



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

NETWORK REVENUE MANAGEMENT EN AEROLÍNEAS RESUELTO A TRAVÉS DE PROGRAMACIÓN DINÁMICA ROBUSTA

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN DE OPERACIONES

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

MARCELO JAVIER JARAMILLO QUIJADA

PROFESOR GUÍA:

FERNANDO ORDÓÑEZ PIZARRO

MIEMBROS DE LA COMISIÓN:

MARCEL GOIC FIGUEROA

SEBASTIÁN RÍOS PÉREZ

MANUEL REYES JARA

SANTIAGO DE CHILE

OCTUBRE 2012

RESUMEN DE LA TESIS
PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL INDUSTRIAL
POR: MARCELO JARAMILLO
FECHA: OCTUBRE DE 2012
PROF. FERNANDO ORDOÑEZ

NETWORK REVENUE MANAGEMENT EN AEROLÍNEAS RESUELTO A TRAVÉS DE PROGRAMACIÓN DINÁMICA ROBUSTA

En el presente trabajo se considera el problema de asignación de asientos para diversas clases de clientes en una red de Aerolíneas, más conocido en inglés como *Network Revenue Management*. Usando la metodología de programación dinámica, se busca maximizar el beneficio esperado sujeto a restricciones de tiempo y capacidad de los aviones en la red.

Uno de los problemas que enfrentan las aerolíneas al momento de vender vuelos interconectados en una red es qué precio fijar para cada clase de cliente de tal forma de no dejar potenciales compradores fuera al fijar precios altos, ni perder potenciales ingresos con precios bajos. Para gestionar la demanda a través del tiempo, las aerolíneas utilizan políticas de control de asientos como *Booking Limits*, *Protection Levels* o *Bid Prices*, cuya solución se obtiene de resolver problemas de optimización dinámicos o estáticos.

Esta tesis aborda este problema cuando la demanda está sujeta a incertidumbre. Bajo este escenario el problema es altamente riesgoso, pues los costos de operación son elevados y el producto que se ofrece es perecible, es decir, los asientos libres no se pueden inventariar lo que genera pérdidas. Para enfrentar esto se propone resolver usando optimización aversa al riesgo más conocida como optimización robusta.

El enfoque de optimización robusta es optimizar contra el peor de los casos que pudiera surgir debido a la incertidumbre en la demanda, encontrando políticas de control en la capacidad de los aviones relativamente insensibles a variaciones en la estimación de demanda.

La formulación robusta intenta mitigar el impacto de los errores en la estimación de probabilidades de transición mediante la elección de una política óptima *maximin*, donde la minimización es sobre un conjunto de probabilidades de transición y el objetivo es escoger una política que maximice los beneficios esperados sobre este conjunto.

Los experimentos que se realizaron muestran que cuando el riesgo supera cierto umbral, el modelo robusto captura de forma más eficiente el riesgo y obtiene resultados esperados mejores que los modelos tradicionales.

A mi abuelita Teresa Zapata (q.e.d.)

ÍNDICE GENERAL

Resumen	ii
Índice general	iv
1. Introducción.....	1
1.1 Planteamiento del problema.....	3
1.2 Solución propuesta y resultados esperados	4
1.3 Objetivos Generales.....	4
1.4 Objetivos Específicos	5
1.5 Organización de la tesis	5
2. Conceptos básicos de Revenue Management.....	7
2.1 Qué significa Revenue Management.....	7
2.2 Marco conceptual del RM	8
2.3 Marco conceptual de Optimización Robusta.....	10
2.4 Investigaciones realizadas	11
3. Modelos y tipos de control en Revenue Management.....	14
3.1 Control por capacidad para un recurso	14
3.1.1 Tipos de control	15
3.1.2 Modelos estáticos	16
3.1.3 Modelos dinámicos.....	20
3.2 Control por capacidad para una red.....	21
3.2.1 Control óptimo para una red mediante programación dinámica.....	22
3.2.1.1 Estructura de control óptimo	24
3.2.1.2 Control mediante Bid Price.....	26

3.2.2 Aproximaciones basados en modelos de redes.....	27
3.2.2.1 Modelo de programación lineal determinista	28
4. Modelo de programación dinámica mediante optimización robusta.....	32
4.1 Motivación.....	33
4.2 Programación dinámica robusta para un horizonte finito.....	34
4.2.1 Descripción del modelo	35
4.3 Conjuntos manejables de medidas condicionales.....	44
4.3.1 Modelo de Verosimilitud	45
4.3.2 Modelo de Entropía Relativa.....	48
4.3.3 Modelo Elipsoidal.....	50
4.3.4 Modelo de Entropía Relativa mediante norma L1.....	52
5. Resultados computacionales.....	56
5.1 Definición del problema	56
5.2 Desempeño computacional.....	58
5.3 Cálculo del valor esperado y Bid Prices.....	60
5.4 Simulación y comparación del modelo robusto vs tradicional.....	63
6. Conclusiones.....	66
Referencias	69

CAPÍTULO 1

INTRODUCCIÓN

Cada vendedor de un producto o servicio se enfrenta a un número de decisiones fundamentales. Por ejemplo, las Aerolíneas tienen que decidir qué precio fijar, cuántos asientos reservar para clase turista y primera clase, cuándo hacer ofertas, etc.

Cualquiera que se haya enfrentado a ese tipo de decisiones, sabe la incertidumbre que hay envuelta. A la Aerolínea le gustaría vender cuando las condiciones del mercado sean las más favorables, fijando el precio óptimo, no tan alto para no dejar potenciales compradores fuera y no tan bajo para no perder potenciales ingresos.

En la industria de servicios tales como Aerolíneas, Transporte y Hoteles, cuya capacidad es limitada, se enfrentan a diario con el problema de asignar en forma eficiente el inventario disponible a diferentes segmentos de la demanda.

Como el producto que se ofrece es perecible, es necesario el uso de herramientas tales como *Revenue Management* que permiten gestionar la demanda mediante distintas técnicas, por ejemplo: Decisiones sobre el precio (cómo establecer el precio publicado, cuándo hacer ofertas), Decisiones estructurales (qué tipo de venta usar, qué segmentación o mecanismos de diferenciación utilizar) y Decisiones basadas en capacidad (si aceptar o rechazar una oferta de compra, cómo asignar la capacidad a distintos segmentos), es en esta última categoría en la cual se enfocara la investigación.

El uso de distintas técnicas de optimización han sido esenciales en el desarrollo de herramientas de *Revenue Management*, por ejemplo, en la industria de las Aerolíneas se utilizan modelos de asignación de asientos para diversas clases de clientes, más conocido como *The passenger-mix problem* o más bien *Network Revenue Management*, en inglés, que permiten maximizar el beneficio esperado, sujeto a restricciones de tiempo y capacidad de los aviones en la red.

Este trabajo de tesis tiene como objetivo estudiar la metodología del cálculo de precios óptimos (*bid prices*) para una red de vuelos interconectados, considerando múltiples clases de clientes e incertidumbre en la estimación de demanda, con el fin de proponer un modelo de programación dinámica basado en las capacidades de la red, tomando en cuenta un enfoque robusto, es decir, encontrar una política de control en las capacidades de los aviones que sea relativamente insensible a variaciones en la estimación de demanda.

La propuesta de este trabajo es el uso de programación dinámica robusta en tiempo discreto para el problema de *revenue management* en aerolíneas en presencia de incertidumbre en la demanda. El objetivo de la formulación robusta es capturar la ambigüedad de las probabilidades de transición subyacentes al buscar la política de control óptima tratando de generar mejores valores futuros en comparación con un modelo no robusto. (Siguiendo la nomenclatura en [28], diremos que existe *incertidumbre* cuando se conoce la distribución de la variable aleatoria y *ambigüedad* cuando el modelo de incertidumbre no es preciso).

La ambigüedad es modelada asociando un conjunto de medidas condicionales¹ a cada par estado-acción. Por lo tanto, en la formulación robusta, cada política de control tiene un set de medidas condicionales asociadas a esta.

Se tratarán técnicas de Nilim y El Ghaoui [31] e Iyengar [28] para la construcción de conjuntos adecuados de medidas condicionales que permitan resolver de manera eficiente la política óptima robusta. Estos conjuntos se construyen a partir de intervalos de confianza

¹ Estos conjuntos se explicarán con más detalle en el capítulo 4 de esta tesis.

asociados a la distribución de probabilidad estimada, por lo tanto puede ser elegido para garantizar distintos niveles de confianza en la política óptima robusta.

1.1 Planteamiento del problema

El problema principal al que se enfrentan las aerolíneas al momento de vender pasajes de vuelos interconectados en una red, es si aceptar o rechazar² a un cliente de clase j en el tiempo t cuando existen $\mathbf{n} = n(1), \dots, n(k)$ asientos disponibles en la red, donde $n(k)$ representa la capacidad disponible del avión k . Para ello en *revenue management* se ha desarrollado una variedad de modelos que permiten tomar la decisión correcta, estos distintos modelos tiene sus ventajas y desventajas que se analizarán en el capítulo 3 de esta tesis.

Al tomador de decisiones le gustaría saber cómo se comportará la demanda futura para así generar una política óptima que maximice los ingresos de la red, se llamará ‘Política’ al conjunto de decisiones que se utilice al momento de aceptar o rechazar clientes para el horizonte finito que se considere, en cada estado posible de la red. Sin embargo, la realidad es distinta, ya que la presencia de incertidumbre en la naturaleza de la demanda puede hacer que políticas deterministas resulten poco eficientes frente a la demanda real.

Por lo tanto, lo complejo del problema es debido a la demanda incierta y la presencia de múltiples clases de clientes con distinta disposición a pagar por el mismo asiento. El desafío de esta situación es tener las herramientas para separar o segmentar los clientes y cobrarles de acuerdo a lo que están dispuestos a pagar. Además la metodología para encontrar una política de control óptima se complica aún más por la estructura de la red que incluye vuelos paralelos e interconectados.

La mayoría de los modelos en *Revenue Management* asumen que el tomador de decisiones conoce la distribución probabilística del proceso de demanda para cada clase, de

² En la realidad las compañías no rechazan a los clientes, si no que suben o bajan el precio de sus productos para lograr este efecto. Por lo tanto cuando se mencione esta palabra, significa que nos referimos a este hecho.

hecho las compañías cuentan con avanzados modelos de predicción de demanda. Estas estimaciones, que se basan en datos históricos, son difíciles de analizar sin un sistema de apoyo en las decisiones. Estos sistemas dependen, y pueden ser muy sensibles, de los supuestos de distribución de la demanda, cuya estimación sólo se conoce estadísticamente.

1.2 Solución propuesta y resultados esperados

En lugar de tratar de calibrar la incertidumbre en la demanda utilizando modelos de pronósticos, la solución que se propone en este trabajo es utilizar un modelo robusto. Es decir, el objetivo principal es encontrar una política de control de asientos que sea relativamente insensible ante variaciones en la demanda. Como resultado se espera que los ingresos esperados, generados por una solución robusta, sean menos variables que los generados por los modelos tradicionales de *Revenue Management*.

La ventaja de usar este nuevo modelo es que en presencia del peor caso, o demandas muy distintas a la esperada, se espera que el modelo robusto genere mejores ingresos promedio. Sin embargo, la desventaja de un modelo robusto es que las soluciones tienden a ser más conservadoras y generan menores ingresos promedio comparado con modelos tradicionales.

1.3 Objetivos Generales

- I. Estudiar en detalle el cálculo de precios óptimos de pasajes aéreos que se realizan en el área de *Revenue Management*, para una red de vuelos interconectados con múltiples clases de clientes.
- II. Generar un modelo matemático robusto para este problema tomando en cuenta la incertidumbre en los supuestos de demanda.
- III. Calcular políticas óptimas del modelo tradicional y robusto.

1.4 Objetivos Específicos

- I. Construir un modelo de Programación Dinámica que genere una política de control óptima, es decir, que indique cuándo aceptar o rechazar un cliente, dependiendo de las capacidades de la red y el tiempo.
- II. Encontrar una política de control de asientos que sea relativamente insensible ante variaciones en la estimación de demanda, mediante un modelo de programación dinámica robusta.
- III. Estudiar y comparar los valores a los cuales una aerolínea esta dispuesta a vender un asiento (*bid prices*) generados por los modelos de programación dinámica y dinámica robusta.
- IV. Construir un simulador que permita validar ambos modelos bajo distintos niveles de riesgo en la demanda.

1.5 Organización de la tesis

En el siguiente capítulo se presenta una visión general de la literatura que describe la metodología *Revenue Management* y optimización robusta, además de investigaciones que se han realizado en estas áreas. En el capítulo 3 se describen varios modelos (estáticos y dinámicos) agrupados en dos categorías, cuando sólo hay disponible un recurso y el caso general de una red. En el capítulo 4 se presenta el modelo con enfoque robusto y tres conjuntos de medidas condicionales, de los cuales dos son usados para encontrar la política óptima. Finalmente, en el capítulo 5 se presentan resultados computacionales de la simulación, donde se comparan los modelos tradicionales explicados en el capítulo 3 con el modelo robusto. La última sección explica las conclusiones a las que se llegaron con este trabajo.

CAPÍTULO 2

CONCEPTOS BÁSICOS DE REVENUE MANAGEMENT

Este capítulo introduce conceptos fundamentales sobre *Revenue Management* (RM), incluye un marco conceptual para entender sus objetivos, en qué condiciones se puede aplicar y una visión general de cómo funciona un sistema de RM. Además ofrece una descripción breve de optimización robusta e investigaciones que se han realizado tanto en el área de RM como en el área de optimización robusta. Estos conceptos están basados en la descripción dada por el libro “*The Theory and Practice of Revenue Management*” escrito por Talluri y Van Ryzing [36].

2.1 Qué significa *Revenue Management*

La definición clásica de RM sostiene que es el arte de vender el producto correcto, al cliente correcto, en el momento correcto. También se puede entender como el proceso de entender, anticiparse e influenciar el comportamiento del consumidor con el fin de maximizar los ingresos esperados [19].

Cualquier comerciante de un producto o servicio se enfrenta a decisiones relacionadas con el proceso de venta, entre ellas: cuándo vender, a qué precio, y por cuánto tiempo. En cada una de estas decisiones juega un rol importante la incertidumbre en la demanda. Por ejemplo, si se rechaza vender a un cliente ahora, ¿existe la seguridad que en el futuro próximo vendrán más clientes con mejor o igual disposición a pagar?

Las empresas por lo general deben tomar decisiones sobre cómo tratar la demanda. Por ejemplo, ¿cómo puede una compañía segmentar a sus clientes ofreciéndoles diferentes términos contractuales y condiciones que exploten los diferentes comportamientos o disponibilidades al comprar sus productos? Una vez segmentados los clientes, ¿qué precio se debe cobrar a cada segmento?

La disciplina conocida como *Revenue Management* se encarga de este tipo de decisiones relacionadas con la demanda, además de la metodología y sistemas requeridos para implementarla.

RM es una idea muy antigua. Todo comerciante en la historia de la humanidad se ha enfrentado a decisiones de tipo RM. Sin embargo, la novedad del RM no son las decisiones por si solas, sino más bien cómo se toman las decisiones. La verdadera innovación del RM, se encuentra en el método de toma de decisiones.

El nuevo enfoque es impulsado por dos fuerzas. Primero, avances científicos en economía, estadística e investigación operativa permiten formular modelos de demanda y condiciones económicas que cuantifican el grado de incertidumbre y restricciones que enfrenta el tomador de decisiones. Segundo, avances en tecnologías de información permiten manejar gran cantidad de datos, resolver algoritmos complejos en poco tiempo y automatizar las decisiones de demanda resultantes. Estas dos fuerzas combinadas hacen posible un enfoque totalmente nuevo para la toma de decisiones.

2.2 Marco conceptual del RM

¿En qué situaciones es posible aplicar RM? Una posible respuesta es: en cualquier negocio donde las decisiones tácticas sobre la demanda son importantes, y existe una cultura relacionada con la tecnología y el *management* que permiten implementarlo. Pero esto a su vez lleva a la pregunta: ¿cuándo surgen estas condiciones? Es lo que se explicará a continuación.

◆ La naturaleza multidimensional de la demanda

La demanda de una empresa tiene múltiples dimensiones, incluyendo (1) los diferentes productos que vende, (2) los tipos de cliente que atiende, sus preferencias por productos y comportamientos de compra, y (3) el tiempo. Si bien es posible agregar otras dimensiones (tales como, ubicación o canal de venta), estas tres sirven como ejemplo para ilustrar la idea.

◆ Vínculos entre decisión y gestión de la demanda

Si la gestión de demanda y el entorno fuesen completamente independientes, el problema de la toma de decisión sería considerablemente más simple. Pero típicamente uno o más de los siguientes tres factores relaciona la demanda a través de las dimensiones.

Primero, múltiples productos pueden compartir capacidad de producción o tener costos de producción conjunta. En estos casos, las decisiones de gestión de demanda para distintos productos están relacionadas. Por ejemplo, aceptar demanda para el producto i en tiempo t , puede significar rechazar demanda para producto k en tiempo $t + 1$.

Segundo, aún cuando las restricciones de producción no relacionen las decisiones de demanda, el comportamiento del cliente a menudo lo hace. Estos pueden elegir entre productos sustitutos o crear estrategias sobre cómo compra un producto determinado en el tiempo.

Finalmente, las decisiones de demanda para diferentes productos, clientes y periodos de tiempo pueden estar relacionadas en términos de cuánta información haya obtenido la compañía hasta ese tiempo.

◆ Condiciones de mercado que conducen al RM

Teniendo esta conceptualización del problema, podemos obtener una mirada más profunda sobre las condiciones que deben darse para que RM sea beneficioso. Estas son algunas:

- I. *Heterogeneidad de los clientes.* Mientras más heterogéneos son los clientes, mayor es el potencial de utilizar este factor para incrementar los ingresos. Los hoteles y aerolíneas proveen un buen ejemplo de esta característica.
- II. *Variabilidad de la demanda.* Mientras más varíe la demanda a través del tiempo (debido a estacionalidades o shocks) y mientras más incertidumbre haya sobre la demanda futura, más complicado será tomar decisiones, aumentando las posibilidades de cometer un error.
- III. *El precio como señal de calidad.* A veces el precio no es símbolo de status ni tampoco señal de calidad. En aerolíneas, generalmente los clientes no asocian la tarifa de un ticket con la calidad del vuelo, pagar US\$300 extras por un pasaje comprado a último minuto no significa tener un mejor vuelo.

◆ Descripción general de un sistema de RM

En la práctica un sistema de RM consta de cuatro pasos:

- I. Recopilación de datos: Recolectar y almacenar datos históricos (precios, demandas, cancelaciones, etc.)
- II. Estimación y predicción: Estimar parámetros del modelo de demanda con los datos recopilados.
- III. Optimización: Encontrar políticas óptimas de control (*Bid Price*, *Booking limits*, precios, descuentos, etc.)
- IV. Control: Controlar la venta de inventario cuya política óptima fue encontrada en III.

