An Optimal Control Problem of Dynamic Pricing

Anton J. Kleywegt *
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0205, USA

July 19, 2001

Abstract

An optimal control problem of dynamic pricing is studied. In the model, prices are chosen to sell multiple products to multiple customer classes over time. The products share a number of scarce resources. All parameters, such as the arrival rates of customers, their purchasing probabilities, their cancellation rates, and the cancellation refunds, are allowed to be time dependent. A solution method for the problem is developed, and is tested with some numerical examples.

*Supported, in part, by the National Science Foundation under grant DMI-9875400.

1 Introduction

The problem of optimal dynamic pricing is of great intellectual, economic, and practical interest. The practical interest follows from the dramatic increase in profitability that can be obtained with improved dynamic pricing. Cases of such improvements were described in Smith, Leimkuhler and Darrow (1992), and Geraghty and Johnson (1997). The economic interest follows from the fact that supply and demand cannot be balanced in a dynamic setting unless dynamic pricing is done efficiently. The intellectual interest has been strong, because in spite of two decades of research by several researchers, optimal dynamic pricing remains a challenging problem. A survey of this research is given in McGill and Van Ryzin (1999).

The model of dynamic pricing introduced in this paper is an improvement over the models of dynamic pricing described in the literature in several ways. Many researchers have worked on problems in which a single product is sold over a time interval, and a variety of results characterizing optimal policies have been obtained for such problems, such as the optimality of simple threshold rules (nesting) for accepting customer requests. The dynamic pricing problem with multiple products and multiple shared resources has turned out to be much more challenging. Research on such problems has been limited to heuristics, some of which were shown to be asymptotically optimal as the number of resources and the demand rates become large. Some of the more widely used methods in industry, such as the EMSR method for booking limit control and the bid price method for revenue threshold booking control are not only heuristics, but they are also based on models of customer behavior that are flawed. Some of the improvements, as well as shortcomings, of the model formulated in this paper are discussed in Section 2.3 after the model has been introduced.

The two major contributions of this paper are the following:

1. The model of dynamic pricing formulated in this paper provides a better combination of realism and tractability than the models described in the literature.

2. A method is developed to solve the optimal dynamic pricing problem introduced here.

In this paper the terms dynamic pricing, revenue management, and yield management are regarded as synonyms, although distinctions are made between these terms in some applications.

The paper is organized as follows: The dynamic pricing problem is introduced and discussed in Section 2. An optimal control formulation of dynamic pricing is given in Section 3. The remaining sections describe a solution method for the optimal control problem. A Lagrangean problem associated with the optimal control problem is discussed in Section 4. Section 5 discusses the numerical solution of the Lagrangean problem. The associated Lagrangean dual problem and its solution is discussed in Section 6. Finally, some computational results are given in Section 7.

2 The Dynamic Pricing Problem

In this section we describe a fairly general dynamic pricing problem, and we discuss how the problem relates to applications and to other formulations of dynamic pricing or revenue management problems.

2.1 Problem Formulation

The problem is formulated in continuous time over an interval $[0, T]$. There is a set of products that can be sold during $[0, T]$. Each product can be sold at several prices. For simplicity of notation, we use a single index to denote a product-price pair. Let $\mathcal{I}$ denote the set of all product-price pairs. For each $i \in \mathcal{I}$ and $t \in [0,T]$, let $s_i(t)$ denote the net revenue associated with selling a unit of product-price pair i at time t, and let $c_i(t)$ denote the net refund if the sale of a unit of product-price pair i is cancelled at time t. Thus, the net refund given to a no-show for a unit of product-price pair i is denoted by $c_i(T)$. Discounting of revenues and costs are easily incorporated, because $s_i(t)$ and $c_i(t)$ may depend on t. There is a set $\mathcal{R}$ of resources; the resources are used to supply the products. A total quantity b_r of each resource $r \in \mathcal{R}$ is available. Each $i \in \mathcal{I}$ uses quantity R_{ri} of resource $r \in \mathcal{R}$.

Customers are classified according to their preferences for different product-price pairs, based on information available to the seller. Customer preferences are modelled with probabilities, or the fractions of customers of each class who accept a product-price pair if the product-price pair is offered to the customer. Let $\mathcal{K}$ denote the set of customer classes. For each $k \in \mathcal{K}$, let $\alpha_k(t)$ denote the arrival rate per unit time of requests from class k customers at time $t \in [0,T]$. For each $i \in \mathcal{I}$, $k \in \mathcal{K}$, and $t \in [0,T]$, let $p_{ik}(t)$ denote the probability that a class k customer who makes a request at time t, purchases the product if product-price pair i is offered, and let $\mu_{ik}(t)$ denote the rate at which class k customers, who have purchased a unit of product-price pair i, cancel their purchases at time t. For each $i \in \mathcal{I}$ and $k \in \mathcal{K}$, let q_{ik} denote the probability that a class k customer, who has purchased a unit of product-price pair i, does not cancel the purchase at the last moment (is not a no-show).

The controls are as follows. For each $i \in \mathcal{I}$, $k \in \mathcal{K}$, and $t \in [0,T]$, let $u_{ik}(t) \in [0,1]$ denote the fraction of class k requests to whom product-price pair i is offered at time t. Thus the $u_{ik}(t)$'s have to satisfy

$$\begin{aligned} \sum_{i \in \mathcal{I}} u_{ik}(t) &\leq 1 \quad \text{for all } k \in \mathcal{K} \text{ and } t \in [0,T] \\ u_{ik}(t) &\geq 0 \quad \text{for all } i \in \mathcal{I}, k \in \mathcal{K}, \text{ and } t \in [0,T] \end{aligned}$$

The quantity $1 - \sum_{i \in \mathcal{I}} u_{ik}(t)$ denotes the fraction of class k requests that are not made any offers at time t.

In many dynamic pricing applications, a set of product-price pairs are offered to a customer when a request is received, and then the customer either chooses one of the product-price pairs in the offered set or none of them. Such a situation can be modelled in the same way as was done above. For a subset $\S \subset \mathcal{I}$ of product-price pairs, $p_{ik\S}(t)$ would denote the probability that a class k customer who makes a request at time t, purchases product $i \in \S$ if subset $\S$ of product-price pairs is offered, and $u_{\S k}(t) \in [0,1]$ would denote the fraction of class k requests to whom subset $\S$ of product-price pairs is offered at time t. To keep the notation and terminology simple, we do not present this enhancement further.

It is assumed that all functions s_i, c_i, α_k, p_{ik}, and μ_{ik} are Lebesgue integrable, that is, s_i, c_i, α_k, p_{ik}, and μ_{ik} are Lebesgue measurable and $\int_0^T |s_i(t)|\,dt < \infty$, $\int_0^T |c_i(t)|\,dt < \infty$, $\int_0^T \alpha_k(t)\,dt < \infty$, and $\int_0^T \mu_{ik}(t)\,dt < \infty$. In addition, attention is restricted to Lebesgue measurable controls u_{ik}.

For each $i \in \mathcal{I}$, $k \in \mathcal{K}$, and $t \in [0,T]$, let $x_{ik}(t)$ denote the amount of product-price pair i that has been sold to class k customers at time t. The vector $x(t)$ is called the state of the process at time t. The initial state $x(0) = x_0 \geq 0$ at time 0 is given.

Note that if the cancellation rate $\mu_{ik}(t)$ and the no-show probability $(1 - q_{ik})$ for product-price pair i do not depend on the class of customer k who made the purchase, then it is sufficient to let the state vector be $x(t) = (x_i(t), i \in \mathcal{I})$, where $x_i(t)$ denotes the total amount of product-price pair i that has been sold at time t.

Because of cancellations and no-shows, the constraint associated with the limited amount b_r of each resource $r \in \mathcal{R}$ is only enforced at terminal time T; at times $t < T$ overbooking is allowed.

Let $Q \in \mathbb{R}^{|\mathcal{I}| \times |\mathcal{I}||\mathcal{K}|}$ denote the matrix with entry q_{ik} in row i and column ik. Then entry i of the vector $Qx(T)$ is equal to $\sum_{k \in \mathcal{K}} q_{ik} x_{ik}(T)$, and is thus equal to the quantity of product i that has to be supplied to customers after no-shows have been deducted. Then entry r of the vector $RQx(T)$ is equal to $\sum_{i \in \mathcal{I}} R_{ri} \sum_{k \in \mathcal{K}} q_{ik} x_{ik}(T)$, and is thus equal to the quantity of resource r that is needed to supply the products that are eventually sold. Hence, the constraints associated with the limited amounts of resources are given by $RQx(T) \leq b$.

At any time $t \in [0, T]$, revenue is earned from sales of product-price pair i to customers of class k at rate $s_i(t)\alpha_k(t)p_{ik}(t)u_{ik}(t)$ per unit time, and refunds are paid due to cancellation of sales of product-price pair i by customers of class k at rate $c_i(t)\mu_{ik}(t)x_{ik}(t)$ per unit time. At the same time, the amount $x_{ik}(t)$ of product-price pair i that has been sold to class k customers increases due to additional sales at rate $\alpha_k(t)p_{ik}(t)u_{ik}(t)$ per unit time, and decreases due to cancellations at rate $\mu_{ik}(t)x_{ik}(t)$ per unit time. The objective is to maximize the profit (net revenues minus refunds) accumulated over the time interval $[0, T]$.

