$J_{\pi}(i) := \lim_{n \rightarrow \infty} \frac{1}{n} E_{\pi}^i \left[\sum_{t=0}^{n-1} C(x_t, a_t) \right].$$

La función de valor está dada por:

$$J^*(i) = \min_{\pi \in \Pi} J_{\pi}(i).$$

El problema de control es determinar una política óptima π^* tal que

$$J^*(i) = J_{\pi^*}(i),$$

utilizando el algoritmo de iteración de valores (AIV) dado en la sección 2.3, de la siguiente manera:

1). Sean $n = 0$ y $u_0 \equiv 0$.

2). Sea

$$w_n(i) = \min_{a \in A(i)} \left\{ C(i, a) + \sum_{j \in S} p_{ij}^*(a) u_n(j) \right\}. \quad (3.1)$$

3). Si $n = 0$, entonces $\delta = 1$.

Si $n \geq 1$, entonces $\delta = \max_{i \in S} |w_n(i) - w_{n-1}(i)|$.

Si $\delta < \epsilon$, se continua con el paso (6).

4). Sea $u_{n+1}(i) = w_n(i) - w_n(x)$.

5). Ir a (2) y reemplazar n por $n + 1$.

6). Parar en una etapa n y escribir la función de valor óptimo $w_n(x)$ y la política óptima que alcanza el mínimo en la ecuación (3.2).

Tomando a

$$x = 0$$

como el estado distinguido, del punto (4) se puede ver que

$$\begin{aligned} u_{n+1}(i) &= w_n(i) - w_n(0) \\ u_{n+1}(0) &= w_n(0) - w_n(0) = 0. \end{aligned}$$

Entonces,

$$u_n(0) \equiv 0, \text{ para cualquier } n.$$

Evaluando

$$w_n(i) = \min_{a \in A(i)} \left\{ C(i, a) + \sum_{j=0}^2 p_{ij}^*(a) u_n(j) \right\},$$

para cada uno de los estados.

Para $i = 0$,

$$\begin{aligned} w_n(0) &= \min_{a \in A(0)} \left\{ C(0, a) + \sum_{j=0}^2 p_{0j}^*(a) u_n(j) \right\} \\ &= \min_{a \in A(0)} \{ C(0, a) + p_{00}^*(a) \underline{u_n(0)} + p_{01}^*(a) u_n(1) \\ &\quad + p_{02}^*(a) u_n(2) \} \\ &= \min_{a \in A(0)} \{ C(0, a) + p_{01}^*(a) u_n(1) \} \\ &= \min_{a \in A(0)} \{ C(0, a_1) + p_{01}^*(a_1) u_n(1), C(0, a_2) \\ &\quad + p_{01}^*(a_2) u_n(1) \} \\ &= \min_{a \in A(0)} \{ 1 + 0.3u_n(1), 2 + 0.15u_n(1) \} \\ &= 1 + \min_{a \in A(0)} \{ 0.3u_n(1), 1 + 0.15u_n(1) \}. \end{aligned}$$

Para $i = 1$,

$$\begin{aligned}
 w_n(1) &= \min_{a \in A(1)} \left\{ C(1, a) + \sum_{j=0}^2 p_{1j}^*(a) u_n(j) \right\} \\
 &= \min_{a \in A(1)} \{ C(1, a) + p_{10}^*(a) \underline{u_n(0)} + p_{11}^*(a) u_n(1) \\
 &\quad + p_{12}^*(a) u_n(2) \} \\
 &= \min_{a \in A(1)} \{ C(1, a) + p_{11}^*(a) u_n(1) + p_{12}^*(a) u_n(2) \} \\
 &= \min_{a \in A(1)} \{ C(1, a_1) + p_{11}^*(a_1) u_n(1) + p_{12}^*(a_1) u_n(2), \\
 &\quad C(1, a_2) + p_{11}^*(a_2) u_n(1) + p_{12}^*(a_2) u_n(2) \} \\
 &= \min_{a \in A(1)} \{ 2 + 0.5u_n(1) + 0.3u_n(2), 3 + 0.5u_n(1) + 0.15u_n(2) \} \\
 &= 2 + 0.5u_n(1) + \min_{a \in A(1)} \{ 0.3u_n(2), 1 + 0.15u_n(2) \}.
 \end{aligned}$$

Para $i = 2$,

$$\begin{aligned}
 w_n(2) &= \min_{a \in A(2)} \left\{ C(2, a) + \sum_{j=0}^2 p_{2j}^*(a) u_n(j) \right\} \\
 &= \min_{a \in A(2)} \{ C(2, a) + \underline{p_{20}^*(a) u_n(0)} + p_{21}^*(a) u_n(1) + p_{22}^*(a) u_n(2) \} \\
 &= \min_{a \in A(2)} \{ C(2, a) + p_{21}^*(a) u_n(1) + p_{22}^*(a) u_n(2) \} \\
 &= \min_{a \in A(2)} \{ C(2, a_1) + p_{21}^*(a_1) u_n(1) + p_{22}^*(a_1) u_n(2), \\
 &\quad C(2, a_2) + p_{21}^*(a_2) u_n(1) + p_{22}^*(a_2) u_n(2) \} \\
 &= \min_{a \in A(2)} \{ 3 + 0.2u_n(1) + 0.8u_n(2), 4 + 0.35u_n(1) + 0.65u_n(2) \} \\
 &= 3 + \min_{a \in A(2)} \{ 0.2u_n(1) + 0.8u_n(2), 1 + 0.35u_n(1) + 0.65u_n(2) \}.
 \end{aligned}$$

De los cálculos anteriores se obtuvo lo siguiente:

$$w_n(0) = 1 + \min_{a \in A(0)} \{ 0.3u_n(1), 1 + 0.15u_n(1) \}.$$

$$\begin{aligned}
w_n(1) &= 2 + 0.5u_n(1) + \min_{a \in A(1)} \{0.3u_n(2), 1 + 0.15u_n(2)\}. \\
w_n(2) &= 3 + \min_{a \in A(2)} \{0.2u_n(1) + 0.8u_n(2), \\
&\quad 1 + 0.35u_n(1) + 0.65u_n(2)\}.
\end{aligned}$$

Ahora, utilizando el método de iteración de valores dado anteriormente.