2.3 Marco conceptual de Optimización Robusta

La investigación actual sobre optimización robusta está estrechamente relacionado con la teoría de control robusto; véase por ejemplo Zhou et al. [45]. Estas ideas sirvieron de contexto para desarrollar diversos programas matemáticos, por ejemplo, en el trabajo de Ben-Tal y

Nemirovski [7] los autores formulan problemas de optimización robusta en programación lineal, programación cuadrática y problemas de programación convexos en general. Independientemente El-Ghaoui et al. [21] estudiaron la misma formulación de optimización robusta para problemas de programación semidefinida. Más recientemente, este enfoque se ha extendido a problemas de programación entera (Bertsimas y Sim [14]).

La ventaja de contar con soluciones robustas es que tienen el potencial de ser soluciones viables en la práctica, ya que tienden a no estar muy lejos de la solución óptima y superan significativamente a la solución óptima en el peor de los casos (Goldfarb y Iyengar [26]; Bertsimas y Sim [15]).

El enfoque general de la optimización robusta es optimizar contra el peor de los casos que pudiera surgir debido a la incertidumbre de los datos, mediante el uso de min-max en la función objetivo. La solución resultante del problema robusto es insensible a variaciones de los datos, ya que minimiza el peor de los casos y por lo tanto es *immune* a esta incertidumbre. La metodología de optimización robusta asume que los parámetros inciertos pertenecen a un conjunto de incertidumbre limitado. Para los conjuntos de incertidumbre bastante generales, la contraparte que resulta del problema robusto puede ser similar al problema original, y por lo tanto puede tener una complejidad comparable. Por ejemplo, un problema lineal con parámetros inciertos perteneciente a un conjunto de incertidumbre poliédrica, tiene un problema robusto asociado que es LP, cuyo tamaño es polinomial en comparación con el problema original (Ben-Tal y Nemirovski [8]). Por lo tanto, surge una pregunta importante: ¿Cómo formular un problema robusto que no sea más difícil de resolver que su homólogo determinista?

2.4 Investigaciones realizadas

El problema ante el control de capacidad con un solo recurso fue introducido por Littlewood [30] con dos clases de clientes que llegan de forma secuencial, y fue posteriormente extendido a múltiples clases de clientes. Con llegadas consecutivas, la política de reserva óptima se encuentra mediante fijar niveles de protección (Talluri y van Ryzin [36]). Estos niveles de

protección (*Nested booking limits* en inglés) se pueden calcular en forma exacta (programación dinámica) o mediante heurísticas (*Expected marginal seat revenue*: EMSR-a y EMSR-b. Belobaba [3], [4]).

Network RM es significativamente más complejo de resolver y encontrar la política de control óptima. En consecuencia, *network RM* es comúnmente resuelto mediante heurísticas, ya sea por medio de aproximaciones en la función de beneficio esperada de la programación dinámica (Bertsimas y Popescu [13], Adelman [1]) o restringiendo el conjunto de políticas factibles. Las políticas de control más conocidas son *partitioned booking limits*, *nested booking limits* y *bid price*, en inglés. Aunque estos controles han sido tradicionalmente obtenidos con formulaciones de programación matemática (Talluri y van Ryzin [36]), últimamente se le ha dado mayor atención a programación dinámica y formulaciones óptimas de control (Adelman [1], Topaloglu [38], Talluri [35], Akan y Ata [2]). Los controles, además pueden ser ajustados con un algoritmo de gradiente estocástico basado en muestras de demanda (Bertsimas y de Boer [16], van Ryzin y Vulcano [39]).

Los modelos tradicionales de RM, que asumen información de demanda probabilística, son muy sensibles a errores en el método de predicción (Weatherford y Belobaba [41]). En lugar de estimar una distribución de probabilidad de la demanda a partir de datos históricos, los métodos de optimización utilizan directamente datos históricos de ventas para optimizar los controles de reserva con un algoritmo de gradiente estocástico (van Ryzin y McGill [40], Bertsimas y de Boer [16], van Ryzin y Vulcano [39]). Métodos no paramétricos de optimización también se han aplicado ampliamente en la gestión del inventario y *pricing* (Godfrey y Powell [25], Levi et al. [29], Erken y Maglaras [22], Rusmevichientong et al. [33], Huh y Rusmevichientong [27], Besbes y Zeevi [17]).

En entornos empresariales nuevos o no estacionarios, los datos históricos no están disponibles o son de poco valor. De hecho, la única información de demanda que está disponible a menudo proviene de expertos de la industria, basados en su juicios y experiencia. Por lo tanto, las decisiones deben tomarse bajo limitada información de la demanda, mediante optimización

robusta. Métodos de optimización robusta han sido ampliamente utilizados en el control de inventario con el criterio *maximin* (Perakis y Roels [32], Scarf [34], Gallego y Moon [23], Gallego et al. [24], Bertsimas y Thiele [12], Ben-Tal et al. [10]) y el criterio de *minimax regret* (Yue et al. [44], Perakis y Roels [32]).

En el trabajo de Perakis y Roels [32], se describen modelos robustos para distribuir la capacidad en el avión (mediante *partitioned booking limits*, *nested booking limits* y *bid price*, estos conceptos se explicarán en el capítulo 3), asumen limitada información sobre la demanda y que pertenece a ciertos conjuntos de incertidumbre poliedricos, no realizan ningún supuesto sobre la secuencia de llegada sobre la demanda. Consideran dos enfoques robustos: *maximin* y *minimax regret*. El criterio *maximin* garantiza un nivel de beneficio esperado, este es más apropiado para tomadores de decisiones adversos al riesgo, por el contrario, el criterio *minimax regret* minimiza el costo de oportunidad de no saber la demanda y genera recomendaciones o políticas menos conservadoras.

Por otro lado, Nilim y El Ghaoui [31] e Iyengar [28], han desarrollado técnicas de programación dinámica robusta que se pueden aplicar al problema de RM. Ellos asumen que las probabilidades de transición en la cadena de markov, correspondientes a cada par estado-acción, no se conocen con exactitud. Esta ambigüedad en las probabilidades de transición son modeladas mediante conjuntos de medidas condicionales. Estos conjuntos están basados en distancias de entropía relativa (Kullback-Leibler y aproximaciones en norma L_2 y L_1) que conducen tanto a una representación estadísticamente precisa de la incertidumbre, y una complejidad de la recursión robusta que es casi la misma que la recursión clásica. Lo que se busca con lo anterior, es una política de control tal que sea la mínima esperada sobre el conjunto de medidas asociadas, y el objetivo del tomador de decisiones es escoger una política con máximo valor, es decir, se adopta un enfoque *maximin*. El nombre de esta formulación es conocida como *Robust Dynamic Programming*, estos modelos se describen en detalle en el capítulo 4.

CAPÍTULO 3

MODELOS Y TIPOS DE CONTROL EN REVENUE MANAGEMENT

En este capítulo se describirán los modelos tradicionales de RM, basados en capacidad, es decir, cómo asignar óptimamente la capacidad de uno o múltiples recursos a diferentes clases de clientes. Primero se explicará el caso de un solo vuelo y luego cuando existen varios aviones interconectados en una red, más conocido como *network* RM.

3.1 Control por capacidad para un recurso

En la realidad, la mayoría de estos problemas son *network* RM, pero en la práctica son resueltos como una colección de problemas de recursos individuales (tratando a cada recurso en forma independiente).

El modelo asume que la compañía vende la capacidad del avión a m distintas clases de clientes que requieren el mismo recurso. Estas clases representan distintas categorías de precio diferenciadas mediante restricciones y condiciones de venta. Los clientes de cada segmento pueden comprar sólo el pasaje perteneciente a su categoría.

Todos los asientos del avión son iguales y los clientes pueden comprar sólo un asiento. Por lo tanto, el problema central de la Aerolínea es cómo asignar óptimamente los recursos de su avión a las distintas clases de clientes. Esta asignación debe hacerse en forma dinámica mientras se observa la demanda y tomando en cuenta la incertidumbre de la demanda que aún no ha llegado.

3.1.1 Tipos de control

En la industria aérea, los sistemas de reservas proporcionan diferentes mecanismos para el control de la disponibilidad. Estas políticas de control son las que deciden cómo manejar la demanda a través del tiempo. A continuación se explican las más conocidas:

- I. *Booking Limits*: Este tipo de control limita la cantidad de capacidad que se puede vender a una determinada clase en un punto dado del tiempo. Existen dos tipos, estos pueden ser *partitioned* o *nested*. El primero divide la capacidad disponible en bloques fijos, una para cada clase, mientras que el segundo se superpone en forma jerárquica, es decir, las clases de mayor rango tienen acceso a toda la capacidad, incluso la que esta reservada para clases de menor rango.
- II. *Protection Levels*: Este control reserva (*protect*) una cantidad específica de capacidad para una clase o conjunto de clases en particular. Esta política puede ser *partitioned* o *nested*, al igual que el caso anterior. Por ejemplo, y_j se define como la cantidad a reservar para las clases $j, j-1, \dots, 1$ en conjunto para el caso *nested*. La figura 3.1 muestra la relación entre *protection levels* y *booking limits*.
- III. *Bid Price*: Lo que distingue a esta política de las otras dos es que ésta se calcula mediante ingresos (*revenue*), no mediante clases. Esta política establece un precio umbral (que dependerá de variables como capacidad restante o tiempo) tal que la demanda por algún producto es aceptada si su tarifa o precio es mayor que este precio umbral y rechazada en caso contrario. La figura 3.1 muestra como los *Bid Prices* $\pi(x)$ pueden ser usados.

Cuando hay 12 o menos unidades disponibles, el precio umbral es de \$75 por lo que sólo la clase 1 es aceptada, cuando quedan menos de 22 asientos y más de 12, el *Bid Price* es de \$50, por lo tanto, las clases 1 y 2 son aceptadas. Y así sucesivamente.

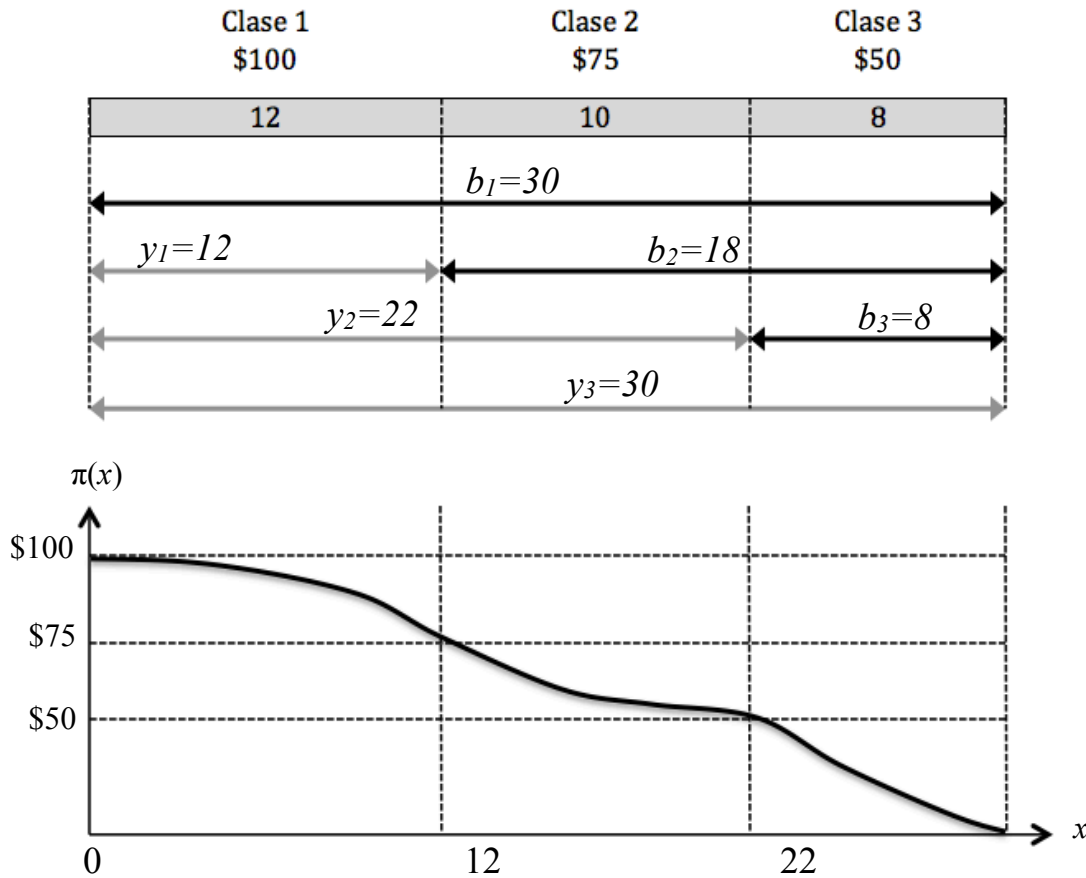


Figura 3.1. Relación entre *booking limits* b_j , *protection levels* y_j , y *bid price* $\pi(x)$.

3.1.2 Modelos estáticos

Estos modelos hacen las siguientes suposiciones: las distintas clases de clientes llegan en intervalos no superpuestos y de forma ordenada, las demandas de cada clase son variables aleatorias independientes y no dependen de la política de control, en particular, no dependen de la capacidad disponible para otras clases y finalmente, no se permiten llegadas grupales. Se explicará primero el caso con dos clases y luego el caso general para n -clases.

I. El modelo de Littlewood

La primera solución al problema de un recurso la propuso Littlewood en 1972 [30], con dos clases de productos, cuyos precios son $p_1 > p_2$ y donde además, la demanda de la clase 2 llega antes que la de la clase 1.

Dada una demanda para la clase j , D_j , la capacidad disponible N y distribución de probabilidad $F_j(\cdot)$, el problema es decidir cuánta demanda de clase 2 aceptar antes de ver la realización de la demanda de clase 1.

La solución óptima se puede derivar mediante un análisis de ingreso/pérdida marginal: Suponer que quedan n unidades de capacidad disponible y llega el cliente de clase 2. Si aceptamos la llegada, se recibirán ingresos de p_2 . Si no se acepta, se venderán n (unidad marginal) unidades al precio p_1 , si y sólo si la demanda de clase 1 es n o mayor. Es decir, si y sólo si $D_1 \geq n$. En consecuencia, el beneficio esperado de reservar la n -ésima unidad para la clase 1 (beneficio marginal esperado) es $p_1 P(D_1 \geq n)$. Por lo tanto, tiene sentido aceptar llegadas de clientes 2 siempre que su precio supere este valor marginal, o equivalentemente

$$p_2 \geq p_1 P(D_1 \geq n) \quad (3.1)$$

Si para modelar la demanda se usa una distribución continua $F_1(x)$, entonces el nivel de protección óptimo y_1^* se encuentra por medio de la siguiente expresión

$$p_2 = p_1 P(D_1 > y_1^*) \iff y_1^* = F_1^{-1} \left(1 - \frac{p_2}{p_1} \right) \quad (3.2)$$

Lo anterior es conocido como la regla de Littlewood.

II. Modelo para n -clases

Ahora se considerará el caso general para $m > 2$ clases. Nuevamente se asume que la demanda llega de forma ordenada, de menor a mayor, de acuerdo a su disposición a pagar, es decir, $p_1 > p_2 > \dots > p_m$.

Este problema se puede resolver a través de programación dinámica, donde el estado n es la capacidad restante del avión. Al comienzo de cada etapa j , no se conoce la demanda D_j, D_{j-1}, \dots, D_1 y cuando el modelo está en la etapa j ocurre lo siguiente:

1. Llega la demanda D_j y se observa su valor.
2. Se decide la cantidad u de demanda a aceptar. La cantidad aceptada debe ser menor que la capacidad restante, es decir, $u \leq n$. Por lo tanto, el control óptimo u^* es una función del estado j , capacidad n , y demanda D_j , $u^* = u^*(j, n, D_j)$.
3. Se recibe un ingreso de $p_j u$, y se pasa a la siguiente etapa $j - 1$ con capacidad restante $n - u$.

Sea $V_j(n)$ la función de beneficio esperada al comienzo de la etapa j . Por consiguiente, la ecuación de Bellman es

$$V_j(n) = E \left[\max_{0 \leq u \leq \min\{D_j, n\}} \{p_j u + V_{j-1}(n - u)\} \right] \quad (3.3)$$

cuya condición de borde es

$$V_0(n) = 0, \quad n = 0, 1, \dots, N$$

Los valores u^* que maximizan la función objetivo $V_j(n)$ para cada etapa j y estado n , forman la política de control óptima para este modelo.

III. Heurísticas

En la práctica es muy común el uso de heurísticas, ya que son más simples de programar, se ejecutan más rápido y generan ingresos que, en muchos casos, están cerca del óptimo. Las heurísticas más conocidas para responder a este problema son EMSR–a y EMSR–b (*Expected Marginal Seat Revenue*, versiones a y b), ambas se atribuyen a Belobaba [3]. La que se expondrá a continuación es EMSR–b, ya que es la más usada.

EMSR–b está basada en aproximaciones que reducen el problema en cada etapa a dos clases, esta aproximación consta de agregar demanda de futuras clases y tratarlas como si fueran una sola con beneficio igual al promedio ponderado de los ingresos.

Se define la demanda agregada futura para las clases $j, j-1, \dots, 1$ como

$$S_j = \sum_{k=1}^j D_k$$

y el promedio ponderado de los ingresos para las clases $1, \dots, j$ denotado como \bar{p}_j

$$\bar{p}_j = \frac{\sum_{k=1}^j p_k E[D_k]}{\sum_{k=1}^j E[D_k]} \quad (3.4)$$

Entonces, el *protection level* para la clase j se encuentra por medio de la regla de Littlewood (3.2)

$$P(S_j > y_j) = \frac{p_{j+1}}{\bar{p}_j} \quad (3.5)$$

Es decir, se está reduciendo el problema anterior a un problema de dos clases, ya que se aceptará alguien de la clase $j+1$ (que es más barata), si y sólo si supera lo que se espera percibir en promedio con las clases siguientes.

3.1.3 Modelos dinámicos

Los modelos dinámicos relajan la condición que la demanda llega de forma ordenada y permiten llegadas aleatorias (por ejemplo, asumen que las llegadas se distribuyen de acuerdo a un proceso de Poisson). Sin embargo, los otros supuestos se mantienen, es decir, la demanda es independiente entre las clases y tiempo, y también son independientes de la política de control.

En el modelo dinámico, existen m clases (indexadas por j) cuyos precios son $p_1 \geq p_2 \geq \dots \geq p_m$. Hay en total T periodos de tiempo, indexados por t ($t = 1$ es el primer periodo y $t = T$ es el último periodo, cuando el avión despegue). En cada periodo se asume que el tiempo es suficientemente discreto tal que llega un solo cliente. La probabilidad que llegue la clase j en el tiempo t es $\lambda_j(t)$. El supuesto que llega un cliente por periodo implica lo siguiente

$$\sum_{j=1}^m \lambda_j(t) \leq 1$$

Sea n la capacidad restante, $V_t(n)$ la función de beneficio en periodo t y $R(t)$ una variable aleatoria tal que, $R(t) = p_j$ si la demanda de clase j llega en t y $R(t) = 0$ en caso contrario. Notar que $P(R(t) = p_j) = \lambda_j(t)$. Sea $u_t = 1$ si se acepta la llegada y $u_t = 0$ en caso contrario. El objetivo es maximizar la suma de ingresos en t y en $t + 1$, por lo tanto la ecuación de Bellman es

$$\begin{aligned} V_t(n) &= E \left[\max_{u_t \in \{0,1\}} \{R(t)u_t + V_{t+1}(n - u_t)\} \right] \\ &= V_{t+1}(n) + E \left[\max_{u_t \in \{0,1\}} \{(R(t) - \Delta V_{t+1}(n)) u_t\} \right] \end{aligned} \quad (3.6)$$

donde $\Delta V_{t+1}(n) = V_{t+1}(n) - V_{t+1}(n - 1)$ es el valor esperado marginal de la capacidad en el periodo $t + 1$.