2.2 Some Applications

The most widely studied revenue management problem is the control of the booking of airline passengers. To apply the problem formulated in Section 2.1 to airline passenger revenue management, a few issues have to be clarified. Typically the products (origin-destination travel on particular flight legs), prices, and resources (seats of each type on each flight leg) have already been determined by the time the revenue management problem has to be solved, and are therefore easy to identify. An issue is how the customer classes are defined. The decision maker should be able to determine the class of a customer at the time that a product-price pair has to be chosen to be offered to the customer. Thus the customer classification must be based on information that is available to the decision maker at the time that a request is received from the customer. Airline customers typically specify the desired origin, destination, and dates of travel when a request is submitted. Additional information provided by the customer includes the number of tickets requested, whether a one-way or return or multistop route is requested, whether travel at a particular time is requested or whether the travel time is flexible, and whether the customer is interested in only a refundable ticket or whether the customer would consider any ticket, refundable or nonrefundable. The customer classification could be based on all such available information. Another issue is how products with different deadlines for bookings (different departure times) fit into a model with a single time horizon T. Such a situation is easily modelled by letting $p_{ik}(t) = 0$ and $\mu_{ik}(t) = 0$ at times t between the deadline for i and the time horizon T. Also, the no-shows $(1 - q_{ik})x_{ik}(T)$ at time T represent no-shows at the deadline for i. However, in such applications the modeler has to choose the time horizon T, and there may be distortions associated with the boundary at T. For example, the optimal control in a neighborhood of time zero may depend on T, and further investigation may be necessary to determine the sensitivity of this dependence.

An emerging application area of revenue management is in freight transportation. Most freight is transported under long term contracts in which the prices of different products have been specified. However, the problem formulated in Section 2.1 still apply to operational booking control. For example, in ocean container transportation, the schedule of voyages and ships that will perform those voyages are determined long in advance. Also, contracts between carriers and customers are signed for often a year at a time. Such a contract typically states that the carrier will be able to carry the freight of the customer, unless circumstances not under the control of the carrier (such as bad weather or war) makes it impossible. However, the carrier does not have to transport the customer's freight on the next voyage. If the voyage is already almost fully booked, and the carrier wants to reserve some capacity for anticipated booking requests from highly valued customers (for example customers with high contract prices), then the carrier may offer a booking to a less valued customer who requests a booking on a later voyage only. Customers, in turn, typically have contracts with several carriers, and a customer who requests a booking may refuse the offered booking if the freight would be delayed too much. Thus, in ocean container transportation applications, the products are slots on future voyages, the prices are given by the contract prices, and the resources are the capacities on the voyages. Containers often have to be carried on several voyages between successive ports between their origins and destinations, and thus products share resources as in airline applications. Bookings for different products have to be received before different deadlines; this aspect is handled as described above for the air passenger application.

Another prominent application area is dynamic pricing for equipment rental, including car rental and hotel stays. The resources are the equipment during a time period, such as a rental car or a hotel room for a day. Customers desire products that require use of the equipment over several time periods, and thus products require and share several resources. Bookings for different products such as car rentals and hotel stays on different days have to be received before different deadlines. Again this aspect can be handled as described above for the air passenger application.

2.3 Comparison with Some Existing Models

The problem formulated in Section 2.1 is an improvement over the revenue management models that have been proposed in the literature in several ways. Some models, such as EMSR models for leg based booking limit control in air passenger applications, assume that there is an exogenous demand for each price class (bucket) that makes use of each resource (flight leg). Sometimes it is even assumed that the exogenous demands for the cheaper price classes occur strictly before the demands for the more expensive flight classes. Other models, such as models for bid price control in air passenger applications, assume that there is an exogenous demand for each product-price pair (origin-destination-fare combination). However, the observed demand for different price classes and for different resources is not just exogenously determined, but is also determined by the revenue management controls that were used. It is therefore more appropriate to model demand for different products, and then the quantities sold at different prices, the quantities of resources used, and the behavior of bookings over time are derived from the demand for the products and the revenue management controls, as in the model of Section 2.1.

In models for bid price control in air passenger applications, the primal decision variables represent the quantities of each product-price pair (origin-destination-fare combination) sold. Because there are so many of these product-price pairs, and the realized demand for many of them is so small and unreliable, the values of the primal decision variables produced by the model are not very useful for booking control. That is one reason why the booking controls based on these bid price models use the values of the dual decision variables instead. In the model of Section 2.1, the primal decision variables represent the product-price pair, or set of product-price pairs, offered to each customer class if a request is received from a customer of the particular class at the particular time. These primal decision variables are more useful than those of the bid price models.

The model formulated in Section 2.1 is a dynamic model, taking time dependencies into account, whereas many revenue management models proposed in the literature are static models.

The model formulated in Section 2.1 and the associated optimal control problem formulated in Section 3 still have several shortcomings. The optimal control problem is a deterministic fluid model, and therefore does not take the effects of random variations as well as the integrality of the demand in many applications into account. A shortcoming that we anticipate to be even more severe, is that the model, similar to most revenue management models, does not take competition explicitly into account, but only implicitly through the acceptance probabilities. Also, in many applications it may be more appropriate to model prices as continuous decision variables, instead of a given set of prices to choose from, as in the model of this paper. Another shortcoming that will cause problems in practical applications is that, similar to many revenue management models, the model of this paper has a large number of input parameters, and the collection of appropriate data and the estimation of these parameters from data can be a more challenging and time consuming task than solving the optimization problem.

3 The Optimal Control Problem

In this section we formulate the dynamic pricing problem described in Section 2 as an optimal control problem. Let $m = n = |\mathcal{I}| \times |\mathcal{K}|$, let $\mathcal{U}$ denote the set of Lebesgue measurable functions $u : [0,T] \mapsto [0,1]^m$, and let $\mathcal{X}$ denote the set of absolutely continuous functions $x : [0,T] \mapsto \mathbb{R}^n$. The optimal control problem has the following form:

$$\sup_{x\in\mathcal{X},u\in\mathcal{U}} \left\{F(x,u) := \int_0^T g(x(t),u(t),t)\,dt + h(x(T))\right\} \tag{1}$$

$$\text{subject to} \quad x(t) \;=\; x_0 + \int_0^t f(x(\tau),u(\tau),\tau)\,d\tau \quad \text{for all } t \in [0,T] \tag{2}$$

$$RQx(T) \;\leq\; b \tag{3}$$

$$D(t)u(t) \;\leq\; d(t) \quad \text{for all } t \in [0,T] \tag{4}$$

Here the revenue rate function $g : \mathbb{R}^n \times [0,1]^m \times [0,T] \mapsto \mathbb{R}$ is given by

$$g(x,u,t) \;:=\; \sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} \left[s_i(t)\alpha_k(t)p_{ik}(t)u_{ik} - c_i(t)\mu_{ik}(t)x_{ik}\right] \tag{5}$$

The terminal value function $h : \mathbb{R}^n \mapsto \mathbb{R}$ is given by

$$h(x) \;:=\; -\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} c_i(T)(1-q_{ik})x_{ik} \tag{6}$$

The transition rate function $f : \mathbb{R}^n \times [0,1]^m \times [0,T] \mapsto \mathbb{R}^n$ is given by

$$f_{ik}(x,u,t) \;:=\; \alpha_k(t)p_{ik}(t)u_{ik} - \mu_{ik}(t)x_{ik} \tag{7}$$

The capacity constraints (3) are given by

$$\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} R_{ri}q_{ik}x_{ik}(T) \;\leq\; b_r \quad \text{for each } r \in \mathcal{R} \tag{8}$$

The allocation constraints (4) are given by

$$\sum_{i\in\mathcal{I}} u_{ik}(t) \;\leq\; 1 \quad \text{for each } k \in \mathcal{K} \text{ and each } t \in [0,T] \tag{9}$$

It is shown in Section 4 that, for each $u \in \mathcal{U}$, the equation $x(t) = x_0 + \int_0^t f(x(\tau),u(\tau),\tau)\,d\tau$ has a solution $x \in \mathcal{X}$. It follows from the assumption that functions s_i, c_i, α_k, and μ_{ik} are integrable that the optimal value of optimal control problem (1)–(4) is finite, unless the problem is infeasible, in which case the optimal value is $-\infty$ by convention.

It also will be useful to have notation $V^*(x_t, b, t)$ for the optimal objective value over the interval $[t, T]$, given that the process is in state x_t at t. That is,

$$
\begin{aligned}
V^*(x_t, b, t) \quad &:= \quad \sup_{x \in \mathcal{X}, u \in \mathcal{U}} \quad \int_t^T g(x(\tau), u(\tau), \tau)\, d\tau + h(x(T)) && (10)\\
&\text{subject to} \quad x(\tau) \ = \ x_t + \int_t^\tau f(x(s), u(s), s)\, ds \quad \text{for all } \tau \in [t, T] && (11)\\
&\qquad RQx(T) \ \leq \ b && (12)\\
&\qquad D(\tau)u(\tau) \ \leq \ d(\tau) \quad \text{for all } \tau \in [t, T] && (13)
\end{aligned}
$$

The function V^* is called the primal function.

4 The Lagrangean Problem

In the remaining sections we describe a solution method for the dynamic pricing optimal control problem (1)–(4). Constraints (3) make the problem hard. Consider the following Lagrangean problem associated with the optimal control problem (1)–(4). Let the Lagrangean objective function $L : \mathcal{X} \times \mathcal{U} \times \mathbb{R}^{|\mathcal{R}|} \mapsto \mathbb{R}$ be given by

$$
L(x, u, \lambda) \quad := \quad \int_0^T g(x(t), u(t), t)\, dt + h(x(T)) + \lambda^T [b - RQx(T)] \tag{14}
$$

Then the Lagrangean dual function $L^* : \mathbb{R}^{|\mathcal{R}|} \mapsto \mathbb{R}$ is given by

$$
\begin{aligned}
L^*(\lambda) \quad &:= \quad \sup_{x \in \mathcal{X}, u \in \mathcal{U}} \quad L(x, u, \lambda) && (15)\\
&\text{subject to} \quad x(t) \ = \ x_0 + \int_0^t f(x(\tau), u(\tau), \tau)\, d\tau \quad \text{for all } t \in [0, T] && (16)\\
&\qquad D(t)u(t) \ \leq \ d(t) \quad \text{for all } t \in [0, T] && (17)
\end{aligned}
$$

It follows from the assumption that functions s_i, c_i, α_k, and μ_{ik} are integrable that the Lagrangean problem (15)–(17) has a finite optimal value $L^*(\lambda)$ for all $\lambda \in \mathbb{R}^{|\mathcal{R}|}$. Also note that the Lagrangean problem (15)–(17) is a relaxation of the optimal control problem (1)–(4) for all $\lambda \geq 0$.