Para $n = 0$, $u_0(i) \equiv 0$, para toda $i \in S$. Entonces

$$\begin{aligned}
w_0(i) &= \min_{a \in A(i)} \left\{ C(i, a) + \sum_{j=0}^2 p_{ij}^*(a) u_0(j) \right\}. \\
w_0(0) &= 1 + \min_{a \in A(0)} \{0.3u_0(1), 1 + 0.15u_0(1)\} \\
&= 1 + \min_{a \in A(0)} \{0, 1\} = 1 + 0 = 1. \\
w_0(1) &= 2 + 0.5u_0(1) + \min_{a \in A(1)} \{0.3u_0(2), 1 + 0.15u_0(2)\} \\
&= 2 + \min_{a \in A(1)} \{0, 1\} = 2 + 0 = 2. \\
w_0(2) &= 3 + \min_{a \in A(2)} \{0.2u_0(1) + 0.8u_0(2), \\
&\quad 1 + 0.35u_0(1) + 0.65u_0(2)\} \\
&= 3 + \min_{a \in A(2)} \{0, 1\} = 3 + 0 = 3.
\end{aligned}$$

Además, $\delta = 1$. Y

$$\begin{aligned}
u_{0+1}(i) &= w_0(i) - w_0(0) \\
u_1(i) &= w_0(i) - w_0(0), \\
u_1(0) &= w_0(0) - w_0(0) = 0, \\
u_1(1) &= w_0(1) - w_0(0) = 2 - 1 = 1, \\
u_1(2) &= w_0(2) - w_0(0) = 3 - 1 = 2.
\end{aligned}$$

$$\text{Para } n = 1, \begin{cases} u_1(0) = 0, \\ u_1(1) = 1, \\ u_1(2) = 2, \end{cases} \quad \text{entonces}$$

$$w_1(i) = \min_{a \in A(i)} \left\{ C(i, a) + \sum_{j=0}^2 p_{ij}^*(a) u_1(j) \right\}.$$

$$\begin{aligned} w_1(0) &= 1 + \min_{a \in A(0)} \{0.3u_1(1), 1 + 0.15u_1(1)\} \\ &= 1 + \min_{a \in A(0)} \{0.3(1), 1 + 0.15(1)\} \\ &= 1 + \min_{a \in A(0)} \{0.3, 1 + 0.15\} \\ &= 1 + \min_{a \in A(0)} \{0.3, 1.15\} = 1 + 0.3 = 1.3. \end{aligned}$$

$$\begin{aligned} w_1(1) &= 2 + 0.5u_1(1) + \min_{a \in A(1)} \{0.3u_1(2), 1 + 0.15u_1(2)\} \\ &= 2 + 0.5(1) + \min_{a \in A(1)} \{0.3(2), 1 + 0.15(2)\} \\ &= 2.5 + \min_{a \in A(1)} \{0.6, 1 + 0.3\} \\ &= 2.5 + \min_{a \in A(1)} \{0.6, 1.3\} \\ &= 2.5 + 0.6 = 3.1. \end{aligned}$$

$$\begin{aligned} w_1(2) &= 3 + \min_{a \in A(2)} \{0.2u_1(1) + 0.8u_1(2), \\ &\quad 1 + 0.35u_1(1) + 0.65u_1(2)\} \\ &= 3 + \min_{a \in A(2)} \{0.2(1) + 0.8(2), 1 + 0.35(1) + 0.65(2)\} \\ &= 3 + \min_{a \in A(2)} \{0.2 + 1.6, 1 + 0.35 + 1.3\} \\ &= 3 + \min_{a \in A(2)} \{1.8, 2.65\} \\ &= 3 + 1.8 = 4.8. \end{aligned}$$

Además,

$$\delta = \max_{i \in S} |w_1(i) - w_0(i)| = \max_{i \in S} \{0.3, 1.1, 1.8\} = 1.8.$$

donde,

$$\begin{aligned} |w_1(0) - w_0(0)| &= |1.3 - 1| = |0.3| = 0.3, \\ |w_1(1) - w_0(1)| &= |3.1 - 2| = |1.1| = 1.1, \\ |w_1(2) - w_0(2)| &= |4.8 - 3| = |1.8| = 1.8. \end{aligned}$$

Además,

$$\begin{aligned} u_{1+1}(i) &= w_1(i) - w_1(0) \\ u_2(i) &= w_1(i) - w_1(0), \end{aligned}$$

$$\begin{aligned} u_2(0) &= w_1(0) - w_1(0) = 0, \\ u_2(1) &= w_1(1) - w_1(0) = 3.1 - 1.3 = 1.8, \\ u_2(2) &= w_1(2) - w_1(0) = 4.8 - 1.3 = 3.5. \end{aligned}$$

$$\text{Para } n = 2, \begin{cases} u_2(0) = 0, \\ u_2(1) = 1.8, \\ u_2(2) = 3.5. \end{cases} \quad \text{Entonces}$$

$$w_2(i) = \min_{a \in A(i)} \left\{ C(i, a) + \sum_{j=0}^2 p_{ij}^*(a) u_2(j) \right\}.$$