Las condiciones de borde son

$$\begin{aligned} V_{T+1}(n) &= 0, & n = 0, 1, \dots, N \\ V_t(0) &= 0, & t = 1, \dots, T \end{aligned}$$

Una consecuencia inmediata de (3.6) es que si llega la clase j , es decir, $R(t) = p_j$, entonces es óptimo aceptar si y sólo si

$$p_j \geq \Delta V_{t+1}(n)$$

Así, la política óptima puede ser implementada usando el tipo de control *Bid Price*, que en este caso es igual al valor marginal,

$$\pi(n) = \Delta V_t(n) \tag{3.7}$$

Precios que superen este umbral son aceptados; aquellos que no, son rechazados.

3.2 Control por capacidad para una red

En el caso de una red, los clientes tienen acceso a productos que hacen uso de varios recursos combinados, lo que crea interdependencia, ya que la falta de disponibilidad en un solo recurso puede limitar la venta, por lo tanto, para maximizar el beneficio esperado, es necesario coordinar en conjunto la política de control en toda la red.

Considerar un modelo en redes puede generar beneficios significativos en RM. En efecto, varios estudios de simulación en aerolíneas han demostrado que el modelo en redes se comporta mejor que el modelo de un solo recurso [9, 10, 15, 16]. En la práctica, estas potenciales mejoras han sido suficientes para justificar inversiones significativas en sistemas *network* RM tanto en las industrias de aerolíneas como en la de hoteles.

Si bien los posibles beneficios pueden ser altos, la metodología de *network* RM plantea un gran desafío en la implementación, ya que aumenta considerablemente el volumen de datos que se debe reunir, almacenar y gestionar, además del impacto que puede generar en la organización.

3.2.1 Control óptimo para una red mediante programación dinámica

El principal problema que se abordará en este trabajo es el de una red compuesta por l legs o arcos (pares consecutivos de aviones en aerolíneas o días en el caso de hoteles) que se utilizan para servir un total de m clases de demanda. El inventario inicial está dado por un vector $\mathbf{N} = (N_1, \dots, N_l)$ de capacidades en los arcos. La red está descrita por una matriz $\mathbf{A} = [a_{ij}]$ de $l \times m$ filas y columnas, y un vector $\mathbf{R} = (R_1, \dots, R_m)$: R_j es la tarifa de la clase j , que utiliza a_{ij} unidades del recurso (*leg*) i . De este modo, una demanda de la clase j se define por su itinerario \mathbf{A}^j (una columna de la matriz \mathbf{A} , que es el conjunto de recursos que usa la clase j) y por su categoría de precio R_j . En una red de líneas aéreas, \mathbf{A} es una matriz de 0 – 1 que puede contener columnas repetidas para cada clase de tarifa en un itinerario determinado. Por lo tanto, $a_{ij} = 1$ si el recurso i es usado por producto j , y $a_{ij} = 0$ en caso contrario.

La Figura 3.2 muestra un pequeño ejemplo correspondiente a cinco nodos conectados en red. Los nodos son aeropuertos en distintas ciudades, mientras que los números entre paréntesis indican la capacidad disponible de cada avión (*legs* = 4), con capacidad total $\mathbf{N} = (450, 300, 275, 200)$.

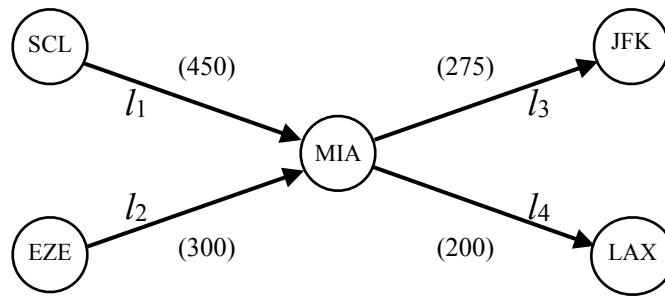


Figura 3.2: Ejemplo pequeño de una red de vuelos interconectados.

Los productos que la aerolínea ofrece pueden ser distintas combinaciones origen-destino, en este caso, el producto puede ser un viaje del aeropuerto SCL hasta JFK a través de MIA (producto 5 cuya tarifa es R_5). Cada itinerario, por ejemplo, se pueden dividir en clase económica (*low*) o *Business (high)*. Para que quede más claro, la siguiente matriz de incidencia, muestra qué *legs* ocupa cada producto

$$\begin{pmatrix} R \\ A \end{pmatrix} = \begin{matrix} l1 \\ l2 \\ l3 \\ l4 \end{matrix} \begin{pmatrix} R_1^l & R_1^h & R_2 & R_3 & R_4 & R_5 & R_6^l & R_6^h & R_7 & R_8 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Para resolver el problema se asume un horizonte finito de tiempo T , con t suficientemente pequeño tal que llega un solo cliente por periodo. El tiempo se cuenta hacia atrás, es decir, $t = T$ es el comienzo del horizonte de reserva y el tiempo $t = 0$ es el final del período de reserva.

La demanda (por venir) en el tiempo t se denota por \mathcal{D}^t , y $\bar{\mathcal{D}}^t$ representa el vector aleatorio de demandas acumuladas. Es decir, $\bar{\mathcal{D}}^t$ es una variable aleatoria que representa el número de solicitudes de la clase j desde el tiempo t hasta que despegue el avión (el orden no importa). Por lo general, existe poca información sobre el proceso de la demanda, que podría consistir en la demanda esperada por venir $\mathbf{D}^t = E[\bar{\mathcal{D}}^t]$, y, posiblemente, otro tipo de información disponible de herramientas de predicción, tales como cancelaciones o probabilidades de no llegada de un cliente.

El estado del sistema \mathbf{S} está dado por el tiempo t (t periodos de la salida) y la capacidad restante de los arcos (*legs*) $\mathbf{n} = (n_1, \dots, n_l)$. Por lo tanto, si se vende el producto j , el estado de la red cambia a $(\mathbf{n} - \mathbf{A}^j, t)$.

Luego, el problema general de control de inventario estocástico para *network* RM puede enunciarse como sigue: En el momento t , y dado que el estado de la red es \mathbf{S} , ¿se debe aceptar o rechazar una nueva solicitud de la clase j ? Tomando en cuenta que el objetivo general es el de

maximizar los ingresos totales esperados.

La decisión de aceptar o rechazar, determina una política de control de admisión, que en general está en función de la configuración de red actual \mathbf{n} , el tiempo t y la tarifa R_j . Se utilizará la notación $\mathbf{u}(t)$ como vector de decisiones, donde $u_j(t) = 1$ si se acepta la solicitud de demanda para el producto j en el periodo t , y $u_j(t) = 0$ en el caso contrario. Dado que a lo más se puede aceptar una solicitud en cada periodo, las decisiones $\mathbf{u}(t)$ están restringidas al conjunto $\mathcal{U}(\mathbf{n}) = \{\mathbf{u} \in \{0,1\}^n : \mathbf{A}\mathbf{u} \leq \mathbf{n}\}$.

En la realidad, la política de control tiene un impacto en el proceso de la demanda, porque la elección del cliente puede depender del conjunto de productos disponibles. La captura y cuantificación de este tipo de fenómeno es una tarea delicada que va más allá del alcance de este trabajo. Por lo tanto se hace la suposición de que el proceso de la demanda es independiente de la política de control.

3.2.1.1 Estructura de control óptimo

El modelo de programación dinámica estocástica proporciona la política óptima para el problema de *network* RM, mediante la evaluación de todo el árbol de posibilidades que permitirá tomar decisiones en cada periodo para generar los mayores ingresos futuros esperados.

Los estados $\mathbf{S} = (\mathbf{n}, t)$ se definen por el vector de capacidad disponible actual \mathbf{n} cuando el tiempo restante es t . Además, la probabilidad de que llegue demanda de clase j en tiempo t es p_j^t .

Por lo tanto, para encontrar la decisión óptima $u^*(\mathbf{n}, t, \mathbf{R})$ se define $DP(\mathbf{n}, t)$ como el beneficio máximo esperado a reunir desde el estado (\mathbf{n}, t) . Asumiendo demandas independientes, esto puede ser calculado mediante la siguiente ecuación de Bellman³:

³ Se usará el operador $(x)^+ = \max(x, 0)$ para $x \in \mathbf{R}$.

$$\begin{aligned}
 DP(\mathbf{n}, t) &= E \left[\max_{\mathbf{u} \in \mathcal{U}} \{ \mathbf{R}^\top \mathbf{u}(\mathbf{n}, t, \mathbf{R}) + DP(\mathbf{n} - \mathbf{A}^j \mathbf{u}, t - 1) \} \right] \quad (3.8) \\
 &= \sum_{j=1}^m p_j^t \max \{ DP(\mathbf{n}, t - 1), R_j + DP(\mathbf{n} - \mathbf{A}^j, t - 1) \} + p_0^t DP(\mathbf{n}, t - 1) \\
 &= DP(\mathbf{n}, t - 1) + \sum_{j=1}^m p_j^t (R_j - DP(\mathbf{n}, t - 1) + DP(\mathbf{n} - \mathbf{A}^j, t - 1))^+
 \end{aligned}$$

para todo $\mathbf{n} \leq N$, $t \leq T$, con las condiciones de borde

$$\begin{aligned}
 DP(\mathbf{n}, t) &= -\infty \text{ si } n_i < 0 \text{ para algún } i, \text{ y} \\
 DP(\mathbf{n}, 0) &= 0, \text{ para } \mathbf{n} \geq 0.
 \end{aligned}$$

El precio umbral o *Bid Price* $BP_j(\mathbf{n}, t)$ se define como

$$BP_j(\mathbf{n}, t) = DP(\mathbf{n}, t - 1) - DP(\mathbf{n} - \mathbf{A}^j, t - 1)$$

Por consiguiente, el valor de la función de beneficio puede ser expresada como sigue

$$DP(\mathbf{n}, t) = DP(\mathbf{n}, t - 1) + \sum_{j=1}^m p_j^t (R_j - BP_j(\mathbf{n}, t))^+. \quad (3.9)$$

Uno puede fácilmente demostrar que la función de valor es no decreciente en \mathbf{n} y t . Por otra parte, en el caso de un solo recurso, sin llegadas grupales, la función de valor es cóncava y los *bid price* son decrecientes en t y \mathbf{n} (K. T. Talluri and G. J. Van Ryzin [36], Diamond y Stone [20]). Tomando en cuenta estas propiedades, se puede demostrar que la política óptima $\mathbf{u}^*(\cdot)$ acepta una solicitud de la clase j , si y sólo si su precio R_j es superior a al precio umbral $BP_j(\mathbf{n}, t)$ y queda capacidad suficiente, esto quiere decir que $\mathbf{u}^*(\cdot)$ satisface lo siguiente:

$$u_j^*(\mathbf{n}, t, \mathbf{R}) = \begin{cases} 1 & \text{si } R_j \geq DP(\mathbf{n}, t - 1) - DP(\mathbf{n} - \mathbf{A}^j, t - 1) \text{ y } \mathbf{A}^j \leq \mathbf{n} \\ 0 & \text{si no.} \end{cases} \quad (3.10)$$

Esto refleja la noción intuitiva de que se acepta una solicitud de reserva para un producto j sólo si su precio es superior al costo de oportunidad de reducir la capacidad de los recursos necesarios para satisfacer la solicitud.

3.2.1.2 Control mediante *Bid Price*

Existen varias formas de calcular los *bid price*, estos se pueden obtener a través de programación dinámica, precios duales de la programación lineal determinista (Williamson [43]) o aproximaciones de estos métodos, por ejemplo, suponiendo que la función de beneficio $DP(\mathbf{n}, t - 1)$ se puede derivar y tiene gradiente $\nabla DP(\mathbf{n}, t - 1)$, entonces la condición para aceptar el producto j puede ser aproximada por:

$$\begin{aligned} R_j &\geq DP(\mathbf{n}, t - 1) - DP(\mathbf{n} - \mathbf{A}^j, t - 1) \\ &\approx \nabla DP^\top(\mathbf{n}, t - 1) \mathbf{A}^j \\ &= \sum_{i \in A^j} \pi_i(\mathbf{n}, t), \end{aligned} \tag{3.11}$$

$$\text{donde } \pi(\mathbf{n}, t) = \frac{\partial}{\partial n_i} DP(\mathbf{n}, t - 1)$$

En otras palabras, este tipo de control define un conjunto de *bid prices* para cada recurso, en cada periodo de tiempo y cada nivel de capacidad disponible, tal que acepta la solicitud por cualquier tipo de producto si y sólo si queda suficiente capacidad y su precio excede la suma de los *bid prices* para todos los recursos usados por ese producto. Esta política de control puede ser implementada mediante una tabla de *bid prices* para cada recurso. La tabla 3.1 ejemplifica lo anterior.

Una observación interesante acerca de los *bid prices* es que no todos los métodos generan los mismos precios umbrales, por ejemplo, el método con gradiente no genera los mismos valores que $BP_j(\mathbf{n}, t)$. Esto significa que se pueden generar instancias de *bid prices* que den una política óptima de decisión pero que al mismo tiempo sean muy malas aproximaciones del valor esperado marginal de la capacidad de un recurso.

Capacidad	t	$t + 1$	$t + 2$...
N	147.3	146.5	145.5	...
N-1	149.5	148.6	147.7	...
N-2	151.8	150.7	150.3	...
...

Tabla 3.1. Ejemplo de cómo implementar *bid price* para un único recurso, en función de capacidad restante y el tiempo.

Esto se puede ver en el siguiente ejemplo, se considera un problema de un solo recurso, en el cual los clientes con alta disposición a pagar llegan estrictamente antes que los de poca disposición a pagar. En este caso, es claro que la política óptima es aceptar a los clientes que primero llegan (*First-come, first-served, FCFS*) hasta agotar la capacidad. Por lo tanto, la solución óptima es un *bid price* con valor constante e igual a cero. Por otro lado, el costo de oportunidad $BP_j(n, t) = BP_j(n, t) - BP_j(n - 1, t)$ en cada periodo de tiempo t será, en general, distinto de cero.

Algunos podrían argumentar que no vale la pena preocuparse por esta diferencia, siempre que existe un *bid price* que sea óptimo para tomar decisiones. Sin embargo, en la práctica se necesita una buena estimación del costo de oportunidad, ya que estos valores a menudo son usados por otros sistemas de RM, tales como fijación de precios (*pricing*) o planificación de la capacidad. Por lo tanto, *bid prices* que aproximan de forma más precisa el costo de oportunidad de usar la capacidad, son más valiosos que aquellos que no lo hacen, incluso si ambos se desempeñan de forma parecida en un sistema de RM.

3.2.2 Aproximaciones basados en modelos de redes

Calcular la función de beneficio esperada $DP(\mathbf{n}, t)$ en (3.9) de forma exacta es muy costoso en términos computacionales, debido a la gran dimensionalidad del espacio de estados (por ejemplo, una red con $m = 20$ recursos cuyas capacidades son $N = 100$ en cada avión, tiene

100²⁰ estados totales). Debido a esto surgen distintos métodos que se aproximan al óptimo encontrado mediante $DP(\mathbf{n}, t)$. Una estrategia posible es plantear el problema como un modelo de programación estático. Estos son análogos deterministas del problema dinámico estocástico, que utilizan sólo la información esperada de la demanda, y suelen ser mucho más simples de resolver.

El resultado más útil de un método aproximado, son las estimaciones de los costos de oportunidad – o *bid prices*. Estos son usados directamente en la política de control o en otros sistemas, como *booking limits* o *virtual nesting*. Entonces, dado un método aproximado M que estima el valor de la función $DP(\mathbf{n}, t)$ mediante $V^M(\mathbf{n}, t)$, se puede aproximar el costo de oportunidad de aceptar el producto j por:

$$V^M(\mathbf{n}, t) - V^M(\mathbf{n} - \mathbf{A}^j, t) \approx (\nabla V^M(\mathbf{n}, t))^\top \mathbf{A}^j,$$

donde $\nabla V^M(\mathbf{n}, t)$ es el gradiente del método aproximado $V^M(\mathbf{n}, t)$, asumiendo que este gradiente existe. Si el gradiente no existe o la aproximación es discreta, entonces se usa la diferencia $V^M(\mathbf{n}, t) - V^M(\mathbf{n} - \mathbf{A}^j, t)$. De esta forma se encuentran buenas aproximaciones de *bid prices* y que además, pueden ser más rápidos de obtener en comparación con $DP(\mathbf{n}, t)$. A continuación se explica uno de los métodos aproximados que se usa en la práctica.

3.2.2.1 Modelo de programación lineal determinista

Se define D_j como la demanda agregada por venir en el tiempo t (demanda para los periodos $t, t - 1, \dots, 0$) para cada producto j con media μ_j . Sea $\mathbf{D} = (D_1, \dots, D_m)$ y $\boldsymbol{\mu} = E[\mathbf{D}]$ el vector de demandas y demandas medias, respectivamente. El modelo de programación lineal determinista (DLP) usa la siguiente aproximación

$$V^{LP}(\mathbf{n}, t) = \max_{\mathbf{y}} \mathbf{R}^\top \mathbf{y} \quad \text{s.a.} \quad \begin{aligned} \mathbf{A}\mathbf{y} &\leq \mathbf{n} \\ 0 &\leq \mathbf{y} \leq \boldsymbol{\mu}. \end{aligned} \quad (3.12)$$

Las variables de decisión $\mathbf{y} = (y_1, \dots, y_m)$ representan la asignación óptima de capacidad para cada uno de los m productos. Esta aproximación trata a la demanda como si fuera determinista e igual a su valor esperado μ y hace una partición óptima de acuerdo a esto. Usando la desigualdad de Jensen, se puede demostrar que $V^{LP}(\mathbf{n}, t)$ es una cota superior de la función óptima $DP(\mathbf{n}, t)$ [18].

Ocasionalmente, la solución \mathbf{y}^* óptima de (3.12) se usa directamente para generar una política de control del tipo *partitioned booking limits*, sin embargo, a menudo esta solución es descartada y sólo se usan las variables óptimas duales, denominadas $\boldsymbol{\pi}^{LP}$ (BPC), asociadas a las restricciones $\mathbf{A}\mathbf{y} \leq \mathbf{n}$, como *bid prices*. Si la solución es no degenerada, es decir, si las restricciones activas son linealmente independientes en la solución óptima, entonces $\nabla V^{LP}(\mathbf{n}, t)$ existe y consiste en un vector único de precios óptimo duales $\boldsymbol{\pi}^{LP}$; si la solución óptima es degenerada, entonces existen múltiples vectores de precios óptimos duales, por lo que hay que usar otro método para estimarlos.