For any $\lambda \in \mathbb{R}^{|\mathcal{R}|}$, the Lagrangean terminal value function $h^\lambda : \mathbb{R}^n \mapsto \mathbb{R}$ is given by

$$
h^\lambda(x) \ := \ -\sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} c_i(T)(1 - q_{ik}) x_{ik} + \sum_{r \in \mathcal{R}} \lambda_r \left[b_r - \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} R_{ri} q_{ik} x_{ik} \right] \tag{18}
$$

For any λ, let (x^λ, u^λ) denote an optimal solution of the Lagrangean problem (15)–(17). (It is shown later that such an optimal solution (x^λ, u^λ) exists for each $\lambda \in \mathbb{R}^{|\mathcal{R}|}$; however, there may not be a unique optimal solution.)

Let $V^\lambda(x_t, b, t)$ denote the optimal objective value of the Lagrangean problem with dual multipliers λ over the interval $[t, T]$, given that the process is in state x_t at t. That is,

$$
\begin{aligned}
V^\lambda(x_t, b, t) \quad &:= \quad \sup_{x \in \mathcal{X}, u \in \mathcal{U}} \quad \int_t^T g(x(\tau), u(\tau), \tau)\, d\tau + h(x(T)) + \lambda^T [b - RQx(T)] && (19)\\
&\text{subject to} \quad x(\tau) \ = \ x_t + \int_t^\tau f(x(s), u(s), s)\, ds \quad \text{for all } \tau \in [t, T] && (20)\\
&\qquad D(\tau)u(\tau) \ \leq \ d(\tau) \quad \text{for all } \tau \in [t, T] && (21)
\end{aligned}
$$

The function V^λ is called the optimal value function. Note that $L^*(\lambda) = V^\lambda(x_0, b, 0)$.

The Hamiltonian $H : \mathbb{R}^n \times [0,1]^m \times [0,T] \times \mathbb{R}^n \mapsto \mathbb{R}$ is given by

$$H(x,u,t,\psi) \; := \; g(x,u,t) + f(x,u,t)^T \psi \tag{22}$$

Note that g is linear in (x,u): $g(x,u,t) = g_x(t)^T x(t) + g_u(t)^T u(t)$, where $g_x : [0,T] \mapsto \mathbb{R}^n$, with

$$g_{x_{ik}}(t) \; := \; -c_i(t)\mu_{ik}(t) \tag{23}$$

and $g_u : [0,T] \mapsto \mathbb{R}^m$, with

$$g_{u_{ik}}(t) \; := \; s_i(t)\alpha_k(t)p_{ik}(t) \tag{24}$$

Also, f is linear in (x,u): $f(x,u,t) = f_x(t)x(t) + f_u(t)u(t)$, where $f_x(t) \in \mathbb{R}^{n\times n}$ is a diagonal matrix with

$$f_{x_{ik,ik}}(t) \; := \; -\mu_{ik}(t) \tag{25}$$

and $f_u(t) \in \mathbb{R}^{n\times m}$ is also a diagonal matrix with

$$f_{u_{ik,ik}}(t) \; := \; \alpha_k(t)p_{ik}(t) \tag{26}$$

(Note that f_x and f_u have a row and column corresponding to each ik pair. This is for simplicity of notation—in applications, customers of a class k are typically interested in only a small subset of product-price pairs i, and hence most ik pairs can be eliminated.) In addition, h^λ is affine in x: $h^\lambda(x) = h_x^{\lambda T} x + \lambda^T b$, where $h_x^\lambda \in \mathbb{R}^n$ with

$$h^\lambda_{x_{ik}} \; := \; -c_i(T)(1-q_{ik}) - \sum_{r\in\mathcal{R}} \lambda_r R_{ri} q_{ik} \tag{27}$$

It follows that a necessary and sufficient condition for the optimality of (x^λ, u^λ) for the Lagrangean problem (15)–(17) is that the following hold (Fleming and Rishel 1975, Theorem II.11.6):

$$x^\lambda(t) \; = \; x_0 + \int_0^t f(x^\lambda(\tau), u^\lambda(\tau), \tau)\, d\tau \quad \text{for all } t \in [0,T] \tag{28}$$

$$\begin{aligned}\frac{d\psi^\lambda}{dt}(t) \; &= \; -\nabla_x H(x^\lambda(t), u^\lambda(t), t, \psi^\lambda(t)) \\ &= \; -g_x(t) - f_x(t)^T \psi^\lambda(t) \quad \text{for all } t \in (0,T)\end{aligned} \tag{29}$$

$$\psi^\lambda(T) \; = \; \nabla h^\lambda(x^\lambda(T)) \; = \; h_x^\lambda \tag{30}$$

$$H(x^\lambda(t), u^\lambda(t), t, \psi^\lambda(t)) \; = \; \max_{\{u\in[0,1]^m \,:\, D(t)u\le d(t)\}} H(x^\lambda(t), u, t, \psi^\lambda(t)) \quad \text{for all } t \in [0,T] \tag{31}$$

The function $\psi^\lambda(t)$ gives the rate of change of the optimal value function $V^\lambda(x_t, b, t)$ with respect to x_t at t, that is, $\psi^\lambda(t) = \nabla_{x_t} V^\lambda(x^\lambda(t), b, t)$. The conditions (28)–(31) are often called Pontryagin's maximum principle.

First, consider the adjoint equation (29). The equation for component ik is

$$\frac{d\psi^\lambda_{ik}}{dt}(t) \; = \; -g_{x_{ik}}(t) - f_{x_{ik,\cdot}}(t)\psi^\lambda(t) \; = \; c_i(t)\mu_{ik}(t) + \mu_{ik}(t)\psi^\lambda_{ik}(t) \quad \text{for all } t \in (0,T) \tag{32}$$

where $f_{x_{ik,\cdot}}(t)$ denotes the row vector corresponding to component ik of matrix $f_x(t)$. The boundary condition is

$$\psi^\lambda_{ik}(T) \; = \; h^\lambda_{x_{ik}} \; = \; -c_i(T)(1-q_{ik}) - \sum_{r\in\mathcal{R}} \lambda_r R_{ri} q_{ik} \tag{33}$$

The solution of the adjoint equation is

$$
\begin{aligned}
\psi_{ik}^{\lambda}(t) &= \exp\left(-\int_t^T \mu_{ik}(s)\,ds\right)\left[\psi_{ik}^{\lambda}(T) - \int_t^T c_i(\tau)\mu_{ik}(\tau)\exp\left(\int_\tau^T \mu_{ik}(s)\,ds\right)\,d\tau\right] \\
&= -\exp\left(-\int_t^T \mu_{ik}(s)\,ds\right)\left[c_i(T)(1-q_{ik}) + \sum_{r\in\mathcal{R}} \lambda_r R_{ri} q_{ik}\right. \\
&\qquad \left. + \int_t^T c_i(\tau)\mu_{ik}(\tau)\exp\left(\int_\tau^T \mu_{ik}(s)\,ds\right)\,d\tau\right] \qquad (34)
\end{aligned}
$$

Note that if $\lambda \geq 0$, then $\psi_{ik}^{\lambda}(t) \leq 0$, which establishes the intuitive result that $\partial V^{\lambda}(x^{\lambda}(t), b, t)/\partial x_{ik} \leq 0$, that is, the optimal objective value over the interval $[t, T]$ decreases if more units have been sold by time t.

Second, consider the optimization problem (31), for given dual multipliers λ, to determine an optimal decision for each t. The objective function is given by

$$
\begin{aligned}
H(x^{\lambda}(t), u, t, \psi^{\lambda}(t)) &= g(x^{\lambda}(t), u, t) + f(x^{\lambda}(t), u, t)^T \psi^{\lambda}(t) \\
&= g_x(t)^T x^{\lambda}(t) + g_u(t)^T u + x^{\lambda}(t)^T f_x(t)^T \psi^{\lambda}(t) + u^T f_u(t)^T \psi^{\lambda}(t) \qquad (35)
\end{aligned}
$$

Note that $\psi^{\lambda}(t)$ does not depend on u. Thus, optimization problem (31) is equivalent to the following optimization problem for each $t \in [0, T]$.

$$
\max_{u\in[0,1]^m} \quad \left[g_u(t)^T + \psi^{\lambda}(t)^T f_u(t)\right] u \qquad (36)
$$

$$
\text{subject to} \quad D(t)u \leq d(t) \qquad (37)
$$

Optimization problem (36)–(37) is particularly easy, for the following reasons.

1. None of $g_u(t)$, $\psi^{\lambda}(t)$, or $f_u(t)$ depends on u, and thus problem (36)–(37) is a linear program.

2. Note that $\psi^{\lambda}(t)$ does not depend on $x^{\lambda}(t)$, and thus we do not need to know the state at time t to solve problem (36)–(37).

3. Furthermore, (37) is a set of independent knapsack constraints of the form (4); one knapsack constraint for each $k \in \mathcal{K}$. Thus the problem (36)–(37) can be solved by solving a separate subproblem for each $k \in \mathcal{K}$.