$$\begin{aligned} w_2(0) &= 1 + \min_{a \in A(0)} \{0.3u_2(1), 1 + 0.15u_2(1)\} \\ &= 1 + \min_{a \in A(0)} \{0.3(1.8), 1 + 0.15(1.8)\} \\ &= 1 + \min_{a \in A(0)} \{0.54, 1.27\} \end{aligned}$$

$$\begin{aligned}
&= 1 + 0.54 = 1.54. \\
w_2(1) &= 2 + 0.5u_2(1) + \min_{a \in A(1)} \{0.3u_2(2), 1 + 0.15u_2(2)\} \\
&= 2 + 0.5(1.8) + \min_{a \in A(1)} \{0.3(3.5), 1 + 0.15(3.5)\} \\
&= 2.9 + \min_{a \in A(1)} \{1.05, 1.525\} \\
&= 2.9 + 1.05 = 3.95. \\
w_2(2) &= 3 + \min_{a \in A(2)} \{0.2u_2(1) + 0.8u_2(2), \\
&\quad 1 + 0.35u_2(1) + 0.65u_2(2)\} \\
&= 3 + \min_{a \in A(2)} \{0.2(1.8) + 0.8(3.5), 1 + 0.35(1.8) + 0.65(3.5)\} \\
&= 3 + \min_{a \in A(2)} \{3.16, 3.905\} \\
&= 3 + 3.16 = 6.16.
\end{aligned}$$

Además,

$$\delta = \max_{i \in S} |w_2(i) - w_1(i)| = \max_{i \in S} \{0.24, 0.85, 1.36\} = 1.36,$$

donde,

$$\begin{aligned}
|w_2(0) - w_1(0)| &= |1.54 - 1.3| = |0.24| = 0.24, \\
|w_2(1) - w_1(1)| &= |3.95 - 3.1| = |0.85| = 0.85, \\
|w_2(2) - w_1(2)| &= |6.16 - 4.8| = |1.36| = 1.36.
\end{aligned}$$

$$\begin{aligned}
u_{2+1}(i) &= w_2(i) - w_2(0) \\
u_3(i) &= w_2(i) - w_2(0),
\end{aligned}$$

$$\begin{aligned}
u_3(0) &= w_2(0) - w_2(0) = 0, \\
u_3(1) &= w_2(1) - w_2(0) = 3.95 - 1.54 = 2.41, \\
u_3(2) &= w_2(2) - w_2(0) = 6.16 - 1.54 = 4.62.
\end{aligned}$$

$$\text{Para } n = 3, \begin{cases} u_3(0) = 0, \\ u_3(1) = 2.41, \\ u_3(2) = 4.62. \end{cases} \quad \text{Entonces}$$

$$\begin{aligned} w_3(0) &= 1 + \min_{a \in A(0)} \{0.3u_3(1), 1 + 0.15u_3(1)\} \\ &= 1 + \min_{a \in A(0)} \{0.3(2.41), 1 + 0.15(2.41)\} \\ &= 1 + \min_{a \in A(0)} \{0.723, 1.3615\} \\ &= 1 + 0.723 = 1.723. \end{aligned}$$

$$\begin{aligned} w_3(1) &= 2 + 0.5u_3(1) + \min_{a \in A(1)} \{0.3u_3(2), 1 + 0.15u_3(2)\} \\ &= 2 + 0.5(2.41) + \min_{a \in A(1)} \{0.3(4.62), 1 + 0.15(4.62)\} \\ &= 3.205 + \min_{a \in A(1)} \{1.386, 1.693\} \\ &= 3.205 + 1.386 = 4.591. \end{aligned}$$

$$\begin{aligned} w_3(2) &= 3 + \min_{a \in A(2)} \{0.2u_3(1) + 0.8u_3(2), \\ &\quad 1 + 0.35u_3(1) + 0.65u_3(2)\} \\ &= 3 + \min_{a \in A(2)} \{0.2(2.41) + 0.8(4.62), 1 + 0.35(2.41) \\ &\quad + 0.65(4.62)\} \\ &= 3 + \min_{a \in A(2)} \{4.178, 4.8465\} \\ &= 3 + 4.178 = 7.178. \end{aligned}$$

Además,

$$\delta = \max_{i \in S} |w_3(i) - w_2(i)| = \max_{i \in S} \{0.183, 0.641, 1.018\} = 1.018,$$

donde,

$$\begin{aligned} |w_3(0) - w_2(0)| &= |1.723 - 1.54| = |0.183| = 0.183, \\ |w_3(1) - w_2(1)| &= |4.591 - 3.95| = |0.641| = 0.641, \\ |w_3(2) - w_2(2)| &= |7.178 - 6.16| = |1.018| = 1.018. \end{aligned}$$

$$\begin{aligned}u_{3+1}(i) &= w_3(i) - w_3(0) \\ u_4(i) &= w_3(i) - w_3(0),\end{aligned}$$

$$\begin{aligned}u_4(0) &= w_3(0) - w_3(0) = 0, \\ u_4(1) &= w_3(1) - w_3(0) = 4.591 - 1.723 = 2.868, \\ u_4(2) &= w_3(2) - w_3(0) = 7.178 - 1.723 = 4.455.\end{aligned}$$

$$\text{Para } n = 4, \begin{cases} u_4(0) = 0, \\ u_4(1) = 2.868, \\ u_4(2) = 4.455. \end{cases} \quad \text{Entonces}$$