La gran ventaja del modelo DLP es que es muy eficiente de resolver en términos computacionales. Debido a su simplicidad y rapidez, es muy popular en la práctica. Por otro lado, la desventaja de esta aproximación es que sólo considera la demanda promedio, ignorando toda la incertidumbre de la predicción. Una consecuencia de esto es que las variables duales serán cero en cualquier recurso que tiene demanda promedio menor a la capacidad, este comportamiento puede provocar problemas en la práctica.

Una variación en el uso de DLP, propuesto por Bertsimas y Popescu [13], es no usar los precios duales de este modelo y en vez de esto calcular diferencias discretas como estimación del valor marginal de la capacidad. Específicamente, para cada producto j , se calcula lo siguiente

$$V^{LP}(\mathbf{n}, t) - V^{LP}(\mathbf{n} - \mathbf{A}^j, t)$$

y luego usar este valor como *bid price* para el producto o clase j . La motivación de esta propuesta, denominada *Certainty equivalent control* (CEC), es eliminar el problema de degeneración y además, para reflejar de una mejor manera los cambios discretos en la capacidad,

que ocurren al aceptar un cliente. En experimentos numéricos, esta modificación obtiene significativas mejoras sobre los precios duales π^{LP} (la figura 3.5 ilustra este hecho), sin embargo, la desventaja es que para determinar una política de control para los m productos, se deben resolver distintos problemas lineales para cada producto j , lo cual aumenta considerablemente el esfuerzo computacional en relación al DLP con precios duales.

Las figuras 3.4 y 3.5 comparan los distintos modelos analizados en este capítulo. Las líneas representan la función de beneficio esperado acumulado (*Revenues*) en un determinado tiempo. La instancia fue tomada del trabajo de Bertsimas y Popescu [13], para una red con dos recursos, $\mathbf{N} = (50, 50)$, y tres productos (AB, AC y BC), como se ilustra en la figura 3.3.

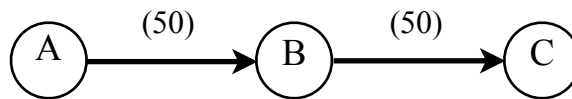


Figura 3.3. Red con tres nodos y dos recursos.

La primera figura 3.4 considera 200 periodos de tiempo, mientras que la segunda figura 3.5 es una ampliación de la primera, desde el tiempo 80 en adelante, debido a que hasta este t todos los modelos se comportan de forma similar.

Como se puede observar, no existe una diferencia tan grande entre los modelos, hasta el tiempo $T = 150$ los modelos aproximados (DLP, CEC, BPC) se acercan bastante al óptimo encontrado mediante programación dinámica (DP). Sin embargo, la aproximación más cercana es mediante el modelo determinista (3.12) usando la política de control de asientos denominada *Certainty equivalent control*, mientras que la política menos eficiente es BPC (*Bid price control* usando variables óptimas duales).

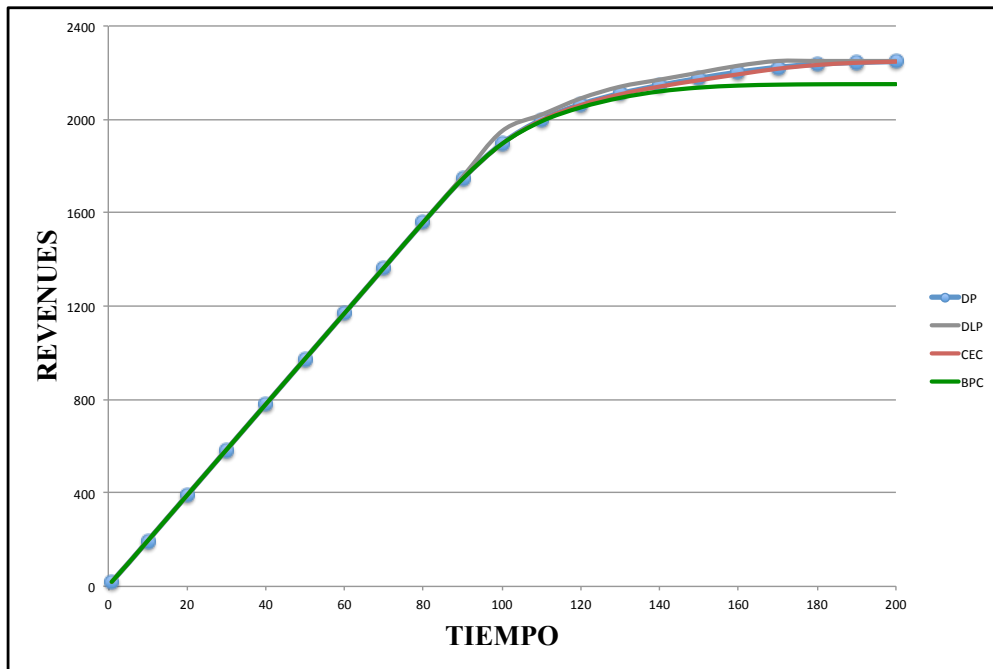


Figura 3.4 Función de beneficio esperado DP, DLP, CEC y BPC. $N = (50, 50)$; $R = (25, 20, 35)$; $p = (0.4, 0.3, 0.1)$; $T = 0 - 200$.

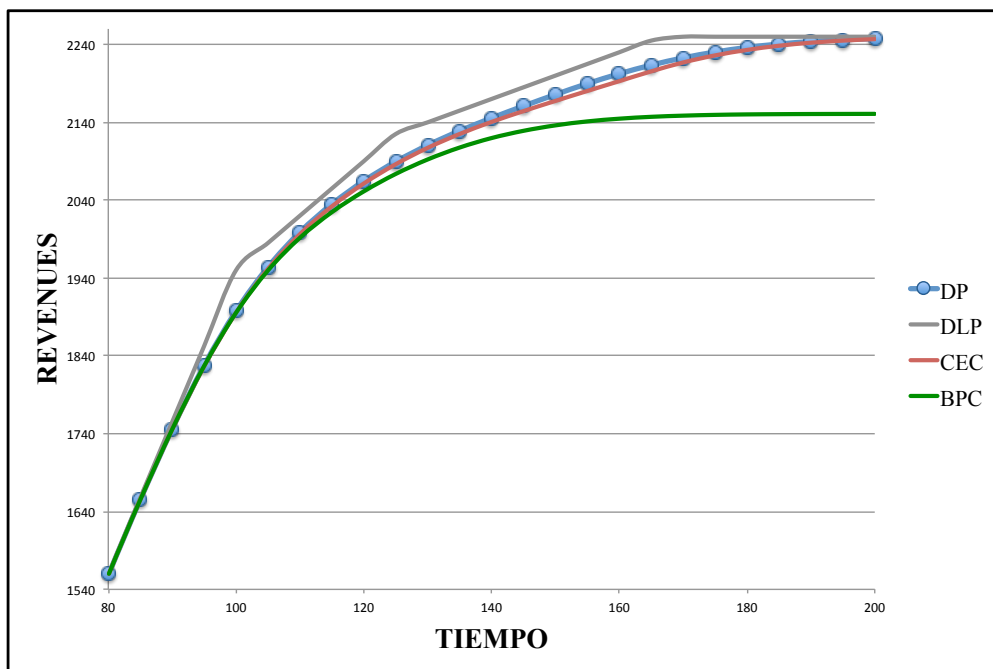


Figura 3.5 Función de beneficio esperado DP, DLP, CEC y BPC. $N = (50, 50)$; $R = (25, 20, 35)$; $p = (0.4, 0.3, 0.1)$; $T = 80 - 200$.

CAPÍTULO 4

MODELO DE PROGRAMACIÓN DINÁMICA MEDIANTE OPTIMIZACIÓN ROBUSTA

En los modelos matemáticos de optimización, comúnmente se asume que las entradas de datos se conocen con precisión y pasan por alto la influencia de la incertidumbre de los parámetros de la optimización y viabilidad de los modelos. Por lo tanto, es esperable que si los datos difieren de los supuestos valores nominales, la solución óptima puede violar restricciones críticas y tener un mal desempeño desde el punto de vista de la función objetivo. Estas observaciones motivan el desarrollo de nuevos modelos de optimización que sean ‘inmunes’ a la incertidumbre de los datos.

La optimización robusta aborda el tema de la incertidumbre de los datos desde la perspectiva del esfuerzo computacional. En la última década, hubo un considerable desarrollo en la teoría de optimización convexa robusta. Sin embargo, en el marco robusto de la literatura, estos modelos en general conducen a un aumento en la complejidad computacional en comparación con el problema original, lo cual significa un gran desafío cuando se quiere resolver problemas grandes.

4.1 Motivación

La incertidumbre de los datos está presente en muchos problemas de optimización del mundo real. Por ejemplo, en aerolíneas, hoteles o cadenas de suministro, la demanda real de productos, rendimientos financieros, las necesidades reales de materiales y otros recursos no se conocen con precisión cuando las decisiones importantes deben tomarse. En la ingeniería y la ciencia, los datos están sometidos a errores de medición, que también constituyen fuentes de incertidumbre en el modelo de optimización.

En la optimización matemática, por lo general se asume que los datos se conocen con precisión. A continuación, se minimiza (o maximiza) una función objetivo sobre un conjunto de variables de decisión de la siguiente manera:

$$\begin{aligned} \min \quad & f_0(\mathbf{x}, \mathbf{D}_0) \\ \text{s.a.} \quad & f_i(\mathbf{x}, \mathbf{D}_i) \geq 0 \quad \forall i \in I, \end{aligned} \tag{4.1}$$

donde \mathbf{x} es el vector de las variables de decisión y \mathbf{D}_i , $i \in I$ son los datos que son parte de las entradas del problema de optimización.

Cuando los parámetros en la función objetivo son inciertos, es poco probable alcanzar el valor óptimo. Es por eso que el grado de variación en la función objetivo es a menudo un motivo de preocupación. Algunos tomadores de decisiones están dispuestos a sacrificar una solución cercana al óptimo nominal por otra solución que garantice un valor objetivo deseado.

Un método clásico para hacer frente a la incertidumbre de los parámetros, es el análisis de sensibilidad. Aquí, el tomador de decisiones ignora la influencia de la incertidumbre de los datos en sus modelos, y posteriormente, realiza análisis de sensibilidad para justificar sus soluciones. Sin embargo, el análisis de sensibilidad es sólo una herramienta para el análisis de la calidad de una solución. No es particularmente útil para generar soluciones que sean robustas frente a variaciones de los datos. Además, no es una herramienta práctica cuando se necesitan analizar

modelos con gran número de parámetros inciertos, como es el caso de programación dinámica, donde en cada estado y etapa se resuelve un problema de optimización.

En vista de estas dificultades, es necesario un buen manejo de la incertidumbre en los parámetros, con el fin de obtener soluciones robustas óptimas que no signifiquen un gran costo en términos de esfuerzo computacional.

A continuación se explicará en detalle la teoría que hay detrás de programación dinámica robusta para un horizonte finito y cómo plantear la ecuación de Bellman cuando se considera el enfoque *maximin*, luego de esto se explicarán algunos conjuntos que modelan la ambigüedad de las probabilidades estimadas, basadas en entropía relativa, y en el siguiente capítulo se realizarán simulaciones tomando en cuenta este modelo.

4.2 Programación dinámica robusta para un horizonte finito

La programación dinámica (DP) esta relacionada con la toma secuencial de decisiones en entornos inciertos. Estas decisiones se toman en etapas y cada decisión, además de ofrecer un beneficio inmediato, cambia el contexto de las decisiones futuras, en consecuencia, esto afecta también los beneficios futuros.

Debido a la naturaleza incierta del entorno, existe escasa información acerca del beneficio inmediato y el estado en el que se encontrará el sistema después de haber tomado una decisión. Por lo tanto, para lograr un buen desempeño en todas las etapas de la toma de decisiones, se tiene que evaluar si es preferible recibir el beneficio de forma inmediata o esperar los pagos futuros.

La incertidumbre del entorno es modelada por un proceso de Markov, cuyas probabilidades de transición dependen tanto del la información del estado como de la decisión escogida. Estas probabilidades de transición que corresponden a cada par estado-acción, se asumen conocidas para el tomador de decisiones, y el objetivo de este es elegir una política (conjunto de decisiones) que maximice la función de beneficio esperada.

Típicamente, las probabilidades de transición del proceso de Markov se estiman a partir de datos históricos, por consiguiente, están sujetas a errores estadísticos. La mayoría de las veces, estos errores son ignorados y la política óptima se calcula suponiendo que la estimación es, de hecho, la probabilidad de transición verdadera. Sin embargo, la política óptima del DP es bastante sensible a perturbaciones en la probabilidad de transición y el ignorar estos errores, puede llevar a resultados incorrectos (Nilim y El Ghaoui [31]). Debido a estos problemas, surge la necesidad de contar con modelos robustos que se hagan cargo de estas probabilidades de transición.

Con el fin de mitigar los errores de estimación, se asume que las probabilidades de cada par estado-acción no se conocen con exactitud, es decir, el modelo de incertidumbre es ‘incierto’. Esta ambigüedad es modelada asociando un conjunto $P(s, a)$ de medidas condicionales a cada par estado-acción (s, a) . En consecuencia, cada política tiene un conjunto de medidas asociadas a ella.

El valor de una política es el beneficio mínimo esperado sobre el conjunto de medidas asociadas, y el objetivo del tomador de decisiones es elegir una política con un valor máximo, es decir, se adopta el enfoque *maximin*, más conocido como *robust DP* (RDP). A continuación se describe matemáticamente el modelo. Como referencia se puede ver el trabajo de Iyengar [28].

4.2.1 Descripción del modelo

Las decisiones se toman de forma discreta en el tiempo $t \in T = \{0, 1, \dots\}$ que corresponden a etapas de decisión. Se asume que T es finito, es decir, $T = \{0, \dots, N - 1\}$ para algún $N \geq 1$. En cada etapa $t \in T$, el sistema ocupa un estado $s \in S_t$, donde S_t se asume que es discreto. En un estado S_t se escoge una acción $a \in A_t(s)$, donde $A_t(s)$ se asume que es discreto.

Para cualquier conjunto discreto B , se define el conjunto de medidas de probabilidad en B como $M(B)$. Los tomadores de decisiones pueden elegir acciones de forma aleatoria o determinista. Una acción aleatoria en el estado $s \in S_t$, es un elemento q_s del conjunto $M(A(s))$, lo

que significa que la acción $a \in A(s)$ es seleccionada con probabilidad $q_s(a)$. En el caso determinista, se asigna toda la masa de probabilidad a una sola acción.

Se define el conjunto de medidas condicionales como $P_t(s, a) \subseteq M(S_{t+1})$ asociado a cada etapa $t \in T$ y par estado-acción (s, a) , $a \in A_s$, $s \in S_t$, con la siguiente interpretación: si en la etapa t se escogió la acción a en el estado s , el siguiente estado $s + 1$ en la etapa $t + 1$ estará determinado por alguna medida condicional $p_{sa} \in P_t(s, a)$. De esta forma, la transición de estado es ambigua.

Se recibe el beneficio $r_t(s_t, a_t, s_{t+1})$ cuando la acción $a_t \in A(s_t)$ es elegida en el estado $s_t \in S$, etapa t y el estado en la siguiente etapa es $s_{t+1} \in S$. Como el estado s_{t+1} es ambiguo, se permite que la función de beneficio en t dependa de s_{t+1} .

Una regla de decisión d_t es el método para seleccionar acciones en cada estado y en una etapa determinada $t \in T$. Se define la regla de decisión histórica dependiente, si depende de toda la historia pasada del sistema representado por la secuencia de estados y acciones pasadas, es decir, d_t es una función de la historia $h_t = (s_0, a_0, \dots, s_{t-1}, a_{t-1}, s_t)$. Se define \mathcal{H}_t como el conjunto de todas las historias h_t . La regla de decisión d_t es determinista si asigna toda la masa de probabilidad en una sola acción $a \in A(s_t)$, y es markoviana si sólo está en función del estado actual s_t .

El conjunto de todas las medidas condicionales consistentes con una regla de decisión determinista y markoviana viene dada por:

$$\mathcal{T}^{d_t} = \{\mathbf{p} : S_t \rightarrow M(S_{t+1}) : \forall s \in S_t, \mathbf{p}_s \in P_t(s, d_t(s))\} \quad (4.3)$$

es decir, para cada estado $s \in S$, el siguiente estado puede ser determinado por cualquier $p \in P_t(s, d_t(s))$.

El conjunto de todas las medidas condicionales consistentes con una regla de decisión histórica dependiente, viene dada por:

$$\mathcal{T}^{d_t} = \left\{ \mathbf{p} : \mathcal{H}_t \rightarrow M(A(s_t) \times S_{t+1}) : \forall h \in \mathcal{H}_t, \begin{array}{l} \mathbf{p}_h(a, s) = q_{d_t(h)}(a) p_{s_t a}(s), \\ p_{s_t a} \in P(s_t, a), a \in A(s_t), s \in S_{t+1} \end{array} \right\} \quad (4.4)$$

Una política determina una regla de decisión a ser usada en todo el periodo de decisiones. De esta forma, una política π es una secuencia de reglas de decisión, es decir, $\pi = (d_t: t \in T)$. Dada la ambigüedad de las medidas condicionales, una política π induce una colección de medidas en el espacio histórico \mathcal{H}_n . Se asume que el conjunto \mathcal{T}^π de medidas consistente con la política π tiene la siguiente estructura:

Supuesto 1 (Ortogonalidad) *El conjunto \mathcal{T}^π de medidas consistente con una política π está dado por*

$$\begin{aligned} \mathcal{T}^\pi &= \left\{ \mathbf{P} : \forall h_N \in \mathcal{H}_N, \mathbf{P}(h_N) = \prod_{t \in T} \mathbf{p}_{h_t}(a_t, s_{t+1}), \mathbf{p}_{h_t} \in \mathcal{T}^{d_t}, t \in T \right\}, \\ &= \mathcal{T}^{d_0} \times \mathcal{T}^{d_1} \times \dots \times \mathcal{T}^{d_{N-1}}, \end{aligned} \quad (4.5)$$

donde la notación en (4.5) indica que cada $p \in \mathcal{T}^\pi$ es el producto de $p_t \in \mathcal{T}^{d_t}$.

El supuesto de ortogonalidad tiene que ver con que las medidas condicionales son independientes entre sí cuando se toma una decisión, es decir, la elección particular de $p \in P(s_t, a_t)$ en el par estado-acción (s_t, a_t) en t , no limita las acciones en el futuro.