Thus, an optimal solution of (36)–(37) is obtained as follows. For each $k \in \mathcal{K}$, if

$$
\max_{i\in\mathcal{I}} \left\{g_{u_{ik}}(t) + \psi_{ik}^{\lambda}(t) f_{u_{ik},ik}(t)\right\} > 0
$$

then let

$$
i_k^{\lambda}(t) \in \arg\max_{i\in\mathcal{I}} \left\{g_{u_{ik}}(t) + \psi_{ik}^{\lambda}(t) f_{u_{ik},ik}(t)\right\} \qquad (38)
$$

and set

$$
u_{i_k^{\lambda}(t),k}^{\lambda}(t) = 1
$$

and

$$
u_{ik}^{\lambda}(t) = 0 \quad \text{for all } i \in \mathcal{I} \setminus \{i_k^{\lambda}(t)\}
$$

Otherwise, set

$$
u_{ik}^{\lambda}(t) = 0 \quad \text{for all } i \in \mathcal{I}
$$

Any consistent tie-breaking rule is sufficient for u^λ_{ik} to be Lebesgue measurable, for example, by ordering the elements of $\mathcal{I}$ and choosing, for each t, $i^\lambda_k(t)$ to be the lowest indexed element of $\mathcal{I}$ that attains the maximum in (38).

Next, (28) can be solved to obtain the optimal state trajectory x^λ corresponding to the dual multipliers λ. The equation for component ik is

$$\frac{dx^\lambda_{ik}}{dt}(t) \;=\; \alpha_k(t)p_{ik}(t)u^\lambda_{ik}(t) - \mu_{ik}(t)x^\lambda_{ik}(t)$$

The solution is

$$\begin{aligned} x^\lambda_{ik}(t) &= \exp\left(-\int_0^t \mu_{ik}(s)\,ds\right)x_{0,ik} + \int_0^t \alpha_k(\tau)p_{ik}(\tau)u^\lambda_{ik}(\tau)\exp\left(-\int_\tau^t \mu_{ik}(s)\,ds\right)\,d\tau \\ &= \exp\left(-\int_0^t \mu_{ik}(s)\,ds\right)x_{0,ik} + \int_{\mathcal{T}^\lambda_{ik}(t)} \alpha_k(\tau)p_{ik}(\tau)\exp\left(-\int_\tau^t \mu_{ik}(s)\,ds\right)\,d\tau \end{aligned} \tag{39}$$

where $\mathcal{T}^\lambda_{ik}(t) := \left\{\tau \in (0,t) : u^\lambda_{ik}(\tau) = 1\right\}$; that is, $\mathcal{T}^\lambda_{ik}(t) := \left\{\tau \in (0,t) : i^\lambda_k(t) = i\right\}$. The factor $\exp\left(-\int_\tau^t \mu_{ik}(s)\,ds\right)$ is an exponential cancellation factor, similar to a discount factor, that discounts the booking rate at time τ with the proportion of cancellations from time τ to time t.

5 Numerical Solution of the Lagrangean Problem

Before we discuss the solution of the Lagrangean dual problem associated with (15), we make a few comments regarding the computational aspects of solving (28)–(31). As shown in (34), (38), and (39), the solution of (28)–(31) involves a number of one-dimensional linear differential equations, or a number of one-dimensional integrals, which often are easy to solve using standard numerical methods.

In specific applications, special structure may allow the solutions to be computed even more efficiently. For example, suppose that the functions $s_i(t)$, $c_i(t)$, $\alpha_k(t)$, $p_{ik}(t)$, and $\mu_{ik}(t)$ are modelled as being piecewise constant, with $0 = t^0_{ik}, t^1_{ik}, \ldots, t^{J_{ik}}_{ik} = T$ being the points of discontinuity of one or more of the mentioned functions. On each piecewise constant interval (t^{j-1}_{ik}, t^j_{ik}), let $s_i(t)$, $c_i(t)$, $\alpha_k(t)$, $p_{ik}(t)$, and $\mu_{ik}(t)$ be given by s^j_i, c^j_i, α^j_k, p^j_{ik}, and μ^j_{ik}, respectively. Then $\psi^\lambda_{ik}(t)$ can be computed inductively, by letting $j = J_{ik}, J_{ik}-1, \ldots, 1$, and for $t \in [t^{j-1}_{ik}, t^j_{ik}]$,

$$\psi^\lambda_{ik}(t) \;=\; e^{-\mu^j_{ik}(t^j_{ik}-t)}\left[\psi^\lambda_{ik}(t^j_{ik}) + c^j_i\left(1 - e^{\mu^j_{ik}(t^j_{ik}-t)}\right)\right] \tag{40}$$

where $\psi^\lambda_{ik}(T)$ is given by (33).

Next, consider the determination of an optimal decision $u^\lambda(t)$ corresponding to the dual multipliers λ, for each t, as in (38). For $t \in [t^{j-1}_{ik}, t^j_{ik}]$, the coefficient of u_{ik} in the objective function (36) is

$$\begin{aligned} C^j_{ik}(t) &:= g_{u_{ik}}(t) + \psi^\lambda_{ik}(t)f_{u_{ik},ik}(t) \\ &= s^j_i\alpha^j_k p^j_{ik} + e^{-\mu^j_{ik}(t^j_{ik}-t)}\left[\psi^\lambda_{ik}(t^j_{ik}) + c^j_i\left(1 - e^{\mu^j_{ik}(t^j_{ik}-t)}\right)\right]\alpha^j_k p^j_{ik} \\ &= \left[s^j_i - c^j_i\right]\alpha^j_k p^j_{ik} + \alpha^j_k p^j_{ik} e^{-\mu^j_{ik}t^j_{ik}}\left[\psi^\lambda_{ik}(t^j_{ik}) + c^j_i\right]e^{\mu^j_{ik}t} \end{aligned}$$

Let

$$\begin{aligned} \beta^j_{ik} &:= \left[s^j_i - c^j_i\right]\alpha^j_k p^j_{ik} \\ \gamma^j_{ik} &:= \alpha^j_k p^j_{ik} e^{-\mu^j_{ik}t^j_{ik}}\left[\psi^\lambda_{ik}(t^j_{ik}) + c^j_i\right] \end{aligned}$$

Thus

$$C^j_{ik}(t) \;=\; \beta^j_{ik} + \gamma^j_{ik} e^{\mu^j_{ik} t}$$

Consider any two product-price pairs i_1 and i_2. The equation

$$C^j_{i_1 k}(t) \;=\; C^j_{i_2 k}(t)$$

has zero, one, or two solutions. It is easy to determine the values of t where $C^j_{i_1 k}(t) = C^j_{i_2 k}(t)$, where $C^j_{i_1 k}(t) < C^j_{i_2 k}(t)$, and where $C^j_{i_1 k}(t) > C^j_{i_2 k}(t)$. This way, it is easy to determine the set of time intervals $\mathcal{T}^\lambda_{ik}(t) \subset (0,t)$ over which it is optimal to offer product-price pair i to customer class k, or equivalently, an optimal product-price pair $i^\lambda_k(t)$ for each customer class k and time t. Let $\mathcal{T}^\lambda_{ik}(T) = \bigcup_{l=1}^{\ell^\lambda_{ik}} (\tau^{\lambda l1}_{ik}, \tau^{\lambda l2}_{ik})$ denote the set of time intervals over which it is optimal to offer product-price pair i to customer class k, with $\tau^{\lambda 11}_{ik} < \tau^{\lambda 12}_{ik} \le \tau^{\lambda 21}_{ik} < \cdots \le \tau^{\lambda \ell^\lambda_{ik} 1}_{ik} < \tau^{\lambda \ell^\lambda_{ik} 2}_{ik}$. Also, for each such interval $(\tau^{\lambda l1}_{ik}, \tau^{\lambda l2}_{ik})$, let $j^{\lambda l}_{ik}$ denote the index of the piecewise constant interval that subinterval $(\tau^{\lambda l1}_{ik}, \tau^{\lambda l2}_{ik})$ belongs to, that is, $(\tau^{\lambda l1}_{ik}, \tau^{\lambda l2}_{ik}) \subseteq (t^{j^{\lambda l}_{ik}-1}_{ik}, t^{j^{\lambda l}_{ik}}_{ik})$. Then, for each $t \in (\tau^{\lambda l1}_{ik}, \tau^{\lambda l2}_{ik})$, the values $s_i(t)$, $c_i(t)$, $\alpha_k(t)$, $p_{ik}(t)$, and $\mu_{ik}(t)$ are given by $s^{j^{\lambda l}_{ik}}_i$, $c^{j^{\lambda l}_{ik}}_i$, $\alpha^{j^{\lambda l}_{ik}}_k$, $p^{j^{\lambda l}_{ik}}_{ik}$, and $\mu^{j^{\lambda l}_{ik}}_{ik}$, respectively.