$$\begin{aligned}w_4(0) &= 1 + \min_{a \in A(0)} \{0.3u_4(1), 1 + 0.15u_4(1)\} \\ &= 1 + \min_{a \in A(0)} \{0.3(2.868), 1 + 0.15(2.868)\} \\ &= 1 + \min_{a \in A(0)} \{0.8604, 1.4302\} \\ &= 1 + 0.8604 = 1.8604. \\ w_4(1) &= 2 + 0.5u_4(1) + \min_{a \in A(1)} \{0.3u_4(2), 1 + 0.15u_4(2)\} \\ &= 2 + 0.5(2.868) + \min_{a \in A(1)} \{0.3(5.455), 1 + 0.15(5.455)\} \\ &= 3.434 + \min_{a \in A(1)} \{1.6365, 1.81825\} \\ &= 3.434 + 1.6365 = 5.0705. \\ w_4(2) &= 3 + \min_{a \in A(2)} \{0.2u_4(1) + 0.8u_4(2), \\ &\quad 1 + 0.35u_4(1) + 0.65u_4(2)\} \\ &= 3 + \min_{a \in A(2)} \{0.2(2.868) + 0.8(5.455), 1 + 0.35(2.868) \\ &\quad + 0.65(5.455)\} \\ &= 3 + \min_{a \in A(2)} \{4.9376, 5.54955\} \\ &= 3 + 4.9376 = 7.9376.\end{aligned}$$

Además,

$$\delta = \max_{i \in S} |w_4(i) - w_3(i)| = \max_{i \in S} \{0.1374, 0.4795, 0.7596\} = 1.7596,$$

donde,

$$\begin{aligned} |w_4(0) - w_3(0)| &= |1.8604 - 1.723| = |0.1374| = 0.1374, \\ |w_4(1) - w_3(1)| &= |5.0705 - 4.591| = |0.4795| = 0.4795, \\ |w_4(2) - w_3(2)| &= |7.9376 - 7.178| = |0.7596| = 0.7596. \end{aligned}$$

$$\begin{aligned} u_{4+1}(i) &= w_4(i) - w_4(0) \\ u_5(i) &= w_4(i) - w_4(0), \end{aligned}$$

$$\begin{aligned} u_5(0) &= w_4(0) - w_4(0) = 0, \\ u_5(1) &= w_4(1) - w_4(0) = 4.0705 - 1.8604 = 3.2101, \\ u_5(2) &= w_4(2) - w_4(0) = 7.9376 - 1.8604 = 6.0772. \end{aligned}$$

$$\text{Para } n = 5, \begin{cases} u_5(0) = 0, \\ u_5(1) = 3.2101, \\ u_5(2) = 6.0772. \end{cases} \quad \text{Entonces}$$

$$\begin{aligned} w_5(0) &= 1 + \min_{a \in A(0)} \{0.3u_5(1), 1 + 0.15u_5(1)\} \\ &= 1 + \min_{a \in A(0)} \{0.3(3.2101), 1 + 0.15(3.2101)\} \\ &= 1 + \min_{a \in A(0)} \{0.96303, 1.4811515\} \\ &= 1 + 0.96303 \\ &= 1.96303. \end{aligned}$$

$$\begin{aligned}
w_5(1) &= 2 + 0.5u_5(1) + \min_{a \in A(1)} \{0.3u_5(2), 1 + 0.15u_5(2)\} \\
&= 2 + 0.5(3.2101) + \min_{a \in A(1)} \{0.3(6.0772), 1 + 0.15(6.0772)\} \\
&= 3.60505 + \min_{a \in A(1)} \{1.82316, 1.91158\} \\
&= 3.60505 + 1.82316 = 5.42821. \\
w_5(2) &= 3 + \min_{a \in A(2)} \{0.2u_5(1) + 0.8u_5(2), \\
&\quad 1 + 0.35u_5(1) + 0.65u_5(2)\} \\
&= 3 + \min_{a \in A(2)} \{0.2(3.2101) + 0.8(6.0772), 1 + 0.35(3.2101) \\
&\quad + 0.65(6.0772)\} \\
&= 3 + \min_{a \in A(2)} \{5.50378, 6.073715\} \\
&= 3 + 5.50378 = 8.50378.
\end{aligned}$$

Además,

$$\begin{aligned}
\delta &= \max_{i \in S} |w_5(i) - w_4(i)| = \max_{i \in S} \{0.10263, 0.35771, 0.56618\} \\
&= 0.56618,
\end{aligned}$$

donde,

$$\begin{aligned}
|w_5(0) - w_4(0)| &= |1.96303 - 1.8604| = |0.10263| = 0.10263, \\
|w_5(1) - w_4(1)| &= |5.42821 - 5.0705| = |0.35771| = 0.35771, \\
|w_5(2) - w_4(2)| &= |8.50378 - 7.9376| = |0.56618| = 0.56618.
\end{aligned}$$

$$\begin{aligned}
u_{5+1}(i) &= w_5(i) - w_5(0) \\
u_6(i) &= w_5(i) - w_5(0),
\end{aligned}$$

$$\begin{aligned}
u_6(0) &= w_5(0) - w_5(0) = 0, \\
u_6(1) &= w_5(1) - w_5(0) = 5.42821 - 1.96303 = 3.46518, \\
u_6(2) &= w_5(2) - w_5(0) = 8.50378 - 1.96303 = 6.54075.
\end{aligned}$$

$$\text{Para } n = 6, \begin{cases} u_6(0) = 0, \\ u_6(1) = 3.46518, \\ u_6(2) = 6.54075. \end{cases} \quad \text{Entonces}$$