El beneficio $V_0^\pi(s)$ generado por la política π a partir del estado inicial $s_0 = s$, se define como sigue

$$V_0^\pi(s) = \min_{\mathbf{P} \in \mathcal{T}^\pi} \mathbf{E}^{\mathbf{P}} \left[\sum_{t \in T} r_t(s_t, d_t(h_t), s_{t+1}) + r_N(s_N) \right], \quad (4.6)$$

donde $\mathbf{E}^{\mathbf{P}}$ denota la esperanza con respecto a la medida $\mathbf{P} \in \mathcal{T}^{\pi}$. La ecuación (4.6) define el beneficio de una política π como el mínimo beneficio esperado sobre el conjunto de medidas condicionales \mathcal{T}^{π} , es decir, se adopta un enfoque del peor caso. En la literatura de optimización este problema es conocido como el enfoque robusto (Ben-Tal y Nemirovski [7]).

Sea Π el conjunto histórico de todas las políticas dependientes, entonces el objetivo del *robust DP* es encontrar el valor de la siguiente función

$$V_0^*(s) = \max_{\pi \in \Pi} \{V_0^{\pi}(s)\} = \max_{\pi \in \Pi} \left\{ \min_{\mathbf{P} \in \mathcal{T}^{\pi}} \mathbf{E}^{\mathbf{P}} \left[\sum_{t \in T} r_t(s_t, d_t(h_t), s_{t+1}) + r_N(s_N) \right] \right\}, \quad (4.7)$$

la política óptima es π^* si se alcanza el máximo.

Con el fin de apreciar las implicaciones del supuesto 1 sobre ortogonalidad, la función objetivo (4.7) se puede interpretar desde el punto de vista de teoría de juegos: existen dos jugadores, el jugador (1) escoge una política π tal que maximiza su beneficio, mientras que el jugador (2) observa la jugada de (1) y escoge una medida $\mathbf{P} \in \mathcal{T}^{\pi}$ que minimiza el beneficio de (1). En este contexto, la ortogonalidad implica que la elección particular de $p \in P(s_t, a_t)$ en el par estado-acción (s_t, a_t) en t , no limita las acciones del adversario en el futuro. Esto, a su vez, conduce a una propiedad de separabilidad que es crucial para poder plantear la ecuación de Bellman del problema robusto.

Sea $V_n^{\pi}(h_n)$ el beneficio obtenido por usar la política π en los periodos $n, n+1, \dots, N-1$, a partir de la historia h_n , es decir,

$$V_n^{\pi}(h_n) = \min_{\mathbf{P} \in \mathcal{T}_n^{\pi}} \mathbf{E}^{\mathbf{P}} \left[\sum_{t=n}^{N-1} r_t(s_t, d_t(h_t), s_{t+1}) + r_N(s_N) \right], \quad (4.8)$$

donde la ortogonalidad implica que el conjunto de medidas condicionales \mathcal{T}_n^{π} consistentes con la política π y la historia h_n está dado por

$$\begin{aligned} \mathcal{T}_n^\pi &= \left\{ \mathbf{P}_n : \mathcal{H}_t \rightarrow \prod_{t=n}^{N-1} (A_t \times S_{t+1}) : \forall h \in \mathcal{H}_t, \mathbf{P}_{h_n}(a_n, s_{n+1}, \dots, a_{N-1}, s_N) = \prod_{t=n}^{N-1} \mathbf{p}_{h_t}(a_t, s_{t+1}), \right. \\ &= \mathcal{T}^{d_n} \times \mathcal{T}^{d_{n+1}} \times \dots \times \mathcal{T}^{d_{N-1}}, \\ &= \mathcal{T}^{d_n} \times \mathcal{T}_{n+1}^\pi. \end{aligned} \quad (4.9)$$

Sea $V_n^*(h_n)$ el beneficio óptimo a partir de la historia h_n en la etapa n , es decir,

$$V_n^*(h_n) = \max_{\pi \in \Pi_n} \{V_n^\pi(h_n)\} = \max_{\pi \in \Pi_n} \left\{ \min_{\mathbf{P} \in \mathcal{T}_n^\pi} \mathbf{E}^{\mathbf{P}} \left[\sum_{t=n}^{N-1} r_t(s_t, d_t(h_t), s_{t+1}) + r_N(s_N) \right] \right\}, \quad (4.10)$$

donde Π_n es el conjunto de todas las políticas para las etapas $t \geq n$.

Teorema 1 (Ecuación de Bellman) *El conjunto de funciones $\{V_n^* : n = 0, 1, \dots, N\}$ satisfacen la siguiente ecuación de Bellman:*

$$\begin{aligned} V_N^*(h_N) &= r_N(s_N) \\ V_n^*(h_n) &= \max_{a \in A(s_n)} \left\{ \min_{p \in P(s_n, a)} \mathbf{E}^p [r_n(s_n, a, s) + V_{n+1}^*(h_n, a, s)] \right\}, \quad n = 0, \dots, N-1 \end{aligned} \quad (4.11)$$

Demostración: de (4.9) se deduce que

$$V_n^*(h_n) = \max_{\pi \in \Pi} \left\{ \min_{\mathbf{P}=(\mathbf{p}, \bar{\mathbf{P}}) \in \mathcal{T}^{d_n} \times \mathcal{T}_{n+1}^\pi} \mathbf{E}^{\mathbf{P}} \left[\sum_{t=n}^{N-1} r_t(s_t, d_t(h_t), s_{t+1}) + r_N(s_N) \right] \right\},$$

Dado que las medidas condicionales $\bar{\mathbf{P}}$ no afectan al primer término $r_n(s_n, d_n(h_n), s_{n+1})$, entonces se cumple que:

$$\begin{aligned} V_n^*(h_n) &= \max_{\pi \in \Pi_n} \left\{ \min_{\mathbf{P}=(\mathbf{p}, \bar{\mathbf{P}}) \in \mathcal{T}^{d_n} \times \mathcal{T}_{n+1}^\pi} \mathbf{E}^{\mathbf{P}} \left[r_n(s_n, d_n(h_n), s_{n+1}) + \mathbf{E}^{\bar{\mathbf{P}}} \left[\sum_{t=n+1}^{N-1} r_t(s_t, d_t(h_t), s_{t+1}) + r_N(s_N) \right] \right] \right\}, \\ &= \max_{\pi \in \Pi_n} \left\{ \min_{\mathbf{p} \in \mathcal{T}^{d_n}} \mathbf{E}^{\mathbf{p}} \left[r_n(s_n, d_n(h_n), s_{n+1}) + \min_{\bar{\mathbf{P}} \in \mathcal{T}_{n+1}^\pi} \mathbf{E}^{\bar{\mathbf{P}}} \left[\sum_{t=n+1}^{N-1} r_t(s_t, d_t(h_t), s_{t+1}) + r_N(s_N) \right] \right] \right\}, \\ &= \max_{\pi \in \Pi_n} \left\{ \min_{\mathbf{p} \in \mathcal{T}^{d_n}} \mathbf{E}^{\mathbf{p}} \left[r_n(s_n, d_n(h_n), s_{n+1}) + V_{n+1}^\pi(h_n, d_n(h_n), s_{n+1}) \right] \right\}, \end{aligned} \quad (4.12)$$

donde la última igualdad se deduce de la definición $V_{n+1}^\pi(h_{n+1})$ en (4.8).

Sea $(d_n(h_n)(w), s_{n+1}(w))$ el par acción-estado correspondiente a una regla de decisión d_n aleatoria. Por lo tanto, el beneficio generado por esta regla decisión partiendo del estado s_{n+1} cumple la siguiente desigualdad

$$V_{n+1}^\pi(h_n, d_n(h_n)(w), s_{n+1}(w)) \leq V_{n+1}^*(h_n, d_n(h_n)(w), s_{n+1}(w))$$

Tomando en cuenta lo anterior, la ecuación (4.12) conlleva a lo siguiente

$$\begin{aligned} V_n^*(h_n) &\leq \max_{\pi \in \Pi_n} \left\{ \min_{\mathbf{p} \in \mathcal{T}^{d_n}} \mathbf{E}^{\mathbf{P}} [r_n(s_n, d_n(h_n), s_{n+1}) + V_{n+1}^*(h_n, d_n(h_n), s_{n+1})] \right\}, \\ &= \max_{d_n \in D_n} \left\{ \min_{\mathbf{p} \in \mathcal{T}^{d_n}} \mathbf{E}^{\mathbf{P}} [r_n(s_n, d_n(h_n), s_{n+1}) + V_{n+1}^*(h_n, d_n(h_n), s_{n+1})] \right\}, \end{aligned} \quad (4.13)$$

donde D_n es el conjunto de todas las decisiones históricas hasta el tiempo n . La última igualdad de (4.13) es debido a que la esperanza depende de la regla de decisión d_n y no de la política π .

Por otra parte, de (4.10) se sabe que $V_{n+1}^*(h_{n+1}) = \max_{\pi \in \Pi_{n+1}} \{V_{n+1}^\pi(h_{n+1})\}$, luego para todo $\varepsilon > 0$ existe una política $\pi_{n+1}^\varepsilon \in \Pi_{n+1}$ tal que

$$V_{n+1}^{\pi_{n+1}^\varepsilon}(h_{n+1}) \geq V_{n+1}^*(h_{n+1}) - \varepsilon, \quad \forall h_{n+1} \in \mathcal{H}_{n+1}$$

luego para todo $d_n \in D_n$, $(d_n, \pi_{n+1}^\varepsilon) \in \Pi_n$ se cumple lo siguiente

$$\begin{aligned} V_n^*(h_n) &= \max_{\pi \in \Pi_n} \left\{ \min_{\mathbf{p} \in \mathcal{T}^{d_n}} \mathbf{E}^{\mathbf{P}} [r_n(s_n, d_n(h_n), s_{n+1}) + V_{n+1}^\pi(h_n, d_n(h_n), s_{n+1})] \right\}, \\ &\geq \max_{d_n \in D_n} \left\{ \min_{\mathbf{p} \in \mathcal{T}^{d_n}} \mathbf{E}^{\mathbf{P}} [r_n(s_n, d_n(h_n), s_{n+1}) + V_{n+1}^{\pi_{n+1}^\varepsilon}(h_n, d_n(h_n), s_{n+1})] \right\}, \\ &\geq \max_{d_n \in D_n} \left\{ \min_{\mathbf{p} \in \mathcal{T}^{d_n}} \mathbf{E}^{\mathbf{P}} [r_n(s_n, d_n(h_n), s_{n+1}) + V_{n+1}^*(h_n, d_n(h_n), s_{n+1})] \right\} - \varepsilon. \end{aligned} \quad (4.14)$$

dado que $\varepsilon > 0$ es arbitrario, se puede hacer tender a cero y por teorema del sándwich, (4.13) y (4.14) implican lo siguiente

$$V_n^*(h_n) = \max_{d_n \in D_n} \left\{ \min_{\mathbf{p} \in \mathcal{T}^{d_n}} \mathbf{E}^{\mathbf{P}} \left[r_n(s_n, d_n(h_n), s_{n+1}) + V_{n+1}^*(h_n, d_n(h_n), s_{n+1}) \right] \right\}$$

La definición de \mathcal{T}^{d_n} en (4.4) implica que $V_n^*(h_n)$ se puede escribir de la siguiente forma

$$\begin{aligned} V_n^*(h_n) &= \max_{q \in M(A(s_n))} \min_{p_{s_n a} \in P_n(s_n, a)} \left\{ \sum_{a \in A(s_n)} q(a) \left[\sum_{s \in S} p_{s_n a} \left[r_n(s_n, a, s) + V_{n+1}^*(h_n, a, s) \right] \right] \right\}, \\ &= \max_{q \in M(A(s_n))} \left\{ \sum_{a \in A(s_n)} q(a) \min_{p_{s_n a} \in P_n(s_n, a)} \left[\sum_{s \in S} p_{s_n a} \left[r_n(s_n, a, s) + V_{n+1}^*(h_n, a, s) \right] \right] \right\}, \\ &= \max_{a \in A(s_n)} \left\{ \min_{p \in P_n(s_n, a)} \left[\sum_{s \in S} p(s) \left[r_n(s_n, a, s) + V_{n+1}^*(h_n, a, s) \right] \right] \right\}, \\ &= \max_{a \in A(s_n)} \left\{ \min_{p \in P_n(s_n, a)} \mathbf{E}^p \left[r_n(s_n, a, s) + V_{n+1}^*(h_n, a, s) \right] \right\}, \end{aligned} \quad (4.15)$$

■

Proposición 1 (Adaptación de la cadena de Markov) *Si se considera un conjunto de acciones contable y una distribución del estado inicial que toma valores en un conjunto numerable, entonces la distribución de probabilidad de cada par (s_k, a_k) y el beneficio esperado correspondiente a una política aleatoria histórica dependiente, también pueden ser obtenidos mediante una política de Markov aleatoria.*

Demostración: Sea $\pi = \{d_0, d_1, \dots\}$ una política histórica dependiente y además $\zeta_k(s_k), \zeta_k(s_k, a_k)$ las distribuciones de s_k y (s_k, a_k) , respectivamente. Sea $\bar{\pi} = \{\bar{d}_0, \bar{d}_1, \dots\}$ la política de Markov aleatoria, donde la regla de decisión \bar{d}_k está definida para todo estado s_k , con $\zeta_k(s_k) > 0$ por

$$\bar{d}_k(a_k | s_k) = \frac{\zeta_k(s_k, a_k)}{\xi_k(s_k)}$$

Sea $\bar{\xi}_k(s_k)$ y $\bar{\zeta}_k(s_k, a_k)$ las distribuciones de s_k y (s_k, a_k) , respectivamente. Se demostrará por inducción que para todo k, s_k y a_k , se cumple que

$$\xi_k(s_k) = \bar{\xi}_k(s_k), \quad \zeta_k(s_k, a_k) = \bar{\zeta}_k(s_k, a_k). \quad (4.16)$$

Para $k = 0$, las distribuciones $\xi_0(s_0)$ y $\bar{\xi}_0(s_0)$ son iguales en el estado inicial, además se cumple lo siguiente

$$\bar{\zeta}_0(s_0, a_0) = \bar{\xi}_0(s_0)\bar{d}_0(a_0|s_0) = \bar{\xi}_0(s_0)\frac{\zeta_k(s_0, a_0)}{\xi_0(s_0)} = \zeta_k(s_0, a_0).$$

Para aplicar el principio de inducción, se asume que la ecuación (4.16) es válida para algún k . Luego para $k + 1$ se tiene que

$$\begin{aligned} \bar{\xi}_{k+1}(s_{k+1}) &= \sum_{s_k, a_k} \bar{\zeta}_k(s_k, a_k) p_{s_k a_k}(s) \\ &= \sum_{s_k, a_k} \bar{\xi}_k(s_k) \bar{d}_k(a_k|s_k) p_{s_k a_k}(s) \\ &= \sum_{s_k, a_k} \bar{\xi}_k(s_k) \frac{\zeta_k(s_k, a_k)}{\xi_k(s_k)} p_{s_k a_k}(s) \\ &= \sum_{s_k, a_k} \zeta_k(s_k, a_k) p_{s_k a_k}(s) \\ &= \xi_{k+1}(s_{k+1}) \end{aligned}$$

donde $p_{s_k a_k}(s)$ son las probabilidades de transición del sistema, y la suma es sobre todos los pares (s_k, a_k) tales que $\zeta_k(s_k, a_k) > 0$. Además,

$$\begin{aligned} \bar{\zeta}_{k+1}(s_{k+1}, a_{k+1}) &= \bar{\xi}_{k+1}(s_{k+1}) \bar{d}_k(a_{k+1}|s_{k+1}) \\ &= \bar{\xi}_{k+1}(s_{k+1}) \frac{\zeta_{k+1}(s_{k+1}, a_{k+1})}{\xi_{k+1}(s_{k+1})} \\ &= \zeta_{k+1}(s_{k+1}, a_{k+1}), \end{aligned}$$

completando así la inducción. De esta forma, π y $\bar{\pi}$ generan la misma distribución estado-acción. De esto se deduce, además, que los correspondientes beneficios esperados de cada etapa son iguales. ■

La proposición anterior muestra que el costo esperado de cualquier política aleatoria histórica dependiente sobre un horizonte finito puede ser replicada por una política de Markov aleatoria. Esto implica que para un problema de horizonte finito, con seguridad se puede restringir la atención a políticas de Markov, y no es necesario considerar políticas que dependen de la historia.

Luego de haber considerado lo anterior, surge una pregunta interesante, ¿Es posible prescindir de las políticas aleatorias y restringirse sólo a políticas de Markov deterministas? Esto es cierto con mucha frecuencia, es decir, para varios problemas se puede demostrar que el beneficio óptimo al usar políticas aleatorias es el mismo que el beneficio óptimo cuando se usan políticas deterministas, además se puede demostrar que si existe una política óptima (posiblemente aleatoria), entonces existe una política determinista óptima. Como ejemplo, están los problemas con horizonte finito, el problema de costo actualizado con costo limitado por etapa, y los modelos con espacios finitos de estados y acciones. Todos estos problemas, incluso se pueden resolver mediante políticas deterministas de Markov *estacionarias*. Estos ejemplos y demostraciones se pueden ver con más detalle en el libro del autor Bertsekas, D. [11].

Por consiguiente, sea Π_D el conjunto de todas las políticas de Markov deterministas, luego, la función V_n^* en (4.11) puede ser caracterizada mediante Π_D como sigue

$$V_n^*(s_n) = \max_{\pi \in \Pi_D} \{V_n^\pi(s_n)\}, \quad n = 0, \dots, N - 1. \quad (4.17)$$

Tomando en cuenta lo anterior, la ecuación de Bellman propuesta en (4.11) se puede plantear de acuerdo al siguiente teorema.

Teorema 2 (Ecuación de Bellman Markoviana) *Para todo $n = 0, \dots, N$, el valor de la función robusta $V_n^*(h_n)$ depende sólo del estado s_n actual, por lo tanto, la ecuación de Bellman robusta (4.11) se reduce a*

$$V_n^*(s_n) = \max_{a \in A(s_n)} \left\{ \min_{p \in P_n(s_n, a)} \mathbf{E}^p \left[r_n(s_n, a, s) + V_{n+1}^*(s) \right] \right\}, \quad n \in T. \quad (4.18)$$

Demostración: Este resultado se concluye a partir de la proposición 1. Los detalles se dejan al lector. ■

La ecuación recursiva (4.18) constituye la base de la programación dinámica robusta, esta relación establece que, dado $V_{n+1}^*(s')$ para todo $s' \in S$, calcular $V_n^*(s)$ se reduce a una colección de problemas de optimización. Por ejemplo, sea $A(s)$ un conjunto finito, luego la regla de decisión óptima d_n^* en la etapa n se encuentra mediante

$$d_n^*(s) = \arg \max_{a \in A(s_n)} \left\{ \min_{p \in P_n(s_n, a)} \mathbf{E}^p \left[r_n(s_n, a, s') + V_{n+1}^*(s') \right] \right\}.$$

Cada etapa de la programación dinámica implica resolver el problema de optimización $\min_{p \in P(s, a)} \mathbf{E}^p[v]$ donde p corresponde a un elemento del conjunto $P(s, a)$ de medidas condicionales. La forma del conjunto $P(s, a)$ es una componente clave de la complejidad computacional del algoritmo de programación dinámica robusta. Sin embargo, más allá de la complejidad numérica, un criterio adicional para la elección de un modelo de incertidumbre, es que los conjuntos $P(s, a)$ deben representar descripciones precisas de la variabilidad estadística sobre las matrices de transición. Estos conjuntos son los que se analizarán en la siguiente sección.