Next we determine the optimal state trajectory x^λ. First focus on the sales of a product during a time interval (τ_1, τ_2), during which the product is sold at rate αp and previous sales are cancelled at rate μ. Let $y(t)$ denote the quantity of the product that remains sold at time $t \in (\tau_1, \tau_2)$. Then y satisfies the differential equation $dy(t)/dt = \alpha p - \mu y(t)$ with $y(\tau_1) = 0$. The solution is $y(t) = [1 - \exp(-\mu(t - \tau_1))]\alpha p/\mu$ if $\mu > 0$, and $y(t) = [t - \tau_1]\alpha p$ if $\mu = 0$. (Note that if $\mu > 0$, then $\lim_{t\to\infty} y(t) = \alpha p/\mu$, that is, the equilibrium quantity sold at which the sales rate equals the cancellation rate is $y(\infty) = \alpha p/\mu$.) For $i \in \mathcal{I}$, $k \in \mathcal{K}$, $l \in \{1, \ldots, \ell^\lambda_{ik}\}$, and $(\tau_1, \tau_2) \subseteq (\tau^{\lambda l1}_{ik}, \tau^{\lambda l2}_{ik})$, let $y^\lambda_{ik}(\tau_1, \tau_2)$ denote the quantity of product-price pair i sold during (τ_1, τ_2) that remains sold at time τ_2. Then

$$y^\lambda_{ik}(\tau_1, \tau_2) \;=\; \left[1 - \exp\left(-\mu^{j^{\lambda l}_{ik}}_{ik} (\tau_2 - \tau_1)\right)\right] \frac{\alpha^{j^{\lambda l}_{ik}}_k p^{j^{\lambda l}_{ik}}_{ik}}{\mu^{j^{\lambda l}_{ik}}_{ik}}$$

if $\mu^{j^{\lambda l}_{ik}}_{ik} > 0$, and

$$y^\lambda_{ik}(\tau_1, \tau_2) \;=\; [\tau_2 - \tau_1]\, \alpha^{j^{\lambda l}_{ik}}_k p^{j^{\lambda l}_{ik}}_{ik}$$

if $\mu^{j^{\lambda l}_{ik}}_{ik} = 0$. Also, for $i \in \mathcal{I}$, $k \in \mathcal{K}$, and $(t_1, t_2) \subset [0, T]$ with $t_1 \in [t^{j^1-1}_{ik}, t^{j^1}_{ik}]$ and $t_2 \in [t^{j^2-1}_{ik}, t^{j^2}_{ik}]$ let

$$\nu_{ik}(t_1, t_2) \;:=\; \exp\left(-\int_{t_1}^{t_2} \mu_{ik}(s)\, ds\right)$$

denote the exponential cancellation factor over (t_1, t_2). If $j^1 < j^2$, then

$$\nu_{ik}(t_1, t_2) \;=\; \exp\left(-\mu^{j^1}_{ik}\left(t^{j^1}_{ik} - t_1\right) - \sum_{j=j^1+1}^{j^2-1} \mu^j_{ik}\left(t^j_{ik} - t^{j-1}_{ik}\right) - \mu^{j^2}_{ik}\left(t_2 - t^{j^2-1}_{ik}\right)\right)$$

and if $j^1 = j^2$, then

$$\nu_{ik}(t_1, t_2) \;=\; \exp\left(-\mu^{j^1}_{ik}(t_2 - t_1)\right)$$

Then the optimal state trajectory x^λ corresponding to dual multipliers $\lambda \in \mathbb{R}^{|\mathcal{R}|}$ is given by

$$x^\lambda_{ik}(t) \;=\; x_{0,ik}\nu_{ik}(0,t) \;+\; \sum_{\{l\,:\,\tau^{\lambda l2}_{ik} \le t\}} y^\lambda_{ik}(\tau^{\lambda l1}_{ik}, \tau^{\lambda l2}_{ik})\nu_{ik}(\tau^{\lambda l2}_{ik}, t) \;+\; \sum_{\{l\,:\,\tau^{\lambda l1}_{ik} < t < \tau^{\lambda l2}_{ik}\}} y^\lambda_{ik}(\tau^{\lambda l1}_{ik}, t) \quad (41)$$

In particular, the final state $x^\lambda(T)$, which appears in the capacity constraints (3), is given by

$$x_{ik}^\lambda(T) \quad = \quad x_{0,ik}\nu_{ik}(0,T) \; + \; \sum_{l=1}^{\ell_{ik}^\lambda} y_{ik}^\lambda(\tau_{ik}^{\lambda l1}, \tau_{ik}^{\lambda l2})\nu_{ik}(\tau_{ik}^{\lambda l2}, T) \tag{42}$$

The optimal value $L^*(\lambda)$ of the Lagrangean problem (the Lagrangean dual function value) can be computed next, as follows. First we consider the cancellation refunds. Again we focus on the sales of a product during a time interval (τ_1, τ_2), during which the product is sold at rate αp and previous sales are cancelled at rate μ. Let $x(t)$ denote the quantity of the product that remains sold at time $t \in (\tau_1, \tau_2)$, with $x(\tau_1) = x_1$. Then x satisfies the differential equation $dx(t)/dt = \alpha p - \mu x(t)$. The solution is $x(t) = \alpha p/\mu + [x_1 - \alpha p/\mu]\exp(-\mu(t - \tau_1))$, if $\mu > 0$, and $x(t) = x_1 + \alpha p[t - \tau_1]$ if $\mu = 0$. Let c denote the cancellation refund during (τ_1, τ_2). Then the cancellation refund rate at time t is equal to $c\mu x(t)$. Let $z(t) := \int_{\tau_1}^t c\mu x(s)\,ds$ denote the total amount of cancellation refunds accumulated during (τ_1, t). Then

$$z(t) \quad = \quad c\alpha p\,(t - \tau_1) + c\left[x_1 - \frac{\alpha p}{\mu}\right][1 - \exp\left(-\mu\,(t - \tau_1)\right)]$$

if $\mu > 0$, and $z(t) = 0$ if $\mu = 0$. Suppose the product is not sold during (τ_1, τ_2), and let $\tilde{x}(t)$ denote the quantity of the product that remains sold at time $t \in (\tau_1, \tau_2)$. Then $\tilde{x}(t) = x_1 \exp(-\mu(t - \tau_1))$,

$$\tilde{z}(t) \quad := \quad \int_{\tau_1}^t c\mu\tilde{x}(s)\,ds \quad = \quad cx_1\,[1 - \exp\left(-\mu\,(t - \tau_1)\right)]$$

if $\mu > 0$, and $\tilde{z}(t) = 0$ if $\mu = 0$. Consider a product-price pair i, a customer class k, and a subinterval $(t_1, t_2) \subseteq (t_{ik}^{j-1}, t_{ik}^j)$ such that $u_{ik}^\lambda(t) = 0$ for all $t \in (t_1, t_2)$. Let $\tilde{z}_{ik}^\lambda(t_1, t_2)$ denote the amount of cancellation refunds for product-price pair i and customer class k resulting from cancellations during time interval (t_1, t_2). Then $\tilde{z}_{ik}^\lambda(t_1, t_2) = 0$ if $\mu_{ik}^j = 0$, and

$$\tilde{z}_{ik}^\lambda(t_1, t_2) \quad = \quad c_i^j x_{ik}^\lambda(t_1)\left[1 - \exp\left(-\mu_{ik}^j\,(t_2 - t_1)\right)\right]$$

if $\mu_{ik}^j > 0$.

Also, consider a product-price pair i, a customer class k, and a subinterval $(\tau_{ik}^{\lambda l1}, \tau_{ik}^{\lambda l2}) \subseteq (t_{ik}^{j-1}, t_{ik}^j)$, where $j := j_{ik}^{\lambda l}$. Let $z_{ik}^\lambda(\tau_{ik}^{\lambda l1}, \tau_{ik}^{\lambda l2})$ denote the amount of cancellation refunds for product-price pair i and customer class k resulting from cancellations during time interval $(\tau_{ik}^{\lambda l1}, \tau_{ik}^{\lambda l2})$. Then $z_{ik}^\lambda(\tau_{ik}^{\lambda l1}, \tau_{ik}^{\lambda l2}) = 0$ if $\mu_{ik}^j = 0$, and

$$z_{ik}^\lambda(\tau_{ik}^{\lambda l1}, \tau_{ik}^{\lambda l2}) \quad = \quad c_i^j \alpha_k^j p_{ik}^j\left(\tau_{ik}^{\lambda l2} - \tau_{ik}^{\lambda l1}\right) + c_i^j\left[x_{ik}^\lambda(\tau_{ik}^{\lambda l1}) - \frac{\alpha_k^j p_{ik}^j}{\mu_{ik}^j}\right]\left[1 - \exp\left(-\mu_{ik}^j\left(\tau_{ik}^{\lambda l2} - \tau_{ik}^{\lambda l1}\right)\right)\right]$$

if $\mu_{ik}^j > 0$.

Next, the total amount Z_{ik}^λ of cancellation refunds for product-price pair i and customer class

k during $[0, T]$ is given by

$$\begin{aligned}
Z_{ik}^{\lambda} &= \sum_{j'=1}^{j_{ik}^{\lambda 1}-1} \tilde{z}_{ik}^{\lambda}(t_{ik}^{j'-1}, t_{ik}^{j'}) \;+\; \tilde{z}_{ik}^{\lambda}(t_{ik}^{j_{ik}^{\lambda 1}-1}, \tau_{ik}^{\lambda 11}) \;+ \sum_{l=1}^{\ell_{ik}^{\lambda}} z_{ik}^{\lambda}(\tau_{ik}^{\lambda l1}, \tau_{ik}^{\lambda l2}) \\
&+ \sum_{l=1}^{\ell_{ik}^{\lambda}-1} \Bigg[\tilde{z}_{ik}^{\lambda}(\tau_{ik}^{\lambda l2}, \min\{t_{ik}^{j_{ik}^{\lambda l}}, \tau_{ik}^{\lambda, l+1,1}\}) \;+ \sum_{j'=j_{ik}^{\lambda l}+1}^{j_{ik}^{\lambda, l+1}-1} \tilde{z}_{ik}^{\lambda}(t_{ik}^{j'-1}, t_{ik}^{j'}) \\
&\qquad + \; \tilde{z}_{ik}^{\lambda}(\max\{t_{ik}^{j_{ik}^{\lambda, l+1}}, \tau_{ik}^{\lambda l2}\}, \tau_{ik}^{\lambda, l+1,1}) \Bigg] \\
&+ \tilde{z}_{ik}^{\lambda}(\tau_{ik}^{\lambda \ell_{ik}^{\lambda} 2}, t_{ik}^{j_{ik}^{\lambda \ell_{ik}^{\lambda}}}) \;+ \sum_{j'=j_{ik}^{\lambda \ell_{ik}^{\lambda}}+1}^{J_{ik}} \tilde{z}_{ik}^{\lambda}(t_{ik}^{j'-1}, t_{ik}^{j'})
\end{aligned}$$