$$\begin{aligned} w_6(0) &= 1 + \min_{a \in A(0)} \{0.3u_6(1), 1 + 0.15u_6(1)\} \\ &= 1 + \min_{a \in A(0)} \{0.3(3.46518), 1 + 0.15(3.46518)\} \\ &= 1 + \min_{a \in A(0)} \{1.039554, 1.519777\} \\ &= 1 + 1.039554 = 2.039554. \\ w_6(1) &= 2 + 0.5u_5(1) + \min_{a \in A(1)} \{0.3u_6(2), 1 + 0.15u_6(2)\} \\ &= 2 + 0.5(3.4651) + \min_{a \in A(1)} \{0.3(6.5407), 1 + 0.15(6.5407)\} \\ &= 3.7325 + \min_{a \in A(1)} \{1.9622, 1.9811\} \\ &= 3.7325 + 1.9622 = 5.6948. \\ w_6(2) &= 3 + \min_{a \in A(2)} \{0.2u_6(1) + 0.8u_6(2), \\ &\quad 1 + 0.35u_6(1) + 0.65u_6(2)\} \\ &= 3 + \min_{a \in A(2)} \{0.2(3.4652) + 0.8(6.5407), 1 + 0.35(3.4651) \\ &\quad + 0.65(6.5407)\} \\ &= 3 + \min_{a \in A(2)} \{5.9256, 6.4643\} \\ &= 3 + 5.9256 = 8.9256. \end{aligned}$$

Además,

$$\delta = \max_{i \in S} |w_6(i) - w_5(i)| = \max_{i \in S} \{0.0765, 0.2666, 0.4218\} = 0.4218,$$

donde,

$$\begin{aligned} |w_6(0) - w_5(0)| &= |2.0395 - 1.9630| = |0.0765| = 0.0765, \\ |w_6(1) - w_5(1)| &= |5.6948 - 5.4282| = |0.2666| = 0.2666, \\ |w_6(2) - w_5(2)| &= |8.9256 - 8.5037| = |0.4218| = 0.4218. \end{aligned}$$

La siguiente tabla muestra los resultados de la etapa 0 hasta la etapa 30 para cada estado, las políticas que se encontraron en este ejemplo fueron dos. La política $e1$ siempre elige la acción a_1 para cada estado y la política $e2$ elige la acción a_1 cuando el sistema está en los estados $\{0, 2\}$ y elige a_2 cuando el sistema está en el estado 1.

		S			S			
	0	1	2	0	1	2		
n		$u_n(i)$			$w_0(i)$		δ	f_n
0	0	0	0	1.000000	2.000000	3.000000	1	$e1$
1	0	1	2	1.300000	3.100000	4.800000	1.800	$e1$
2	0	1.800000	3.500000	1.540000	3.950000	6.160000	1.360	$e1$
3	0	2.410000	4.620000	1.723000	4.591000	7.178000	1.018	$e1$
4	0	2.868000	5.455000	1.860400	5.070500	7.937600	0.760	$e1$
5	0	3.210100	6.077200	1.963030	5.428210	8.503780	0.566	$e1$
6	0	3.465180	6.540750	2.039554	5.694815	8.925636	0.422	$e1$
7	0	3.655261	6.886082	2.096578	5.860543	9.239918	0.314	$e2$
8	0	3.763965	7.143340	2.129189	5.953483	9.467465	0.228	$e2$
9	0	3.824294	7.338275	2.147288	6.012888	9.635479	0.168	$e2$
10	0	3.865600	7.488191	2.159680	6.056029	9.763673	0.128	$e2$
11	0	3.896349	7.603993	2.168905	6.088773	9.862464	0.099	$e2$
12	0	3.919869	7.693559	2.175961	6.113968	9.938821	0.076	$e2$
13	0	3.938008	7.762861	2.181402	6.133433	9.997890	0.059	$e2$
14	0	3.952031	7.816488	2.185609	6.148488	10.043596	0.046	$e2$
15	0	3.962879	7.857987	2.188864	6.160138	10.078965	0.035	$e2$
16	0	3.971274	7.890102	2.191382	6.169152	10.106336	0.027	$e2$
17	0	3.977770	7.914954	2.193331	6.176128	10.127517	0.021	$e2$

		S			S			
	0	1	2	0	1	2		
n		$u_n(i)$			$w_0(i)$		δ	f_n
18	0	3.982797	7.934186	2.194839	6.181526	10.143908	0.016	e2
19	0	3.986687	7.949069	2.196006	6.185704	10.156593	0.013	e2
20	0	3.989698	7.960587	2.196909	6.188937	10.166409	0.010	e2
21	0	3.992028	7.969500	2.197608	6.191439	10.174005	0.008	e2
22	0	3.993830	7.976397	2.198149	6.193375	10.179884	0.006	e2
23	0	3.995226	7.981734	2.198568	6.194873	10.184433	0.005	e2
24	0	3.996305	7.985865	2.198892	6.196032	10.187953	0.004	e2
25	0	3.997141	7.989061	2.199142	6.196930	10.190677	0.003	e2
26	0	3.997787	7.991535	2.199336	6.197624	10.192786	0.002	e2
27	0	3.998288	7.993449	2.199486	6.198161	10.194417	0.002	e2
28	0	3.998675	7.994931	2.199602	6.198577	10.195680	0.001	e2
29	0	3.998975	7.996077	2.199692	6.198899	10.196657	0.001	e2
30	0	3.999206	7.996964	2.199762	6.199148	10.197413	0.001	e2
31	0	3.999386	7.997651	2.199816	6.199341	10.197998	0.001	e2

Por lo tanto, el sistema debe operar utilizando la razón de servicio más lenta si el sistema tiene 0 ó 2 clientes y operar utilizando la razón de servicio más rápida cuando hay 1 cliente en el sistema.

EJEMPLO 3.1.2. Una sala de espera es observada a tiempo continuo y puede almacenar un número infinito de clientes, ésta es atendida por un solo servidor el cual sólo atiende un cliente por etapa, éste cuenta con tres tasas de servicio una de 4, 5 y otra de 6.