4.3 Conjuntos manejables de medidas condicionales

En esta sección se analizarán qué conjuntos se pueden considerar para resolver el problema interior de la ecuación (4.18), tomando en cuenta un espacio de estados y acciones finito.

En ausencia de cualquier ambigüedad, el valor de una acción $a \in A(s)$ en el estado $s \in S$ está dado por $\mathbf{E}^p[v] = p^T v$, donde p es la medida condicional y v una variable aleatoria que toma valores en $v(s') = r(s, a, s') + V(s')$ en el estado $s' \in S$. En consecuencia, la complejidad de calcular el valor de un par estado-acción es $\mathcal{O}(|S|)$.

Cuando la medida condicional es ambigua, el valor del par estado-acción (s, a) está dado por

$$\min_{p \in \mathcal{P}(s,a)} v^\top p \quad (4.19)$$

Para resolver el problema anterior, se presentarán cuatro familias de conjuntos de medidas condicionales $P(s, a)$, que involucran regiones de verosimilitud y entropía relativa.

4.3.1 Modelo de Verosimilitud

El primer modelo se basa en restricciones de verosimilitud para describir la incertidumbre de cada matriz de transición. Se denota F como la matriz de frecuencias empíricas de transición, cuya i -ésima fila es f_i . Se tiene que $F \geq 0$ y $F\mathbf{1} = \mathbf{1}$, donde $\mathbf{1}$ representa el vector de unos.

Al modelo se le entrega como dato la estimación de $\hat{P} = F$ que es la solución al problema de máxima-verosimilitud

$$\max_p L(P) = \sum_{i,j} F(i,j) \log P(i,j) : P \geq 0, P\mathbf{1} = \mathbf{1} \quad (4.20)$$

donde el *log-likelihood* óptimo es

$$\beta_{max} = \sum_{i,j} F(i,j) \log F(i,j)$$

Una descripción clásica de la incertidumbre en un entorno de máxima verosimilitud es a través de la región de verosimilitud (Lehmann y Casella 1998, Malo 1988)

$$\left\{ P \in \mathbf{R}^{n \times n} : P \geq 0, P\mathbf{1} = \mathbf{1}, \sum_{i,j} F(i,j) \log P(i,j) \geq \beta \right\} \quad (4.21)$$

donde $\beta < \beta_{max}$ es un valor arbitrario, que representa el nivel de incertidumbre. En la práctica, el tomador de decisiones puede elegir un nivel de incertidumbre y β puede ser estimado mediante

métodos de muestreo, o una muestra grande de aproximaciones gaussianas, con el fin de que el conjunto (4.21) tenga un nivel deseado de confianza.

La descripción anterior es clásica en el sentido que las regiones de máxima verosimilitud son el punto de partida para el desarrollo de modelos elipsoidales o de intervalos de confianza, por lo tanto, son estadísticamente más precisos (Lehmann y Casella, 1998). El conjunto anterior es estadísticamente significativo, ya que describe cuán informativa es la data. Este conjunto tiene algunas características interesantes. En primer lugar, no es el resultado de una aproximación cuadrática, sino que es una descripción válida de la incertidumbre, incluso para valores β que están muy alejados de β_{max} . En segundo lugar, este conjunto no es simétrico en torno al punto de máxima verosimilitud, lo cual implica que la incertidumbre estadística depende de una dirección. Por último, por como está construido el conjunto, excluye aquellas matrices que no son las matrices de transición, lo mismo no puede decirse para aproximaciones elipsoidales.

Para aplicar la recursión robusta, el conjunto P en (4.21) debe cumplir con la propiedad de ortogonalidad. Sin embargo, este conjunto no tiene esta característica, pero se puede aproximar por un conjunto que sí la tiene, mediante la proyección de esta región en n -subespacios correspondientes a las filas de la matriz de transición.

Debido a la naturaleza separable de la función logarítmica de verosimilitud, la proyección del conjunto sobre p_i (es decir, la fila i) de la matriz P está dado por

$$\mathcal{P}_i(\beta_i) = \left\{ p \in \mathbf{R}_+^n, p^\top \mathbf{1} = 1 : \sum_j f_i(j) \log p_i(j) \geq \beta_i \right\}$$

donde

$$\beta_i = \beta - \sum_{k \neq i} \sum_j F(k, j) \log F(k, j).$$

representa el nivel de incertidumbre para la fila i de la matriz de transición P .

Tomando en cuenta lo anterior, el problema planteado en (4.19), es resolver un problema de optimización que tiene la siguiente forma

$$\begin{aligned} \min_p \quad & \sum_i p(i)v(i) \\ \text{s.a.} \quad & \sum_{j \in S} f_i(j) \log p_i(j) \geq \beta_i \quad \forall i \in S \\ & \sum_i p(i) = 1 \end{aligned} \quad (4.22)$$

En esta sección, β_{max} representa el máximo valor de la función verosimilitud que aparece en el conjunto anterior, es decir, $\beta_{max}^i = \sum_{j \in S} f_i(j) \log f_i(j)$. Se asume que $\beta_i < \beta_{max}^i$ que, junto con $f > 0$, asegura que el conjunto anterior tenga interior no vacío.

El problema Dual

El Lagrangiano L asociado al problema (4.22) se puede escribir de la siguiente forma

$$L(v, \mu, \lambda) = \sum_i p(i)v(i) - \lambda \left(\sum_i f(i) \log p(i) - \beta \right) - \mu \left(\sum_i p(i) - 1 \right)$$

donde $\lambda \geq 0$ y $\mu \in \mathbf{R}$, tomando la derivada con respecto a $p(i)$ e igualando a cero se obtiene

$$\frac{\partial L}{\partial p(i)} = v(i) - \lambda f(i) \frac{1}{p(i)} - \mu = 0, \quad i \in S$$

$$p^*(i) = \frac{\lambda f(i)}{v(i) - \mu}$$

Luego la función dual del lagrangiano es la siguiente

$$\begin{aligned} \max_{\lambda, \mu} L &= \max_{\lambda, \mu} \left[\sum_i \frac{\lambda f(i)}{v(i) - \mu} v(i) - \lambda \left(\sum_i f(i) \log \frac{\lambda f(i)}{v(i) - \mu} - \beta \right) - \mu \left(\sum_i \frac{\lambda f(i)}{v(i) - \mu} - 1 \right) \right] \\ &= \max_{\lambda, \mu} \left[\mu + (1 + \beta)\lambda - \lambda \sum_i f(i) \log \frac{\lambda f(i)}{v(i) - \mu} \right], \quad \lambda \geq 0, v - \mu \mathbf{1} > 0 \end{aligned} \quad (4.23)$$

Debido a que el problema anterior es cóncavo, y tiene un conjunto factible no vacío, no hay una brecha dual (*duality gap*) o en otras palabras, existe dualidad perfecta, es decir, los problemas (4.22) y (4.23) generan la misma solución.

Por lo tanto, la ecuación de Bellman robusta en (4.18) se ha simplificado al siguiente problema

$$V_n^*(s_n) = \max_{\lambda, \mu, a \in A(s_n)} \left[\mu + (1 + \beta)\lambda - \lambda \sum_i f(i) \log \frac{\lambda f(i)}{r_n(s_n, a, s) + V_{n+1}^*(s) - \mu} \right]$$

$$\lambda \geq 0, \mu < v_{min} = \min_i v(i) \quad (4.24)$$

4.3.2 Modelo de Entropía Relativa

La entropía es una medida de la incertidumbre de una variable aleatoria; o también se puede entender como la información requerida en promedio para describir la variable aleatoria.

La entropía relativa es una medida de la distancia entre dos distribuciones. En estadística, se define como el valor esperado del logaritmo de la razón de verosimilitud. La entropía relativa $D(p||q)$ es una medida de la ineficiencia de asumir que la distribución es q cuando la verdadera distribución es p . La distancia $D(p||q)$ más conocida como *Kullback-Leibler divergence* se define como

$$D(p \parallel q) = \sum_j p(j) \log \frac{p(j)}{q(j)}$$

por lo tanto, el problema (4.19) se puede describir mediante el siguiente conjunto

$$\mathcal{P} = \{p \in \mathbf{R}_+^n, p^\top \mathbf{1} = 1 : D(p \parallel q) \leq \beta\}$$

donde $\beta > 0$ es un valor arbitrario y $q > 0$ es una distribución dada.

Ahora se abordará el problema interior (4.19) con el conjunto descrito anteriormente, es decir, se resolverá el siguiente problema de optimización

$$\begin{aligned} \min_p \quad & \sum_i p(i)v(i) \\ \text{s.a.} \quad & \sum_{i \in S} p(i) \log \frac{p(i)}{q(i)} \leq \beta \\ & \sum_i p(i) = 1 \end{aligned} \tag{4.25}$$

El problema Dual

El Lagrangiano L asociado al problema (4.25) se puede escribir de la siguiente forma

$$L(v, \mu, \lambda) = \sum_i p(i)v(i) - \lambda \left(\beta - \sum_i p(i) \log \frac{p(i)}{q(i)} \right) - \mu \left(\sum_i p(i) - 1 \right)$$

donde $\lambda \geq 0$ y $\mu \in \mathbf{R}$, tomando la derivada con respecto a $p(i)$ e igualando a cero se obtiene

$$\frac{\partial L}{\partial p(i)} = v(i) + \lambda \left(\log \frac{p(i)}{q(i)} + 1 \right) - \mu = 0, \quad i \in S$$

$$p^*(i) = q(i) \exp \left(\frac{\mu - v(i)}{\lambda} - 1 \right)$$

luego, resolver el problema primal (4.25) es equivalente a resolver su dual:

$$\max_{\lambda > 0, \mu} \mu - \beta \lambda - \lambda \sum_i q(i) \exp \left(\frac{\mu - v(i)}{\lambda} - 1 \right)$$

si se toma la derivada con respecto a μ igual a cero, se obtiene la siguiente condición de optimalidad

$$\sum_i q(i) \exp \left(\frac{\mu - v(i)}{\lambda} - 1 \right) = 1$$

por consiguiente, se deriva que

$$\mu = -\lambda \log \left[\sum_i q(i) \exp \left(\frac{-v(i)}{\lambda} \right) \right] + \lambda$$

entonces, el problema (4.25) se redujo a un problema cóncavo de una sola variable, es decir

$$\max_{\lambda > 0} -\lambda \log \left[\sum_i q(i) \exp \left(\frac{-v(i)}{\lambda} \right) \right] - \beta \lambda \quad (4.26)$$

Luego, el problema de programación dinámica robusta planteado en (4.18) se puede resolver mediante la siguiente recursión

$$V_n^*(s_n) = \max_{\lambda > 0, a \in A(s_n)} -\lambda \log \left[\sum_i q(i) \exp \left(-\frac{r_n(s_n, a, s) + V_{n+1}^*(s)}{\lambda} \right) \right] - \beta \lambda \quad (4.27)$$

4.3.3 Modelo Elipsoidal

Estos modelos surgen de aproximaciones de segundo orden de la función verosimilitud $L(P)$ definida en (4.20). Se desarrolla una serie de Taylor de segundo orden en torno a la estimación de probabilidad F de máxima verosimilitud de la siguiente forma

$$\begin{aligned} Y(P) &= L(F) + L'(F)(P - F) + \frac{1}{2}L''(F)(P - F)^2 \\ &= \sum_{i,j} F(i,j) \log F(i,j) + \sum_{i,j} \frac{F(i,j)}{F(i,j)} (P(i,j) - F(i,j)) - \frac{1}{2} \sum_{i,j} \frac{F(i,j)}{F(i,j)^2} (P(i,j) - F(i,j))^2 \\ &= \beta_{max} - \frac{1}{2} \sum_{i,j} \frac{(P(i,j) - F(i,j))^2}{F(i,j)} \end{aligned}$$

por lo tanto, se considerará esta aproximación para definir un nuevo conjunto dado por

$$\mathcal{P} = \{P \in \mathbf{R}^{n \times n} : P \geq 0, P\mathbf{1} = \mathbf{1}, Y(P) \geq \beta\}$$

El conjunto anterior es un elipsoide interceptado por el politopo de matrices de transición. Este conjunto no cumple con la propiedad de ortogonalidad, sin embargo se puede proyectar esta región a las filas de la matriz de transición, como se hizo en el caso del primer conjunto, específicamente:

$$\mathcal{P} = \left\{ p \in \mathbf{R}_+^n, p^\top \mathbf{1} = 1 : \sum_{i \in S} \frac{(p(i) - f(i))^2}{f(i)} \leq k^2 \right\}$$

donde⁴ $k^2 = 2(\beta_{max} - \beta)$. Por lo tanto, el problema interior (4.19) tiene la siguiente forma

$$\begin{aligned} \min_p \quad & \sum_i p(i)v(i) \\ \text{s.a.} \quad & : \sum_{i \in S} \frac{(p(i) - f(i))^2}{f(i)} \leq k^2 \\ & \sum_i p(i) = 1 \\ & p \geq 0 \end{aligned} \tag{4.28}$$

que es equivalente al dual

$$\max_{\lambda, \mu \geq 0} \sum_{i \in S} f(i)(v(i) - \mu(i)) - \sqrt{k^2 \sum_{i \in S} f(i)(v(i) - \lambda - \mu(i))^2} \tag{4.29}$$

Demostración: Si consideramos el siguiente cambio de variable $y = p - f$ el problema (4.28) se transforma en

$$\begin{aligned} \sum_i f(i)v(i) + \min_p \quad & \sum_i y(i)v(i) \\ \text{s.a.} \quad & \sum_i \frac{y(i)^2}{f(i)} \leq k^2 \\ & \sum_i y(i) = 1 \\ & y \geq -f \end{aligned}$$

⁴ Se ha eliminado el índice i correspondiente a la fila i de la matriz de transición.

Este, a su vez, tiene asociado el siguiente dual:

$$\begin{aligned}
 D &= \sum_i f(i)v(i) + \max_{\lambda, \mu} \min_{\sum_i \frac{y(i)^2}{f(i)} \leq k^2} \left[\sum_i y(i)v(i) - \lambda \sum_i y(i) - \sum_i \mu(i)(y(i) + f(i)) \right] \\
 &= \max_{\lambda, \mu} \sum_i f(i)(v(i) - \mu(i)) + \min_{\sum_i \frac{y(i)^2}{f(i)} \leq k^2} \left[\sum_i y(i)(v(i) - \lambda - \mu(i)) \right] \\
 &= \max_{\lambda, \mu \geq 0} \sum_{i \in S} f(i)(v(i) - \mu(i)) - \sqrt{k^2 \sum_{i \in S} f(i)(v(i) - \lambda - \mu(i))^2}
 \end{aligned}$$

donde la última igualdad se obtiene resolviendo el problema de minimización. Aplicando las condiciones de Karush–Kuhn–Tucker (KKT) se obtiene que

$$\begin{aligned}
 y(i) &= -\frac{v(i) - \lambda - \mu(i)}{2\lambda} q(i) \\
 (2\lambda)^2 &= \frac{1}{k^2} \sum_i q(i)(v(i) - \lambda - \mu(i))^2
 \end{aligned}$$

reemplazando estos valores en la función objetivo de minimización, se obtiene el resultado deseado. ■

4.3.4 Modelo de Entropía Relativa mediante norma L_1

El último conjunto que se analizará es una aproximación o cota inferior de la entropía relativa, utilizando la norma o distancia L_1 .

La distancia L_1 entre dos distribuciones cualquiera se define como

$$\| P_1 - P_2 \|_1 = \sum_{a \in X} |P_1(a) - P_2(a)| \tag{4.30}$$

Utilizando esta distancia, se puede demostrar lo siguiente:

$$D(P_1 \parallel P_2) \geq \frac{1}{2 \ln 2} \|P_1 - P_2\|_1^2 \quad (4.31)$$

Demostración: Considerar la siguiente propiedad, sea A el conjunto que cumple lo siguiente, $P_1(x) > P_2(x)$, luego

$$\begin{aligned} \|P_1 - P_2\|_1 &= \sum_{x \in X} |P_1(x) - P_2(x)| \\ &= \sum_{x \in A} (P_1(x) - P_2(x)) + \sum_{x \in A^c} (P_2(x) - P_1(x)) \\ &= P_1(A) - P_2(A) + P_2(A^c) - P_1(A^c) \\ &= P_1(A) - P_2(A) + 1 - P_2(A) - 1 + P_1(A) \\ &= 2(P_1(A) - P_2(A)) \end{aligned} \quad (4.32)$$

Primero se demostrará para el caso binario. Sean dos distribuciones binarias con parámetros p y q , tal que $p \geq q$, se mostrará que

$$p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q} \geq \frac{4}{2 \ln 2} (p-q)^2$$

La diferencia $g(p, q)$ entre ambos lados de la desigualdad es

$$g(p, q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q} - \frac{4}{2 \ln 2} (p-q)^2$$

luego

$$\begin{aligned} \frac{dg(p, q)}{dq} &= -\frac{p}{q \ln 2} + \frac{1-p}{(1-q) \ln 2} - \frac{4}{2 \ln 2} 2(q-p) \\ &= \frac{q-p}{q(1-q) \ln 2} - \frac{4}{\ln 2} (q-p) \\ &\leq 0 \end{aligned}$$

ya que $q(1-q) \leq 1/4$ y $q \leq p$. Para $q = p$, $g(p, q) = 0$, entonces $g(p, q) \geq 0$ para $q \leq p$, en el caso de dos distribuciones binarias.

Para el caso general, considerar dos distribuciones P_1 y P_2 cualquiera, sea

$$A = \{x: P_1(x) > P_2(x)\}$$

Se define una nueva variable aleatoria binaria $Y = \Phi(X)$ que es el indicador del conjunto A , y sea \hat{P}_1 y \hat{P}_2 las distribuciones de Y . Tanto \hat{P}_1 como \hat{P}_2 corresponden a la versión cuantificada de P_1 y P_2 . Luego se cumple que

$$\begin{aligned} D(P_1 \parallel P_2) &\geq D(\hat{P}_1 \parallel \hat{P}_2) \\ &\geq \frac{4}{2 \ln 2} (P_1(A) - P_2(A))^2 \\ &= \frac{1}{2 \ln 2} \|P_1 - P_2\|_1^2 \end{aligned}$$

la última igualdad se debe a la propiedad (4.32), por lo tanto (4.31) queda demostrado. ■

Por consiguiente, se utilizará esta aproximación para definir el siguiente conjunto

$$\mathcal{P} = \left\{ p \in \mathbf{R}_+^n, p^\top \mathbf{1} = 1 : \|p - f\|_1 \leq \sqrt{2 \ln(2) \beta} \right\}$$

Entonces, el problema (4.19) se plantea como sigue

$$\begin{aligned} \min_p \quad & \sum_i p(i) v(i) \\ \text{s.a.} \quad & \|p - f\|_1 \leq \sqrt{2 \ln(2) \beta} \\ & \sum_i p(i) = 1 \\ & p \geq 0 \end{aligned} \tag{4.33}$$

que es equivalente a resolver

$$\max_{\mu \geq 0} \left\{ \sum_i f(i) (v(i) - \mu(i)) - \frac{1}{2} \sqrt{2 \ln(2) \beta} \left(\max_i \{v(i) - \mu(i)\} - \min_i \{v(i) - \mu(i)\} \right) \right\} \tag{4.34}$$

Los detalles de la demostración se dejan al lector, ya que se usa la misma estrategia que los conjuntos anteriores. Como indicación, se puede usar un cambio de variable $y(i) = p(i) - f(i)$, y luego plantear el lagrangiano que conduce al dual.