Then the optimal value $L^*(\lambda)$ of the Lagrangean problem is given by

$$\begin{aligned}
L^*(\lambda) &= \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \int_0^T \Big[s_i(t)\alpha_k(t)p_{ik}(t)u_{ik}^{\lambda}(t) - c_i(t)\mu_{ik}(t)x_{ik}^{\lambda}(t) \Big] \; dt \\
&\quad - \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} c_i(T)(1 - q_{ik})x_{ik}^{\lambda}(T) + \lambda^T [b - RQx^{\lambda}(T)] \\
&= \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} \sum_{l=1}^{\ell_{ik}^{\lambda}} s_i^{j_{ik}^{\lambda l}} \alpha_k^{j_{ik}^{\lambda l}} p_{ik}^{j_{ik}^{\lambda l}} \Big[\tau_{ik}^{\lambda l2} - \tau_{ik}^{\lambda l1} \Big] \; - \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} Z_{ik}^{\lambda} \\
&\quad - \sum_{i \in \mathcal{I}} \sum_{k \in \mathcal{K}} c_i(T)(1 - q_{ik})x_{ik}^{\lambda}(T) \;+ \sum_{r \in \mathcal{R}} \lambda_r \left[b_r - \sum_{i \in \mathcal{I}} R_{ri} \sum_{k \in \mathcal{K}} q_{ik} x_{ik}^{\lambda}(T) \right] \qquad (43)
\end{aligned}$$

6 The Lagrangean Dual Problem

In this section we investigate the Lagrangean dual problem

$$\inf_{\lambda \geq 0} \; L^*(\lambda) \qquad (44)$$

Several issues are of interest, including the following:

1. The relation between the original (primal) problem (1)–(4) and the Lagrangean dual problem (44).
2. The existence of an optimal solution for the Lagrangean dual problem (44).
3. How to compute an optimal (or ε-optimal) solution λ^* for the Lagrangean dual problem (44).
4. How to compute an optimal (or ε-optimal) solution (x^*, u^*) for the primal problem (1)–(4).

These issues are addressed in this section.

6.1 Strong Duality

Recall that for any primal solution (x,u) that satisfies $RQx(T) \le b$ and any feasible dual solution $\lambda \ge 0$,

$$L(x,u,\lambda) \;:=\; \int_0^T g(x(t),u(t),t)\,dt + h(x(T)) + \lambda^T[b - RQx(T)] \;\ge\; \int_0^T g(x(t),u(t),t)\,dt + h(x(T))$$

Thus the following weak duality result is obtained:

$$L^*(\lambda) \;:=\; \sup_{\substack{x\in\mathcal{X},u\in\mathcal{U} \\ \text{subject to (2),(4)}}} L(x,u,\lambda) \;\ge\; \sup_{\substack{x\in\mathcal{X},u\in\mathcal{U} \\ \text{subject to (2),(3),(4)}}} \int_0^T g(x(t),u(t),t)\,dt + h(x(T)) \tag{45}$$

for all $\lambda \ge 0$, and hence

$$\inf_{\lambda\ge 0} L^*(\lambda) \;\ge\; \sup_{\substack{x\in\mathcal{X},u\in\mathcal{U} \\ \text{subject to (2),(3),(4)}}} \int_0^T g(x(t),u(t),t)\,dt + h(x(T)) \tag{46}$$

The next question is whether equality holds in (46), that is, whether strong duality holds, or whether there is a duality gap. To answer this question, recall the following result from convex analysis (Rockafellar 1970, Bertsekas 1996, Hiriart-Urruty and Lemaréchal 1993, Bonnans and Shapiro 2000): Consider primal problem

$$\sup\{F(y) \;:\; y \in \mathcal{Y},\, G(y) \ge 0\}$$

where $\mathcal{Y}$ is a convex subset of a (possibly infinite dimensional) vector space, $F : \mathcal{Y} \mapsto \mathbb{R}$ is concave, and $G : \mathcal{Y} \mapsto \mathbb{R}^r$ is concave. Consider Lagrangean dual function $L^* : \mathbb{R}^r \mapsto \mathbb{R}$, given by

$$L^*(\lambda) \;:=\; \sup\{F(y) + \lambda^T G(y) \;:\; y \in \mathcal{Y}\}$$

Suppose there exists a $\lambda \ge 0$ such that $L^*(\lambda) < \infty$, and a $y \in \mathcal{Y}$ such that $G(y) > 0$. Then

$$\inf\{L^*(\lambda) \;:\; \lambda \ge 0\} \;=\; \sup\{F(y) \;:\; y \in \mathcal{Y},\, G(y) \ge 0\}$$

that is, there is no duality gap. In addition, the set of optimal solutions of the Lagrangean dual problem $\inf\{L^*(\lambda) : \lambda \ge 0\}$ is a nonempty convex set.

In the primal problem (1)–(4),

$$\mathcal{Y} \;:=\; \left\{(x,u) \in \mathcal{X}\times\mathcal{U} \;:\; x(t) = x_0 + \int_0^t f(x(\tau),u(\tau),\tau)\,d\tau,\; D(t)u(t) \le d(t),\; \forall\, t \in [0,T]\right\}$$

Note that $\mathcal{X}$ and $\mathcal{U}$ are convex, thus $\mathcal{X}\times\mathcal{U}$ is convex. Consider any $(x^1,u^1),(x^2,u^2) \in \mathcal{Y}$. Let $\theta \in [0,1]$, and let $(\bar{x},\bar{u}) := \theta(x^1,u^1) + (1-\theta)(x^2,u^2) \in \mathcal{X}\times\mathcal{U}$. Then it follows from the linearity of f in (x,u) that

$$\bar{x}(t) \;=\; x_0 + \int_0^t f(\bar{x}(\tau),\bar{u}(\tau),\tau)\,d\tau$$

for all $t \in [0,T]$. Also from linearity, $D(t)\bar{u}(t) \le d(t)$ for all $t \in [0,T]$. Thus $(\bar{x},\bar{u}) \in \mathcal{Y}$, and hence $\mathcal{Y}$ is convex. Further, g and h are affine in (x,u), and thus $F(x,u) := \int_0^T g(x(t),u(t),t)\,dt + h(x(T))$

is concave. Also, $G(x,u) := b - RQx(T)$ is affine in (x,u), and thus concave. It was noted in Section 4 that $L^*(\lambda) < \infty$ for all $\lambda \in \mathbb{R}^{|\mathcal{R}|}$.

The remaining question is whether there exists a $(x,u) \in \mathcal{Y}$ such that $G(x,u) > 0$, that is, such that a positive amount of each resource remains at time T. Intuitively, $G(x(u),u) \leq G(x(0),0)$ for all $(x,u) \in \mathcal{Y}$, where $x(u)$ is used here to indicate that x depends on u, and thus there exists a $(x,u) \in \mathcal{Y}$ such that $G(x,u) > 0$ if and only if $G(x(0),0) > 0$, that is, if and only if not selling any products during $[0,T]$ leaves a positive amount of each resource remaining at time T. The intuition can be confirmed as follows: Let $\lambda \in (0,\infty)^{|\mathcal{R}|}$. Consider the optimal control problem with only the Lagrangean term in the objective:

$$\sup_{x\in\mathcal{X},u\in\mathcal{U}} \quad \lambda^T G(x,u) \tag{47}$$

$$\text{subject to} \quad x(t) \;=\; x_0 + \int_0^t f(x(\tau),u(\tau),\tau)\,d\tau \quad \text{for all } t \in [0,T] \tag{48}$$

$$D(t)u(t) \;\leq\; d(t) \quad \text{for all } t \in [0,T] \tag{49}$$

The Hamiltonian is given by

$$H(x,u,t,\psi) \;:=\; f(x,u,t)^T \psi \tag{50}$$

A necessary and sufficient condition for the optimality of (x^λ, u^λ) for the problem (47)–(49) is that the following hold:

$$\begin{aligned}
x^\lambda(t) &= x_0 + \int_0^t f(x^\lambda(\tau),u^\lambda(\tau),\tau)\,d\tau \quad \text{for all } t \in [0,T] \\
\frac{d\psi^\lambda}{dt}(t) &= -\nabla_x H(x^\lambda(t),u^\lambda(t),t,\psi^\lambda(t)) \;=\; -f_x(t)^T\psi^\lambda(t) \quad \text{for all } t \in (0,T) \\
\psi^\lambda(T) &= -Q^T R^T \lambda \\
H(x^\lambda(t),u^\lambda(t),t,\psi^\lambda(t)) &= \max_{\{u\in[0,1]^m \,:\, D(t)u\leq d(t)\}} H(x^\lambda(t),u,t,\psi^\lambda(t)) \quad \text{for all } t \in [0,T]
\end{aligned}$$

The solution of the adjoint equation is

$$\psi_{ik}^\lambda(t) \;=\; \exp\left(-\int_t^T \mu_{ik}(s)\,ds\right)\psi_{ik}^\lambda(T) \;=\; -\exp\left(-\int_t^T \mu_{ik}(s)\,ds\right)\sum_{r\in\mathcal{R}} \lambda_r R_{ri} q_{ik}$$

Note that $\psi_{ik}^\lambda(t) \leq 0$ for all t. Thus, the coefficient of u_{ik} in

$$H(x^\lambda(t),u,t,\psi^\lambda(t)) \;=\; \sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}} \Big(\alpha_k(t)p_{ik}(t)u_{ik} - \mu_{ik}(t)x_{ik}^\lambda(t)\Big)\,\psi_{ik}^\lambda(t)$$

is nonpositive, and thus $u^\lambda = 0$ and the corresponding x^λ is an optimal solution. That is, $u^\lambda = 0$ maximizes each individual component of G, and thus $G(x(u),u) \leq G(x(0),0)$. In addition, if $\sum_{r\in\mathcal{R}} \lambda_r R_{ri} q_{ik} > 0$, that is, if selling product-price pair i to customer class k uses any resource at all, then $u_{ik}^\lambda(t) = 0$ is the unique optimal solution for all t such that $\alpha_k(t)p_{ik}(t) > 0$, and at times t such that $\alpha_k(t)p_{ik}(t) = 0$, nothing can be sold in any case. Thus all optimal solutions of the above problem are such that nothing is sold.