Se genera un costo de 2, 5 y 5.5 unidades respectivamente de acuerdo a la tasa de servicio empleada, pero si el sistema está vacío no se emplea ninguna acción por lo que no se genera

ningún costo.

Las probabilidades de transición están determinadas de la siguiente manera:

$$p_{01} = p_{21} = 1, p_{10} = \frac{a(1-p)}{p+a-2ap}, p_{12} = \frac{p(1-a)}{p+a-2ap}.$$

Con los datos que se proporcionaron anteriormente se puede modelar el problema de control de la siguiente manera. El modelo de control Markoviano es

$$(S, A, \{A(i)|i \in S\}, P, C),$$

donde,

- $S = \{0, 1, 2, \dots\}$ es el espacio de estados.
- $A = A(i) = \{a_1, a_2, a_3\} = \{4, 5, 6\}$ es el espacio de acciones (que en este ejemplo son las razones de servicio) donde $i \in S$.
- La matriz de probabilidades de transición con $\lambda = 3$ es la siguiente:

$$P = [p_{ij}(a)] = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ \frac{a}{\lambda+a} & 0 & \frac{\lambda}{\lambda+a} & 0 & 0 & \dots \\ 0 & \frac{a}{\lambda+a} & 0 & \frac{\lambda}{\lambda+a} & 0 & \dots \\ 0 & 0 & \frac{a}{\lambda+a} & 0 & \frac{\lambda}{\lambda+a} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots \\ \frac{a}{3+a} & 0 & \frac{3}{3+a} & 0 & 0 & \dots \\ 0 & \frac{a}{3+a} & 0 & \frac{3}{3+a} & 0 & \dots \\ 0 & 0 & \frac{a}{3+a} & 0 & \frac{3}{3+a} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

- Los costos que se generan por elegir cada una de las acciones se obtienen como sigue:

$$C(i, a) = c(a),$$

con $c(a_1) = 2$, $c(a_2) = 5$ y $c(a_3) = 5.5$. Tomando encuentan que solo en el estado 0 no se elige ninguna acción, los costos que se generan por elegir cualquiera de las tres acciones para cualquier estado son:

$$C(0, a) = c(0) \equiv 0,$$

$$C(i, a_1) = 2,$$

$$C(i, a_2) = 5,$$

$$C(i, a_3) = 5.5.$$

Para resolver este ejemplo a tiempo continuo se aplica el método de Uniformización (ver Sección 2.5.1), es decir, el problema original se lleva a un problema a tiempo discreto. La probabilidad de transición propuesta en (2.50) es

$$p_{ij}^*(a) = \begin{cases} b\nu(i, a)p_{ij}(a), & j \neq i, \\ 1 - b\nu(i, a), & j = i, \end{cases}$$

con $b = \frac{1}{2(\lambda + a_3)} = \frac{1}{17}$ y el tiempo de permanencia en cada uno de los estados está determinado por: $\nu(0, a) = \nu(0) = \lambda = 3$ y $\nu(i, a) = \lambda + a = 3 + a$.

Entonces la nueva matriz de probabilidades de transición es

$$\begin{aligned}
 P^* &= \begin{pmatrix} 1-b\lambda & b\lambda & 0 & 0 & \cdots \\ ba & 1-b(\lambda+a) & b\lambda & 0 & \cdots \\ 0 & ba & 1-b(\lambda+a) & b\lambda & \cdots \\ 0 & 0 & ba & 1-b(\lambda+a) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \\
 &= \begin{pmatrix} \frac{14}{17} & \frac{3}{17} & 0 & 0 & \cdots \\ \frac{a}{17} & 1-\frac{3+a}{17} & \frac{3}{17} & 0 & \cdots \\ 0 & \frac{a}{17} & 1-\frac{3+a}{17} & \frac{3}{17} & \cdots \\ 0 & 0 & \frac{a}{17} & 1-\frac{3+a}{17} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.
 \end{aligned}$$

La matriz de probabilidad de transición anterior es para el espacio de estados infinito numerable y tiempo discreto. Utilizando el método de aproximaciones sucesivas para hacer el espacio de estados de numerable a finito, es decir, $S = \{0, 1, 2, \dots, N\}$ la matriz es como sigue

$$\begin{aligned}
 P^* &= \begin{pmatrix} 1-b\lambda & b\lambda & 0 & \cdots & 0 & 0 \\ ba & 1-b(\lambda+a) & b\lambda & \cdots & 0 & 0 \\ 0 & ba & 1-b(\lambda+a) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & ba & 1-ba \end{pmatrix} \\
 &= \begin{pmatrix} \frac{14}{17} & \frac{3}{17} & 0 & \cdots & 0 & 0 \\ \frac{a}{17} & 1-\frac{3+a}{17} & \frac{3}{17} & \cdots & 0 & 0 \\ 0 & \frac{a}{17} & 1-\frac{3+a}{17} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{a}{17} & 1-\frac{a}{17} \end{pmatrix}.
 \end{aligned}$$

Denotando a $P^*(a)$ como la matriz de probabilidades de transi-

ción por elegir cada una de entonces se tiene que:

$$\begin{aligned}
 P^*(a_1) &= \begin{pmatrix} \frac{14}{17} & \frac{3}{17} & 0 & \dots & 0 & 0 \\ \frac{4}{17} & \frac{10}{17} & \frac{3}{17} & \dots & 0 & 0 \\ 0 & \frac{4}{17} & \frac{10}{17} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \frac{4}{17} & \frac{13}{17} \end{pmatrix}, \\
 P^*(a_2) &= \begin{pmatrix} \frac{14}{17} & \frac{3}{17} & 0 & \dots & 0 & 0 \\ \frac{5}{17} & \frac{9}{17} & \frac{3}{17} & \dots & 0 & 0 \\ 0 & \frac{5}{17} & \frac{9}{17} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \frac{5}{17} & \frac{12}{17} \end{pmatrix}, \\
 P^*(a_3) &= \begin{pmatrix} \frac{14}{17} & \frac{3}{17} & 0 & \dots & 0 & 0 \\ \frac{6}{17} & \frac{8}{17} & \frac{3}{17} & \dots & 0 & 0 \\ 0 & \frac{6}{17} & \frac{8}{17} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \frac{6}{17} & \frac{11}{17} \end{pmatrix}.
 \end{aligned}$$

Entonces el problema a tiempo continuo y con espacio de estados numerable se ha llevado a un problema a tiempo discreto con espacio de estados finitos.