En el siguiente capítulo se analizará un ejemplo donde se puede aplicar la metodología de programación dinámica robusta bajo estos conjuntos de incertidumbre.

CAPÍTULO 5

RESULTADOS COMPUTACIONALES

En este capítulo se presentan resultados computacionales que ilustran el comportamiento de las distintas políticas encontradas con los modelos descritos en las secciones anteriores. El objetivo principal es comparar como se comporta el modelo robusto con los modelos tradicionales descritos en el capítulo tres y además como puede afectar la variabilidad de la demanda en los beneficios esperados, especialmente se busca la peor demanda esperada bajo un nivel de riesgo dado. En particular se analizará lo siguiente: Control óptimo para una red mediante programación dinámica cuyos resultados se compararán con el modelo de programación dinámica mediante optimización robusta utilizando el conjunto de entropía relativa mediante norma L_1 .

Las simulaciones se realizaron en Matlab R2010a mediante un procesador 2.3 GHz Intel i5 (4GB RAM, OS X 10.8.1). Debido a la gran dimensión del problema, se resolvió una instancia pequeña tomada del trabajo de Bertsimas y Popescu [13].

5.1 Definición del problema

El experimento que se realizó consiste en tres nodos y dos arcos, los nodos representan los destinos o aeropuertos mientras que los arcos representan la capacidad en asientos de cada avión. La figura 5.1 describe esta red de forma más clara.

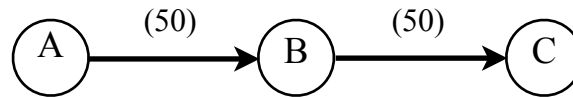


Figura 5.1. Red con tres nodos y dos recursos.

Los nodos están representados por A, B y C, los tramos son (1) AB y (2) BC, la capacidad de la red es $\mathbf{N} = (50, 50)$. Los destinos que ofrece la aerolínea son (1) AB, (2) AC y (3) BC, sólo existe una clase por producto o destino. El vector de precios o tarifas es $\mathbf{R} = (25, 20, 35)$.

La matriz de incidencia junto el vector de tarifas es la siguiente:

$$\begin{pmatrix} R \\ A \end{pmatrix} = \begin{pmatrix} R_1 & R_2 & R_3 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

Por lo tanto, existen dos *legs* y tres productos o destinos, lo que implica una matriz de incidencia de 2×3 , las filas representan cada recurso o avión, mientras que las columnas representan los productos. Esta matriz debe interpretarse como sigue: El producto 1 ocupa un asiento del recurso o avión 1 y su tarifa corresponde a R_1 , mientras que el producto 3 utiliza un asiento del avión 1, un asiento del avión 2 y su tarifa corresponde a R_2 .

En estos experimentos no se aceptan ventas grupales, ni tampoco se considerarán cancelaciones u *overbooking* (vender o reservar más asientos de los que dispone un determinado recurso).

Para modelar o simular el proceso de venta de pasajes aéreos, se asume que el cliente accede a algún punto de venta (tienda física o sitio web en internet), este observa los productos disponibles y compra el pasaje correspondiente a su clase, que en este caso son R_1 , R_2 o R_3 . Este proceso de llegada de clientes se modela mediante un proceso de Poisson Homogéneo con un vector de tasa de llegada constante \mathbf{p} , para el caso en que sólo se considera la demanda esperada (como se analizó en el capítulo 3, sección 3.2.1). Sin embargo, para observar o más bien, estudiar

el comportamiento del modelo robusto descrito en el capítulo 4, es necesario considerar un proceso de Poisson No-Homogéneo, con tasas de llegadas variables a través del tiempo y el estado en particular en que se encuentre el sistema. La justificación del uso de este proceso no-homogéneo es debido a que el modelo robusto se debería comportar mejor en presencia de una demanda que esta por debajo de la demanda esperada, es decir, el peor caso de demanda esperada, por lo que hay que simular con esta demanda en cada etapa. Las tasas de llegada se obtienen a partir de cada iteración en la programación dinámica, resolviendo el problema de optimización descrito en la sección 4.18, para algún conjunto de incertidumbre deseado. Para esta simulación se escogió el último conjunto descrito en 4.3.4 correspondiente al modelo de entropía relativa mediante norma L_1 .

5.2 Desempeño computacional

Resolver en forma exacta el problema *Network Revenue Management* es muy costoso en términos computacionales, aún cuando la tecnología que existe en este tiempo es bastante avanzada, esto es debido a que en programación dinámica hay que calcular la función de beneficio para cada estado posible en que se pueda encontrar el sistema, por lo tanto, si existen dos recursos esto implica $51 \times 51 = 2601$ estados posibles (considerando el estado en que no quedan asientos disponibles, es decir, el estado $(0, 0)$). Si se quisiera resolver en forma exacta una red real mediante programación dinámica, se requeriría un considerable esfuerzo en términos de *hardware* y tiempo para obtener la solución. Debido a esto se consideró una instancia pequeña para ver cuáles son los efectos de usar un modelo robusto en vez del modelo tradicional que no considera riesgo.

El tiempo que tarda el modelo tradicional en entregar una solución es de 40 segundos aproximadamente (para la instancia descrita anteriormente) mientras que el modelo robusto tarda 180 segundos aproximadamente (utilizando el conjunto aproximado de entropía relativa). La figura 5.2 ilustra el costo de agregar un recurso a la red.

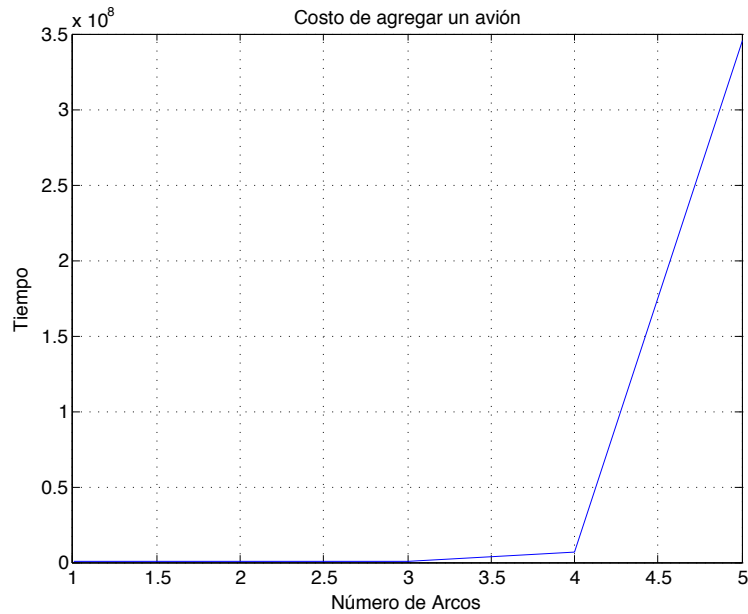


Figura 5.2. Gráfico de Tiempo vs Número de recursos

Como se puede apreciar en la ilustración anterior, agregar un quinto avión a la red con capacidad constante igual a 50 número de asientos en cada recurso, equivale a calcular $51^5 = 345.025.251$ estados posibles en la red. Además hay que considerar otra variable importante en el modelo, el horizonte de tiempo del problema. Para la instancia que se resolvió, se consideró un horizonte igual a doscientos periodos de tiempo, estos pueden ser días en un comienzo, pero cada vez que nos acercamos al tiempo en que partirá el avión, estos periodos de tiempo pueden llegar a ser segundos. Por lo tanto, para estimar el tiempo que tarda el modelo en resolver el problema hay que considerar el número de estados y horizonte de tiempo, es decir, $51^5 \times 200 = 69.005.050.200$ estados, que en tiempo computacional equivale a dos meses, utilizando el hardware descrito en el comienzo de este capítulo.

Debido a la gran dimensión de este problema, es que se usan otros métodos aproximados, como por ejemplo el que se describió en la sección 3.2.2.1 (Modelo de programación lineal determinista), estos modelos estáticos son muy eficientes y se acercan con bastante precisión al óptimo encontrado mediante programación dinámica.

5.3 Cálculo del valor esperado y Bid Prices

El objetivo de resolver los problemas de optimización descritos en 3.9 para el modelo tradicional o sin riesgo y 4.18 para el modelo robusto, es obtener una política de control de asientos, en el capítulo tres se describieron algunas de las políticas más usadas en la industria de las aerolíneas, la que se usará en este trabajo es la política mediante *bid prices*, la diferencia con respecto a los otros tipos de control (*Booking limits* o *Protection levels*) es que ésta se calcula mediante ingresos (*revenues*) y no mediante clases.

Los resultados de los valores esperados se muestran en la siguiente tabla 5.1

T	DP	RDPL1	RDPL2
10	195	80.931	120.1
20	390	161.86	240.19
30	585	242.79	360.29
40	780	323.72	480.38
50	975	404.65	600.48
60	1,170	485.58	720.58
70	1,365	566.51	840.67
80	1,559.5	647.45	960.77
90	1,745.7	728.38	1,080.9
100	1,897.5	809.31	1,200.9
110	1,998.9	890.22	1,309.9
120	2,064.6	970.96	1,325
130	2,110.6	1,050.3	1,331.4
140	2,145.5	1,122.8	1,337.9
150	2,175.7	1,169.8	1,344.4
160	2,202.4	1,191.4	1,350.8
170	2,222.8	1,205.9	1,357.3
180	2,236.2	1,218.8	1,363.7
190	2,243.8	1,231.4	1,370.2
200	2,247.5	1,242.7	1,376.7

Tabla 5.1. Beneficios esperados de los modelos DP y robusto (L1 y L2)

$$\beta = 0.5$$

La tabla anterior muestra los resultados de tres modelos distintos, el primero corresponde a la programación dinámica sin considerar riesgo (DP) explicada en la sección 3.9, mientras que las otras dos columnas representan los resultados de dos modelos robustos, el primero corresponde al modelo de entropía relativa mediante norma L1 (RDP L1) explicado en sección 4.3.4, y el segundo corresponde al modelo elipsoidal (RDP L2) explicado en sección 4.3.3. La figura 5.3 ilustra de mejor manera la tabla anterior.

Cabe destacar que las últimas dos columnas son el resultado de haber considerado un riesgo equivalente a 0.5 en el valor de beta, este riesgo representa un radio o distancia de la probabilidad esperada. Mientras más grande sea este valor, más riesgo se espera que exista, por lo tanto es esperable que la demanda en el peor de los casos cada vez sea más pesimista, y esto se ve reflejado en los valores esperados obtenidos para los modelos robustos. En la siguiente sección 5.4 se realizarán distintas simulaciones aumentando el nivel de riesgo hasta un punto crítico donde los valores esperados de los modelos robusto y tradicional llegan a ser iguales.

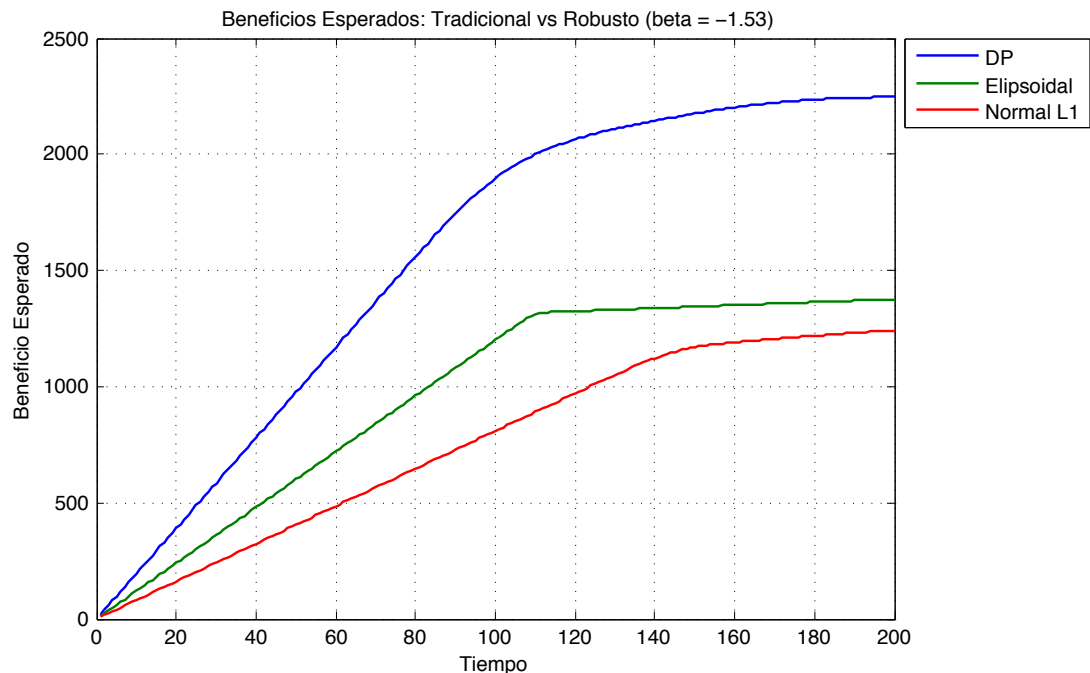


Figura 5.3. Beneficios esperados de los modelos DP y robusto (L1 y L2).

$$\beta = 0.5$$

Una vez calculado los valores esperados de los modelos mencionados anteriormente, es posible encontrar la política de control de asientos, es decir, los *bid prices*. La forma de obtener este tipo de control fue descrito en la sección 3.2.1.1 y 3.2.1.2. A modo de recordatorio, la noción intuitiva de esta política es aceptar una solicitud de reserva para un producto j sólo si su precio es superior al costo de oportunidad de reducir la capacidad de los recursos necesarios para satisfacer dicha solicitud. La figura 5.4 muestra la evolución de los *bid prices* a medida que van quedando menos asientos.

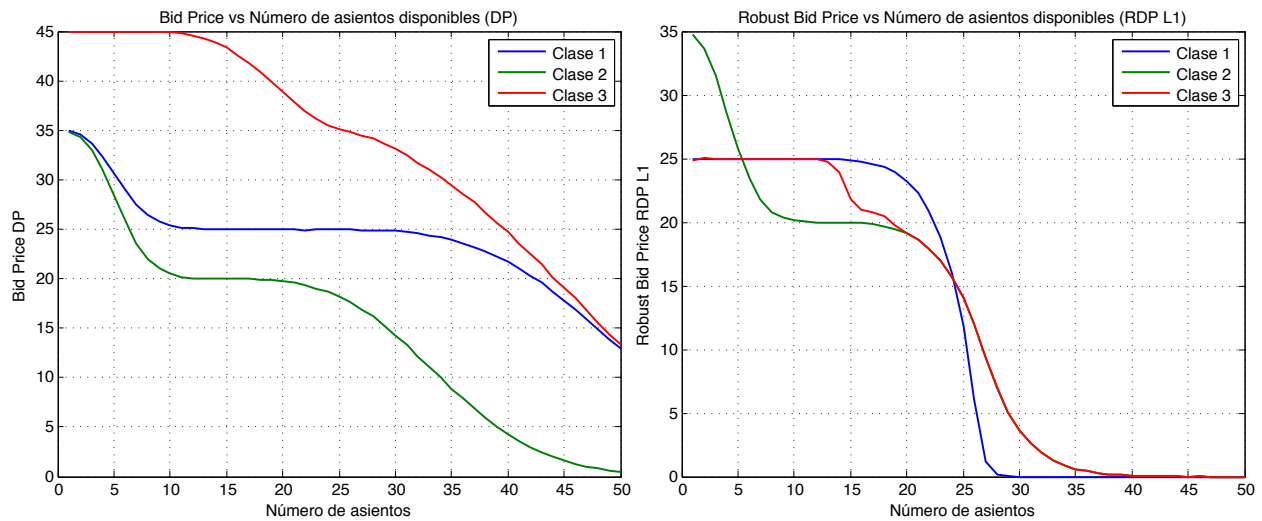


Figura 5.4. *Bid prices* vs Número de asientos disponibles.

El gráfico de la izquierda representa los *bid prices* del modelo DP mientras que el de la derecha el modelo RDP L1. Era de esperar que los *bid prices* obtenidos mediante el modelo robusto fueran más conservadores que el modelo original y esto se ve reflejado en que las curvas cortan el eje y en valores más bajos, comparados con el modelo DP. Además se observa en el modelo robusto que mientras exista capacidad mayor a la mitad disponible, los *bid prices* se mantiene casi en cero.

Como se explicó en la sección 3.1.1 los *bid prices* representan un precio umbral, que depende de variables como capacidad restante en el avión o tiempo, tal que la demanda por algún producto es aceptada si su tarifa o precio es mayor que este precio umbral y rechazada en caso contrario.

5.4 Simulación y comparación del modelo robusto vs tradicional

Los resultados de la simulación para la red con tres nodos y dos recursos descrita en la sección 5.1 se muestran en la siguiente tabla 5.2. La primera columna representa el nivel de riesgo, mientras que las otras dos columnas representan el beneficio esperado junto con la desviación estándar (σ).

β	DP (σ)	RDP L1 (σ)
0.1	3,060.4 (230.5)	1,891.9 (91.4)
0.2	2,729.3 (175.9)	1,592.3 (69.0)
0.3	2,068.0 (138.9)	1,402.5 (37.3)
0.4	1,004.4 (85.0)	1,306.2 (24.1)
0.5	1,007.3 (163.1)	1,244.3 (47.3)
0.6	802.8 (82.0)	1,128.6 (44.1)
0.7	744.8 (77.4)	1,019.0 (29.1)
0.8	796.0 (103.1)	981.6 (49.5)
0.9	857.2 (96.0)	930.7 (78.2)
1.0	822.9 (96.2)	843.8 (108.0)
1.25	566.1 (99.9)	566.1 (99.9)
1.5	316.3 (70.1)	316.3 (70.1)

Tabla 5.2. Beneficios esperados de la simulación para distintos niveles de riesgo (β) y un horizonte de $T = 200$.

La simulación se realizó mediante un proceso de poisson no-homogéneo, es decir, con tasa de llegada variable a través del horizonte de tiempo y los estados del sistema. Esta simulación se realizó considerando la peor demanda.

Debido a que la dimensión del problema es considerablemente grande, como se explicó en la sección 5.2, no se resolvieron instancias más grandes, sin embargo los resultados confirman que a determinado nivel de riesgo el modelo robusto captura de forma más eficiente el riesgo que el modelo DP. La figura 5.5 muestra de forma gráfica los resultados de la tabla anterior.