It is not restrictive to assume that $G(x(0),0) > 0$. One can simply do the following preprocessing to ensure that $G(x(0),0) > 0$, and thus to ensure that strong duality holds. Suppose the problem input has a set $\mathcal{I}'$ of product-price pairs, a set $\mathcal{K}'$ of customer classes, and a set $\mathcal{R}'$ of resources. Compute $G(x(0),0) := b - RQx(0)(T)$. For each resource $r \in \mathcal{R}'$ such that $G(x(0),0)_r < 0$, there is no control that can achieve $G(x(0),0)_r \geq 0$, and thus the best that can be done is not to sell any products that would use resource r. Similarly, for each resource $r \in \mathcal{R}'$ such that $G(x(0),0)_r = 0$, the only feasible control is not to sell any products that would use resource r. Thus, as preprocessing step, $G(x(0),0)$ is computed, and for each resource $r \in \mathcal{R}'$ such that $G(x(0),0)_r \leq 0$, u_{ik} is set to zero for each product-price pair $i \in \mathcal{I}'$ and customer class $k \in \mathcal{K}'$ such that $R_{ri}q_{ik} > 0$. Then, let $\mathcal{R} := \{r \in \mathcal{R}' : G(x(0),0)_r > 0\}$, $\mathcal{I} := \mathcal{I}' \setminus \{i \in \mathcal{I}' : G(x(0),0)_r \leq 0$ for some $r \in \mathcal{R}'$ such that $R_{ri}q_{ik} > 0$ for all $k \in \mathcal{K}'\}$, and $\mathcal{K} := \{k \in \mathcal{K}' : \alpha_k(t)p_{ik}(t) > 0$ for some $i \in \mathcal{I}$ and some $t \in [0,T]\}$. The problem with input sets $\mathcal{I}$, $\mathcal{K}$, and $\mathcal{R}$ has a feasible primal solution and no duality gap.

Let $\mathcal{XU}^*$ denote the set of optimal solutions of the primal problem (1)–(4). Let $\Lambda^* := \arg\min\{L^*(\lambda) : \lambda \geq 0\}$ denote the set of optimal solutions of the Lagrangean dual problem; recall that $\Lambda^* \neq \emptyset$. Then $(x^*,u^*,\lambda^*) \in \mathcal{XU}^* \times \Lambda^*$ if and only if

1. (x^*,u^*) is primal feasible, that is, (x^*,u^*) satisfies (2)–(4),
2. λ^* is dual feasible, that is, $\lambda^* \geq 0$,
3. (x^*,u^*) is an optimal solution of the Lagrangean problem (15)–(17) with multiplier λ^*, that is $(x^*,u^*) \in \mathcal{Y}^{\lambda^*}$, and
4. complementary slackness holds, that is $\lambda^*_r G(x^*,u^*)_r = 0$ for each $r \in \mathcal{R}$.

Recall the primal function V^* defined in (10)–(13). It follows from results in convex analysis that V^* is concave in b, and that $\Lambda^* = \partial_b V^*(x_0,b,0)$. That is, a Lagrange multiplier (or bid price in revenue management terminology) $\lambda^* \in \Lambda^*$ is an estimate of the rate of increase in the optimal value with an increase in capacity b. It is interesting to observe the behavior of the Lagrange multipliers over time. Suppose that $(x^*,u^*,\lambda^*) \in \mathcal{XU}^* \times \Lambda^*$, and that optimal trajectory (x^*,u^*) is followed over $[0,t] \subset [0,T]$. Thus, to control the process over the remaining interval $[t,T]$, we would like to solve the optimal control problem (10)–(13), with associated Lagrangean problem (19)–(21) over $[t,T]$, and associated Lagrangean dual problem $\inf_{\lambda\geq 0} V^\lambda(x^*(t),b,t)$. Note that (x^*,u^*) is a feasible solution for the primal problem (10)–(13), $\lambda^* \geq 0$ is a feasible solution for the Lagrangean dual problem $\inf_{\lambda\geq 0} V^\lambda(x^*(t),b,t)$, and complementary slackness $\lambda^*_r G(x^*,u^*)_r = 0$ for all $r \in \mathcal{R}$ holds from the optimality of $(x^*,u^*,\lambda^*) \in \mathcal{XU}^* \times \Lambda^*$. Thus, to show that (x^*,u^*,λ^*) is also optimal over the remaining interval $[t,T]$, it remains to be shown that (x^*,u^*) is an optimal solution of the Lagrangean problem (19)–(21) over $[t,T]$ with multiplier λ^*. This follows from noting that the solution ψ^{λ^*} of the adjoint equation over $[t,T]$ is the same as the part over $[t,T]$ of the solution ψ^{λ^*} over $[0,T]$. Thus, for each $\tau \in [t,T]$, the Hamiltonian $H(x^{\lambda^*}(\tau),u,\tau,\psi^{\lambda^*}(\tau))$ is the same as before, and hence $u^*(\tau)$ still optimizes the Hamiltonian for each $\tau \in [t,T]$. Thus (x^*,u^*) is an optimal solution of the Lagrangean problem (19)–(21) over $[t,T]$ with multiplier λ^*. Therefore, in revenue management terminology, the bid prices remain constant along an optimal trajectory.

6.2 Solving the Lagrangean Dual Problem

Next we address the problem of computing an optimal (or ε-optimal) solution λ^* for the Lagrangean dual problem (44). First recall the following results from convex analysis, as well as from results in previous sections.

1. The Lagrangean dual function $L^*(\lambda) := \sup_{(x,u)\in\mathcal{Y}} \left\{F(x,u) + \lambda^T G(x,u)\right\}$ is the pointwise supremum of a collection (indexed by $(x,u) \in \mathcal{Y}$) of affine functions in λ, and hence is convex. Also, L^* is lower semi-continuous.

2. For $\lambda \in \mathbb{R}^{|\mathcal{R}|}$, let $\mathcal{Y}^\lambda$ denote the set of optimal solutions (x^λ, u^λ) of the Lagrangean problem (15)–(17). Recall that it was shown in Section 4 that, for each $\lambda \in \mathbb{R}^{|\mathcal{R}|}$, $\mathcal{Y}^\lambda \neq \emptyset$. For any $\lambda_1, \lambda_2 \in \mathbb{R}^{|\mathcal{R}|}$,

$$\begin{aligned} L^*(\lambda_2) &\geq F(x^{\lambda_1}, u^{\lambda_1}) + \lambda_2^T G(x^{\lambda_1}, u^{\lambda_1}) \\ &= F(x^{\lambda_1}, u^{\lambda_1}) + \lambda_1^T G(x^{\lambda_1}, u^{\lambda_1}) + \left(\lambda_2^T - \lambda_1^T\right) G(x^{\lambda_1}, u^{\lambda_1}) \\ &= L^*(\lambda_1) + \left(\lambda_2^T - \lambda_1^T\right) G(x^{\lambda_1}, u^{\lambda_1}) \end{aligned}$$

and thus $G(x^{\lambda_1}, u^{\lambda_1})$ is a subgradient of L^* at λ_1. Thus L^* is differentiable at λ only if there is a unique optimal constraint value $G(x^\lambda, u^\lambda)$ of the Lagrangean problem. Hence, L^* is not differentiable in general.

3. Note that even if $\lambda^* \in \Lambda^*$, each optimal solution $(x^{\lambda^*}, u^{\lambda^*})$ of the Lagrangean problem with multiplier λ^* may not be a primal optimal solution. That is, it is possible that $\lambda^* \in \Lambda^*$ and $(x^{\lambda^*}, u^{\lambda^*}) \in \mathcal{Y}^{\lambda^*}$, but $(x^{\lambda^*}, u^{\lambda^*}) \notin \mathcal{XU}^*$, because $(x^{\lambda^*}, u^{\lambda^*})$ is not primal feasible or complementary slackness does not hold. Thus, even if the Lagrangean dual problem $\inf\{L^*(\lambda) : \lambda \geq 0\}$ can be solved and an optimal solution (x^λ, u^λ) of the Lagrangean problem (15)–(17) can be found for each λ, it still remains to be addressed how to find an optimal primal solution.

Thus the Lagrangean dual problem $\inf\{L^*(\lambda) : \lambda \geq 0\}$ is a constrained nonsmooth convex optimization problem.