En este ejemplo se utiliza el criterio de costo promedio dado en la siguiente ecuación

$$J_{\pi}(i) := \lim_{n \rightarrow \infty} \frac{1}{n} E_{\pi}^x \sum_{t=0}^{n-1} C(x_t, a_t).$$

El problema de control es

$$J^*(i) = \min_{\pi \in \Pi} J_{\pi}(i).$$

Utilizando el algoritmo de iteración de valores (AIV) dado en la sección 2.3 de la siguiente manera:

1). Sea $n = 0$ y $u_0 \equiv 0$.

2). Sea

$$w_n(i) = \min_{a \in A(i)} \left\{ C(i, a) + \sum_{j \in S} p_{ij}^*(a) u_n(j) \right\}. \quad (3.2)$$

3). Si $n = 0$, entonces $\delta = 1$.

Si $n \geq 1$, entonces $\delta = \max_{i \in S} |w_n(i) - w_{n-1}(i)|$.

Si $\delta < \epsilon$, se continua con el paso (6).

4). Sea $u_{n+1}(i) = w_n(i) - w_n(x)$.

5). Ir a (2) y reemplazar n por $n + 1$.

6). Parar en una etapa n y escribir la función de valor óptimo $w_n(x)$ y la política óptima que alcanza el mínimo en la ecuación (3.2).

Tomando a $x = 0$ como el estado distinguido, del punto (4) se puede ver que

$$\begin{aligned} u_{n+1}(i) &= w_n(i) - w_n(0) \\ u_{n+1}(0) &= w_n(0) - w_n(0) = 0. \end{aligned}$$

Entonces, $u_n(0) \equiv 0$, para cualquier n .

Calculando primero

$$w_n(i) = \min_{a \in A(i)} \left\{ C(i, a) + \sum_{j=0}^N p_{ij}^*(a) u_n(j) \right\},$$

para cada uno de los estados.

Para $i = 0$,

$$\begin{aligned}
 w_n(0) &= \min_{a \in A(0)} \left\{ C(0, a) + \sum_{j=0}^N p_{0j}^*(a) u_n(j) \right\} \\
 &= \min_{a \in A(0)} \left\{ c(0) + \sum_{j=0}^1 p_{0j}^*(a) u_n(j) \right\} \\
 &= \min_{a \in A(0)} \{ 0 + p_{00}^*(a) u_n(0) + p_{01}^*(a) u_n(1) \} \\
 &= \min_{a \in A(0)} \left\{ \frac{3}{17} u_n(1) \right\} \\
 &= \frac{3}{17} u_n(1).
 \end{aligned}$$

Para $1 \leq i \leq N - 1$,

$$\begin{aligned}
 w_n(i) &= \min_{a \in A(i)} \left\{ C(i, a) + \sum_{j=0}^N p_{ij}^*(a) u_n(j) \right\} \\
 &= \min_{a \in A(i)} \left\{ c(a) + \sum_{j=i-1}^{i+1} p_{ij}^*(a) u_n(j) \right\} \\
 &= \min_{a \in A(i)} \{ c(a) + p_{ii-1}^*(a) u_n(i-1) + p_{ii}^*(a) u_n(i) \\
 &\quad + p_{ii+1}^*(a) u_n(i+1) \}.
 \end{aligned}$$

Para $i = N$,

$$\begin{aligned}
 w_n(N) &= \min_{a \in A(N)} \left\{ C(i, a) + \sum_{j=0}^N p_{ij}^*(a) u_n(j) \right\} \\
 &= \min_{a \in A(N)} \left\{ c(a) + \sum_{j=N-1}^N p_{ij}^*(a) u_n(j) \right\} \\
 &= \min_{a \in A(N)} \{ c(a) + p_{NN-1}^*(a) u_n(N-1) + p_{NN}^*(a) u_n(N) \}
 \end{aligned}$$

De los cálculos anteriores se obtuvo lo siguiente:

$$\begin{aligned}
 w_n(0) &= \frac{3}{17} u_n(1). \\
 w_n(i) &= \min_{a \in A(i)} \{ c(a) + p_{i,i-1}^*(a) u_n(i-1) + p_{ii}^*(a) u_n(i) \\
 &\quad + p_{ii+1}^*(a) u_n(i+1) \}, \text{ para } 1 \leq i \leq N-1. \\
 w_n(N) &= \min_{a \in A(N)} \{ c(a) + p_{NN-1}^*(a) u_n(N-1) + p_{NN}^*(a) u_n(N) \}
 \end{aligned}$$

Además, utilizando la ecuación (2.56) de la proposición 2.4.7, con $H = 1$ y $\rho_a = \frac{\lambda}{a} = \frac{3}{a}$ se tiene que

$$\begin{aligned}
 J_d^\Psi &= \min_{a > \lambda} \{ J_{d(a)}^\Psi \} = \min_{a > \lambda} \{ \rho_a c(a) \} \\
 &= \min_{a > \lambda} \{ \rho_{a_1} c(a_1), \rho_{a_2} c(a_2), \rho_{a_3} c(a_3) \} \\
 &= \min_{a > \lambda} \left\{ \frac{3}{4}(2), \frac{3}{5}(5), \frac{1}{2}(5.5) \right\} \\
 &= \min_{a > \lambda} \{ 1.5, 3, 2.75 \} \\
 &= 1.5
 \end{aligned}$$

Por lo tanto, el costo promedio mínimo es 1.5 y la política óptima es utilizar la tasa de servicio $a_1 = 4$.