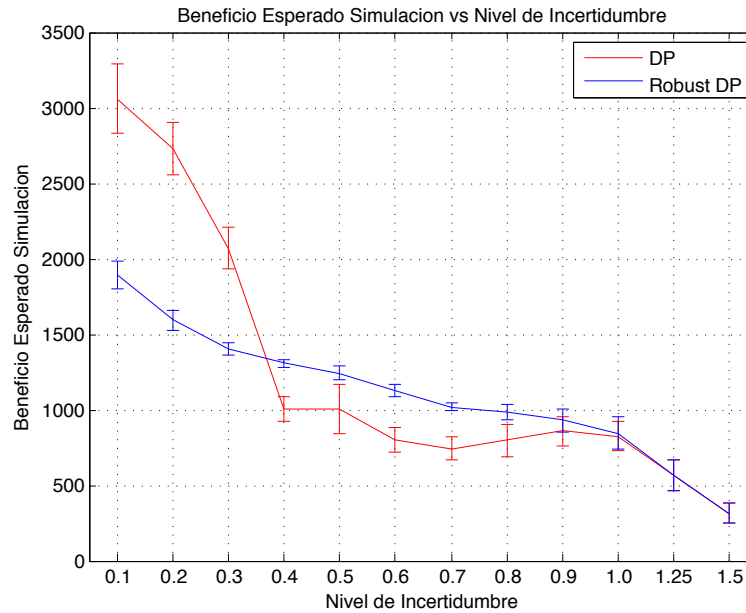


Figura 5.5. Beneficios esperados de la simulación para distintos niveles de riesgo (β) y un horizonte de $T = 200$.

Para la simulación anterior se consideraron $NE = 300$ números de ensayos para cada nivel de incertidumbre, en la figura se observa el promedio junto con la desviación estándar de cada nivel de riesgo.

Por lo tanto, para un nivel de incertidumbre pequeño, el conjunto de medidas condicionales estará en torno a las probabilidades estimadas y en consecuencia el modelo DP se comporta mejor que el robusto, mientras que si el nivel de incertidumbre es más grande, el conjunto de medidas condicionales va a ser más grande por lo que el peor de los casos va a estar más alejado del punto estimado. Es en estos casos donde el modelo robusto es más eficiente, pero llega un punto en que el nivel de riesgo es tan alto, que ambos modelos se comportan igual.

Otro punto a considerar es que el modelo robusto también se puede equivocar, en el sentido que si el tomador de decisiones considera que el riesgo esperado va a ser 0.6 pero en realidad el riesgo fue 0.7, ¿cómo se comportan los modelos? la siguiente figura 5.6 ilustra este hecho.

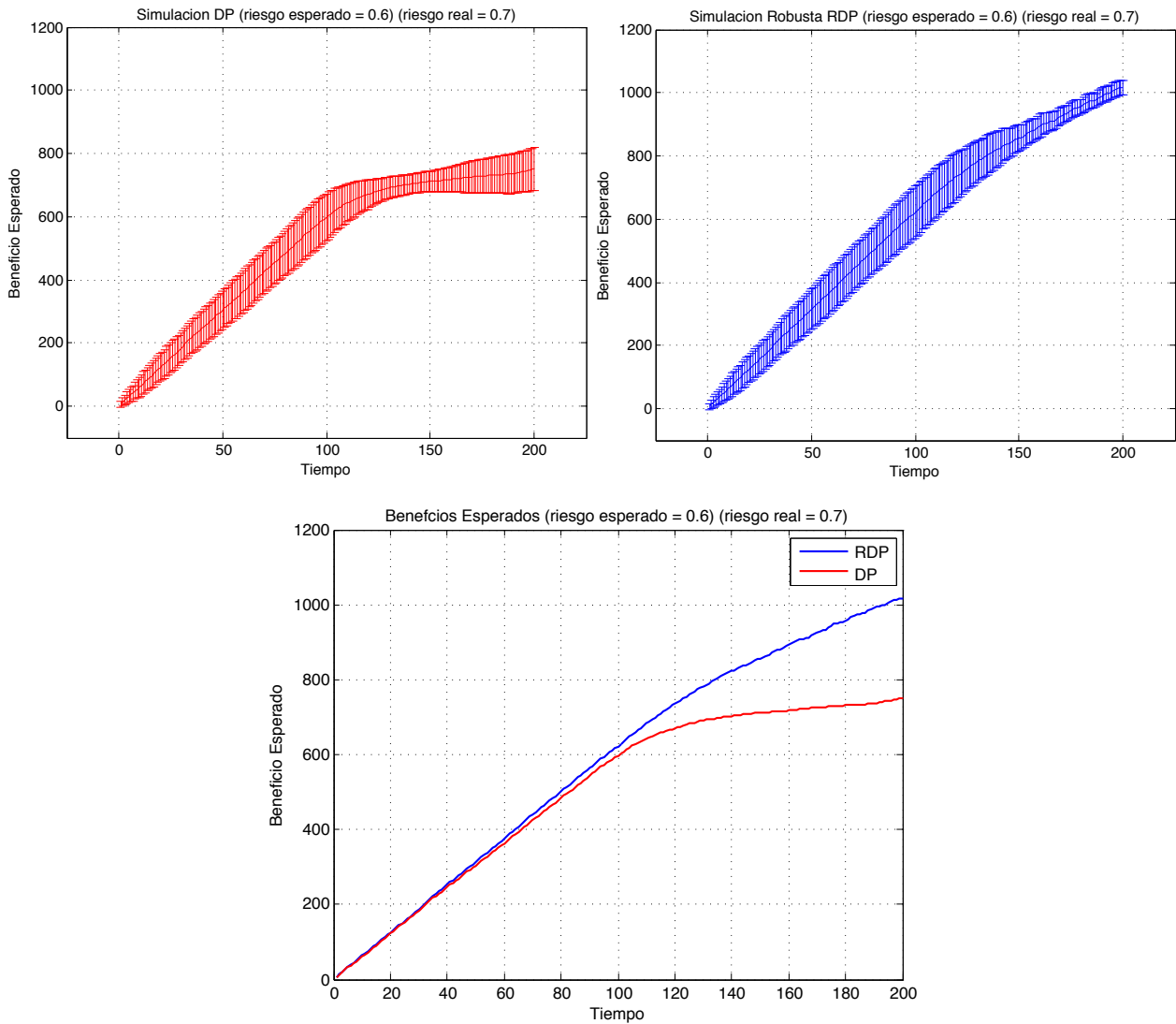


Figura 5.6. Beneficios esperados de la simulación con nivel de riesgo esperado = 0.6 y riesgo real = 0.7

Los dos gráficos que aparecen al principio representan el promedio junto con la desviación estándar para un horizonte de tiempo $T = 200$. Y el tercer gráfico es una comparación de ambos modelos considerando sólo los promedios. Se observa que aún cuando el riesgo real fue mayor que el considerado, el modelo robusto se comporta mejor.

CAPÍTULO 6

CONCLUSIONES

En esta tesis se estudió cómo resolver el problema *Network Revenue Management* para tiempo discreto mediante programación dinámica y su versión robusta, además de mostrar otras técnicas usadas mediante modelos estáticos. La formulación robusta intenta mitigar el impacto de los errores en la estimación de las probabilidades de transición mediante la elección de una política óptima *maximin*, donde la minimización es sobre un conjunto de probabilidades de transición y el objetivo es escoger una política que maximice los beneficios esperados sobre este conjunto. Este conjunto de medidas condicionales intenta modelar el conocimiento limitado o ambigüedad que el tomador de decisiones tiene en torno a las probabilidades de transición de la cadena de markov subyacente. Para ello en el cuarto capítulo se mostraron cuatro conjuntos que capturan de forma eficiente el riesgo.

Los resultados del experimento muestran que cuando hay poco riesgo, es decir, el conjunto de medidas condicionales es pequeño, la probabilidad en el peor caso para cada estado va a estar a poca distancia de la probabilidad estimada, por lo que el modelo original DP se comporta mejor que el modelo robusto, alrededor de 35 puntos porcentuales por sobre el modelo robusto cuando el riesgo es menor que 0.3, sin embargo cuando el riesgo es mayor que 0.3 el modelo robusto captura de forma más eficiente el riesgo y obtiene resultados esperados que son mejores que el modelo original, entre 25 y 40 puntos porcentuales mejor que DP cuando el riesgo está entre 0.4 – 0.8. Ahora si la demanda es muy variable (riesgo muy alto) ambos modelos se

comportan de forma similar, es decir, están llegando tan pocos clientes a comprar pasajes, que el tomador de decisiones acepta cualquier solicitud de compra para así ocupar los asientos disponibles de los recursos en la red. El análisis anterior también se puede ver desde los *bid prices*, mientras mayor sea el riesgo, el modelo robusto va a poner mayor peso a la probabilidad de que no lleguen clientes por lo que los *bid prices* serán cada vez más bajos, de esa forma no se rechazan ventas.

Debido a la configuración del problema, obtener una solución para una red con varios destinos y recursos resulta prácticamente imposible, ya que la metodología de programación dinámica implica resolver un problema de optimización para cada estado y tiempo en el sistema. Además si se agregan más variables como es el caso del modelo robusto, el tiempo para obtener una solución por cada iteración se verá incrementado. Debido a este problema, a la “maldición de la dimensionalidad” como le llama Bellman, es que las aerolíneas consideran modelos estáticos, como el que se analizó en la sección 3.2.2.1 (Modelo de programación lineal determinista) y variaciones de este modelo, también es común el uso de heurísticas como por ejemplo EMSR-a y EMSR-b, entre otras. En el caso del modelo determinista mediante programación lineal, se analizó una aproximación que resulta muy eficiente para calcular los *bid prices*, esta propuesta llamada CEC (*Certainty equivalent control*) es la que está más cerca del óptimo calculado mediante el modelo de programación dinámica (DP). Estas aproximaciones y el uso de modelos estáticos son más rápidos de resolver que el modelo dinámico, ya que no es necesario calcular los *bid prices* para cada estado posible de la red. A estos modelos también se les puede agregar la componente robusta, pero de forma distinta a la que se analizó en esta tesis, en el trabajo de Perakis y Roels [32] se explica cómo construir conjuntos de incertidumbre de la demanda que dependen del criterio del tomador de decisiones.

Si se desea aplicar la metodología robusta mediante programación dinámica a otro problema distinto al de las aerolíneas, el capítulo cuatro explica cómo hacerlo, ya que la solución no está restringida al problema de las aerolíneas, basta definir cuál es la función de beneficio, las variables de decisión, los parámetros con incertidumbre, las etapas y condiciones de borde para aplicarlo a un problema distinto.

Existen varios problemas que se pueden resolver mediante programación dinámica robusta⁵, uno de ellos es el problema de venta de piezas en los hoteles para una fecha dada, un número determinado de reservas en días y diferentes clases de tarifas. En este caso, el problema es gestionar la capacidad de piezas del hotel en días consecutivos cuando los clientes se quedan por varias noches. Este problema es conocido en la literatura como *length of stay control*.

Dado que esta tesis pretende ser una primera aproximación al problema de *Network Revenue Management* mediante programación dinámica robusta, hay diversos aspectos que se deben abordar para lograr mejores resultados: Por ejemplo, considerar data histórica del problema real y estimar las probabilidades de transición mediante máxima verosimilitud. Considerar restricciones adicionales, en el caso de aerolíneas estas pueden ser cancelaciones, *overbooking*, llegadas grupales, entre otras, de esta forma el modelo es más completo y se acerca a la realidad. Dado que obtener soluciones mediante la metodología de programación dinámica es costoso en términos computacionales, es recomendable contar con hardware que sea rápido y solvers (como los que incluye Matlab) que sean eficientes al resolver problemas de optimización lineales y no lineales. También se pueden agregar condiciones de mercado a las variables aleatorias, de esta forma se estaría abarcando un problema más realista.

⁵ Industria de trenes, venta de carga en aerolíneas, turismo en cruceros, entre otros.

REFERENCIAS

- [1] Adelman, D. *Dynamic bid-prices in revenue management*. Oper. Res. 55(4) 647–661, 2007.
- [2] Akan, M., B. Ata. *On bid price controls for network revenue management*. Northwestern University, Chicago. 2006.
- [3] Belobaba, P. P. *Air travel demand and airline seat inventory management*. Ph.D. thesis, MIT Cambridge, MA. 1987.
- [4] Belobaba, P. P. *Optimal vs. heuristic methods for nested seat allocation*. Proc. AGIFORS Reservations and Yield Management Study Group, Brussels, Belgium, 1992.
- [5] Belobaba, P. P. *Revenue and competitive impacts of O-D controls: Summary of PODS result*. INFORMS Revenue Management Section Meeting, NY, 2001.
- [6] Belobaba, P. P., S. Lee. *PODS update: Large network O-D control results*. AGIFORS Reservation and Yield Management Study Group Symposium Proceedings, NY, 2000.
- [7] Ben-Tal, A., A. Nemirovski. *Robust convex optimization*. Math. Oper. Res., 23(4): 769–805, 1998.
- [8] Ben-Tal, A., A. Nemirovski. *Robust solutions to uncertain programs*. Operations Research Letters 25, 1–13, 1999.
- [9] Ben-Tal, A., A. Nemirovski. *Robust solutions of linear programming problems contaminated with uncertain data*, Math. Program, 88, 411-424, 2000.

- [10] Ben-Tal, A., B. Golany, A. Nemirovski, J.-P. Vial. *Retailer- supplier flexible commitments contracts: A robust optimization approach*. Manufacturing Service Oper. Management 7(3) 248–271, 2005.
- [11] Bertsekas, D. *Dynamic Programming and Optimal Control*. Tercera edición Volúmenes 1 y 2, 2005.
- [12] Bertsimas, D., A. Thiele. 2006. *A robust optimization approach to inventory theory*. Oper. Res. 54(1) 150–168, 2006.
- [13] Bertsimas, D., I. Popescu. *Revenue Management in a Dynamic Network Environment*. Transportation Science Informs Vol. 37, No. 3, pp. 257-277, 2003.
- [14] Bertsimas, D., M. Sim. *Robust discrete optimization and network flows*. Mathematical Programming 98(1-3), 49–71, 2003.
- [15] Bertsimas, D., M. Sim. *The price of robustness*. Operations Research 52(1), 35–53, 2004.
- [16] Bertsimas, D., S. de Boer. *Simulation-based booking limits for airline revenue management*. Oper. Res. 53(1) 90–106, 2005.
- [17] Besbes, O., A. Zeevi. *Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms*. Oper. Res. 2009.
- [18] Cooper, W. L. *Asymptotic behavior of an allocation policy for revenue management*. Operation Research, 50:720–727, 2002.
- [19] Cross, R. G. *Revenue Management: Hardcore Tactics for Market Domination*. Broadway Books, New York, NY, 1997.

- [20] Diamond, M., R. Stone. *Dynamic yield management on a single leg*. Unpublished manuscript, Northwest Airlines, 1991.
- [21] El-Ghaoui, L., F. Oustry, y H. Lebret. *Robust solutions to uncertain semidefinite programs*. SIAM J. Optimization 9(1), 1998.
- [22] Erken, S., C. Maglaras. *Revenue management heuristics under limited market information: A maximum entropy approach*. 6th Annual INFORMS Revenue Management and Pricing Section Conference, INFORMS, Hanover, MD, 2006.
- [23] Gallego, G., I. Moon. *The distribution free newsboy problem: Review and extensions*. J. Oper. Res. Soc. 44(8) 825–834, 1993.
- [24] Gallego, G., J. K. Ryan, D. Simchi-Levi. *Minimax analysis for finite-horizon inventory models*. IIE Trans. 33 861–874, 2001.
- [25] Godfrey, G., W. B. Powell. *An adaptive, distribution-free algorithm for the newsvendor problem with censored demands, with application to inventory and distribution problems*. Management Sci. 47(8) 1101–1112, 2001.
- [26] Goldfarb, D., G. Iyengar. *Robust portfolio selection problems*. Mathematics of Operations Research 28(1), 1–38, 2003.
- [27] Huh, W. T., P. Rusmevichientong. *A nonparametric asymptotic analysis of inventory planning with censored demand*. Math. Oper. Res. 34(1) 103–123, 2009.
- [28] Iyengar, G. *Robust Dynamic Programming*. Mathematics of Operation Research Vol. 30, No 2, pp. 257-280, 2005.

- [29] Levi, R., R. Roundy, D. B. Shmoys. *Provably near-optimal sample-based policies for stochastic inventory control models*. Math. Oper. Res. 32(4) 821–838, 2007.
- [30] Littlewood, K. *Forecasting and control of passenger bookings*. Twelfth Annual AGIFORS Symposium, Nathanya, Israel, 1972.
- [31] Nilim, A., L. El Ghaoui. *Robust solutions to Markov decision problems with uncertain transition matrices*. UC Berkeley Tech Report UCB - ERL-M02/31. 2005.
- [32] Perakis G. and Roels G. *Robust Controls for Network Revenue Management*. Manufacturing & Service Operations Management 12(1), pp. 56-76, 2010.
- [33] Rusmevichientong, P., P. Van Roy, P. W. Glynn. *A non- parametric approach to multiproduct pricing*. Oper. Res. 54(1) 89–98, 2006.
- [34] Scarf, H. E. *A min-max solution to an inventory problem*. Stanford University Press, Stanford, CA, 201–209, 1958.
- [35] Talluri, K. *On bounds for network revenue management*. Universitat Pompeu, Fabra, Barcelona, Spain, 2008.
- [36] Talluri, K. T., G. J. Van Ryzin. *The Theory and Practice of Revenue Management*. Springer; XXXII, 714 pg. edition 2004.
- [37] Talluri, K. T., G. J. Van Ryzin. *An analysis of bid-price controls for network revenue management*. Management Science, 44:1577-1593, 1999.
- [38] Topaloglu, H. *Using Lagrangian relaxation to compute capacity-dependent bid-prices in network revenue management*. Oper. Res. 57(3) 637–649, 2009.

- [39] van Ryzin, G. J., G. Vulcano. *Simulation-based optimization of virtual nesting controls for network revenue management*. Oper. Res. 56(4) 865–880, 2008.
- [40] van Ryzin, G. J., J. I. McGill. *An adaptive algorithm for determining seat protection levels*. Management Sci. 46(6) 760–775, 2000.
- [41] Weatherford, L. R., P. P. Belobaba. *Revenue impacts of fare input and demand forecast accuracy in airline yield management*. J. Oper. Res. Soc. 53(8) 811–821, 2002.
- [42] Williamson, E. *Comparison of optimization techniques for origin-destination seat inventory control*. Thesis, Flight Transportation Laboratory, MIT, Cambridge, MA 1988.
- [43] Williamson, E. *Airline network seat control*. Ph.D. thesis, MIT, Cambridge, MA. 1992.
- [44] Yue, J., B. Chen, M. C. Wang. *Expected value of distribution information for the newsvendor problem*. Oper. Res. 6(54) 1128–1136, 2006.
- [45] Zhou, K., J. Doyle, y K. Glover. *Robust and optimal control*. Prentice Hall, Upper Saddle River, NJ. 1996.