Universidad Juárez Autónoma de Tabasco.
México.

Conclusión

En esta tesis se planteó y se resolvió un problema de líneas de espera controladas con parámetro a tiempo continuo y con espacio numerable de estados. Para alcanzar esta meta se desarrolló la teoría básica de procesos estocásticos.

La teoría de procesos estocásticos fundamentó la teoría de líneas de espera, en el presente trabajo se ha dado lo elemental para definir los problemas de interés.

La teoría de líneas de espera proporciona buenos modelos para problemas interesantes de la vida real. Sin embargo, nuestro interés no sólo es modelar la cola sino controlarla. Para este fin se estudió la teoría de control estocástico en el caso de espacio de estados y acciones discretos, este contexto se conoce como Cadenas de Decisión de Markov o Cadenas de Control de Markov.

En este trabajo se aplicó el método de uniformización para resolver problemas a tiempo continuo, es decir, se lleva el problema original a un problema a tiempo discreto y con espacio numerable de estados. Además, se utilizó el método de Aproximaciones Sucesivas para aproximar el espacio numerable de estados por un espacio finito de estados (ver [14]). El método de

programación dinámica nos proporciona la solución a los problemas de cadenas de decisión de Markov, a tiempo discreto y con espacio finito de estados.

El ejemplo que se consideró para este trabajo fue resuelto utilizando Excel, la ejemplificación de la teoría queda cubierta satisfactoriamente.

Así, este trabajo nos proporciona la base para continuar con estudios más avanzados sobre el tema de líneas de espera controladas, por ejemplo, utilizar un espacio de estados y acciones continuos.

Bibliografía

- [1] Ash R.B., Probability and Measure Theory, Academic Press, Massachusetts, 2000.
- [2] Brzeźniak Zdzisław and Zastawniak Tomasz, Basic Stochastic Processes, Springer, 2002.
- [3] Dionicio Arévalo M. y Chablé de la Cruz J., Tesis: Cadenas de Decisión de Markov, División Académica de Ciencias Básicas, Universidad Juárez Autónoma de Tabasco, 2009.
- [4] Gou X., Hernández-Lerma O., Continuous-Time Markov Decision Processes, Springer, 2009.
- [5] ¹⁰ Gunter Bolch, Stefan Greiner, Hermann de Meer and Kishor S. Trivedi, Queueing Networks and Markov Chains, A Wiley-Interscience Publication, 1998.
- [6] ³ Hernández-Lerma O., Procesos Estocásticos: Introducción a la Teoría de colas, 2 Coloquio del Departamento de Matemáticas del Centro de Investigación y de Estudios

- Avanzados del IPN, Oaxtepec, 1981.
- [7] Hernández-Lerma O. y Bernart Lasserre J., Discrete-Time Markov Control Processes, Springer, 1996.
- [8] Lieberman G. J. y Hillier F. S., Introducción a la investigación de operaciones, Cuarta edición, 1997.
- [9] Lotfi Tadj., Optimal Design and Control of Queues, Sociedad de Estadística e Investigación Operativa, Vol. 13, No. 2, 2005, pg. 359-412.
- [10] Luis Rincón, Introducción a los Procesos Estocásticos, Departamento de Matemáticas Facultad de Ciencias UNAM Circuito Exterior de CU 04510 México DF, Enero 2012.
- [11] Puterman M. L., Markov Decision Processes, Wiley, New York, 1994.
- [12] Ross S.M., Stochastic Processes, Cánada, 1996.
- [13] Sennott L. I., Constrained discounted Markov decision chains, Probability in the Engineering and Informational Sciences, 5, 1991,pg. 463-475.

- [14] Sennott L. I., Stochastic Dynamic Programming and the Control of Queueing Systems, Wiley C  nada, 1999.
- [15] Serfozo R.F., An Equivalence Between Continuous and Discrete Time Markov Decision Processes, Operations Research, Vol. 27, No. 3, 1979, pg. 616-620.

CONTROL DE LINEAS DE ESPERA CON COSTO PROMEDIO

ORIGINALITY REPORT

8%

SIMILARITY INDEX

PRIMARY SOURCES

1	mat.izt.uam.mx Internet	859 words — 5%
2	www.monografias.com Internet	184 words — 1%
3	core.ac.uk Internet	95 words — 1%
4	ubaguayana-io.blogspot.com Internet	75 words — < 1%
5	www.coursehero.com Internet	70 words — < 1%
6	docplayer.es Internet	47 words — < 1%
7	es.slideshare.net Internet	29 words — < 1%
8	www.repositorio.usac.edu.gt Internet	29 words — < 1%
9	www.slideshare.net Internet	28 words — < 1%

10 Daichi Yanagisawa. "Designing method for large queueing system by walking-distance introduced queueing theory", 2008 SICE Annual Conference, 08/2008 21 words — < 1%
Crossref

11 Luis Carlos Palacios Acero. "Dirección estratégica", Ecoe Ediciones S. A. S., 2023 20 words — < 1%
Crossref

12 www.scribd.com 20 words — < 1%
Internet

EXCLUDE QUOTES ON

EXCLUDE BIBLIOGRAPHY ON

EXCLUDE SOURCES

EXCLUDE MATCHES

OFF

< 20 WORDS