Several algorithms have been developed for constrained nonsmooth convex optimization problems; see Hiriart-Urruty and Lemaréchal (1993) for a number of these algorithms belonging to the class of bundle methods. One approach is to search for an optimal solution by generating a sequence of solutions; in such an algorithm each solution is generated from the previous solution by choosing a direction and doing a line search from the previous solution in the chosen direction. Another approach is to construct a sequence of approximations $\hat{L}^v$ to the objective function L^*; when the approximation is sufficiently accurate, especially in a neighborhood of an optimal solution, then optimizing the approximating function provides a good solution. For such an algorithm to be practical, it should be easy to construct the approximation, and it should be easier to optimize the approximating function than to optimize the original function directly. For these reasons, a piecewise affine convex (polyhedral) approximation of the form $\hat{L}^v(\lambda) := \max\{A^j + G^{jT}\lambda, : (A^j, G^j) \in \mathcal{B}^v\}$, where $A^j \in \mathbb{R}$, $G^j \in \mathbb{R}^{|\mathcal{R}|}$, and $\mathcal{B}^v$ denotes the set of affine pieces at iteration v, is often proposed. For stopping tests, it is also useful for $\hat{L}^v$ to be a lower bound on L^*: $\hat{L}^v(\lambda) \leq L^*(\lambda)$ for all λ. The affine pieces $A^j + G^{jT}\lambda$ of such an approximating function can be generated as follows. At iteration v, suppose we have an approximation $\hat{L}^v \leq L^*$. Let $\lambda^v \in \arg\min\{\hat{L}^v(\lambda) : \lambda \geq 0\}$, that is, λ^v is an optimal solution of the approximating problem $\min\{\hat{L}^v(\lambda) : \lambda \geq 0\}$, let $G^v \in \partial L^*(\lambda^v)$, that is, G^v is a subgradient of L^* at λ^v, and let $A^v := L^*(\lambda^v) - G^{vT}\lambda^v$. Then let $\mathcal{B}^{v+1} := \mathcal{B}^v \cup \{(A^v, G^v)\}$. Note that, due to the convexity of L^*, $L^*(\lambda) \geq L^*(\lambda^v) + G^{vT}(\lambda - \lambda^v) = A^v + G^{vT}\lambda$ for all λ, and thus $\hat{L}^{v+1}(\lambda) \leq L^*(\lambda)$. Also, $\hat{L}^{v+1}(\lambda^v) = L^*(\lambda^v)$, that is, the approximating function $\hat{L}^{v+1}$ is tight at the new point λ^v. For it to be easy to construct such an approximation $\hat{L}^v$, the following properties are desirable:

1. For any $\lambda \geq 0$, it should be easy to compute the objective value $L^*(\lambda)$. In the primal-dual context, that means it should be easy to compute the optimal value of the Lagrangean problem. For the optimal control problem, it was shown in Section 4 how that can be accomplished by solving the Lagrangean problem.

2. For any $\lambda \geq 0$, it should be easy to compute a subgradient $G \in \partial L^*(\lambda)$. In the primal-dual context, that means it should be easy to compute the values $G(y^\lambda)$ of the dualized

constraints $G(y) \geq 0$ at an optimal solution y^λ of the Lagrangean problem. For the optimal control problem, it was shown in Section 4 how that can be accomplished, by computing an optimal solution (x^λ, u^λ) of the Lagrangean problem, and computing the constraint values $G(x^\lambda, u^\lambda) := b - RQx^\lambda(T)$.

3. It should be easy to solve the approximating problem $\min\{\hat{L}^v(\lambda) : \lambda \geq 0\}$. Note that the approximating problem with piecewise affine objective function $\hat{L}^v$ is the following linear program, and is therefore easy to solve.

$$\begin{array}{rll} \min & z & \\ \text{subject to} & z \geq A^j + G^{jT}\lambda & \text{for all } (A^j, G^j) \in \mathcal{B}^v \\ & z \in \mathbb{R},\ \lambda \in [0,\infty)^{|\mathcal{R}|} & \end{array} \tag{51}$$

   Also note that, because $\hat{L}^v \leq L^*$, the optimal value z^* of the approximating problem (51) provides a lower bound on the dual (and primal) optimal value: $z^* \leq \inf\{L^*(\lambda) : \lambda \geq 0\}$.

4. The number of affine pieces $A^j + G^{jT}\lambda$ required to find a good solution for the original problem should not be prohibitive. This property is more elusive than the previous properties.

The type of algorithm discussed above is called a cutting-plane algorithm (Kelley 1960). A basic cutting-plane algorithm is stated next, in which a compact set $\mathcal{C}$ that contains an optimal solution is specified a-priori.

Cutting-Plane Algorithm
Step 0 (Initialization): Choose a compact set $\mathcal{C}$ that contains an optimal solution, an initial solution $\lambda^0 \in \mathcal{C}$, and a stopping tolerance $\varepsilon \geq 0$. Solve the Lagrangean problem with dual multipliers λ^0, producing objective value $L^*(\lambda^0)$, subgradient $G^0 \in \partial L^*(\lambda^0)$, and $A^0 := L^*(\lambda^0) - G^{0T}\lambda^0$. Let $\mathcal{B}^1 := \{(A^0, G^0)\}$, and let $\hat{L}^1(\lambda) := A^0 + G^{0T}\lambda$. Set $v = 1$.
Step 1 (Approximating Problem): Solve the approximating problem (51), producing optimal solution (z^v, λ^v).
Step 2 (Lagrangean Problem): Solve the Lagrangean problem (15)–(17) with dual multipliers λ^v, producing solution $(x^{\lambda^v}, u^{\lambda^v})$, objective value $L^*(\lambda^v) = L(x^{\lambda^v}, u^{\lambda^v}, \lambda^v)$, subgradient $G^v \in \partial L^*(\lambda^v)$, and $A^v := L^*(\lambda^v) - G^{vT}\lambda^v = F(x^{\lambda^v}, u^{\lambda^v})$.
Step 3 (Stopping Test): If $L^*(\lambda^v) < z^v + \varepsilon$, then stop with ε-optimal dual solution λ^v.
Step 4 (Continue): Let $\mathcal{B}^{v+1} := \mathcal{B}^v \cup \{(A^v, G^v)\}$, and let $\hat{L}^{v+1}(\lambda) := \max\{A^j + G^{jT}\lambda : (A^j, G^j) \in \mathcal{B}^{v+1}\}$. Set $v \leftarrow v + 1$. Go to step 1.

It was shown in Hiriart-Urruty and Lemaréchal (1993) that if $\varepsilon > 0$, then the algorithm stops after a finite number v_ε of iterations, and $L^*(\lambda^{v_\varepsilon}) \leq \inf\{L^*(\lambda) : \lambda \geq 0\} + \varepsilon$, and if $\varepsilon = 0$, then $L^*(\lambda^v), z^v \to \inf\{L^*(\lambda) : \lambda \geq 0\}$ as $v \to \infty$.

However, the basic cutting-plane algorithm has several shortcomings. The performance of the algorithm depends to a large extent on the size of the compact set $\mathcal{C}$ specified initially. An example is given in Hiriart-Urruty and Lemaréchal (1993) that provides some insight into why this happens. Also, it has been observed empirically that the successive iterates λ^v generated by the algorithm can jump around a lot. That is because the iterates λ^v tend to be generated in a part of $\mathcal{C}$ where $\hat{L}^v$ does not give a good approximation to L^*, and not close to an optimal solution until $\hat{L}^v$ has become quite close to L^* over the whole set $\mathcal{C}$. Thus simply choosing a very large set $\mathcal{C}$ that is guaranteed to contain an optimal solution is usually not a good idea. For these reasons several methods have been proposed to stabilize the successive iterates λ^v, and hopefully make the algorithm more efficient. One approach adds a quadratic regularization (penalty) term $\vartheta^v\|\lambda-\lambda^v\|^2$ to the objective function of (51), where $\vartheta^v \geq 0$ is the penalty parameter at iteration v, and λ^v is the stability center at iteration v, for example the best solution found so far. The resulting problem is a quadratic program, which is still quite tractable. Another approach is to adjust the trust region $\mathcal{C}$ between iterations, depending on the properties of the results obtained. This can be accomplished by adding trust region constraints $\|\lambda - \lambda^v\|^2 \leq \Delta^v$ to (51), where $\Delta^v \geq 0$ is the size parameter at iteration v, and λ^v is the stability center as before. The two approaches are equivalent in the sense that for any $\vartheta \geq 0$, there is a $\Delta \geq 0$ such that any optimal solution of the problem with penalty term is also an optimal solution of the problem with trust region constraint, and for any $\Delta^v > 0$, there is a $\vartheta \geq 0$ such that any optimal solution of the problem with trust region constraint is also an optimal solution of the problem with penalty term (Hiriart-Urruty and Lemaréchal 1993). If $\|\cdot\|$ is chosen to be $\|\cdot\|_1$ ($\|\lambda\|_1 := \sum_{r\in\mathcal{R}}|\lambda_r|$) or $\|\cdot\|_\infty$ ($\|\lambda\|_\infty := \max\{|\lambda_1|,\ldots,|\lambda_{|\mathcal{R}|}|\}$), then the problem with penalty term $\vartheta^v\|\lambda-\lambda^v\|$ and the problem with trust region constraint $\|\lambda-\lambda^v\| \leq \Delta^v$ are linear programs. Stabilization methods have been studied by many, including Ruszczyński (1986), Kiwiel (1990), and Moré (1997). The method used in the numerical work discussed in Section 7 is similar to an algorithm studied by Linderoth and Wright (2001), and is discussed next.

The method used has trust region constraints with $\|\cdot\|_\infty$ as norm. Thus the approximating problem is the following linear program:

$$\begin{array}{lll} \min & z & \\ \text{subject to} & z \;\geq\; A^j + G^{jT}\lambda & \text{for all } (A^j, G^j) \in \mathcal{B}^v \\ & \lambda_r^v - \Delta_r^v \;\leq\; \lambda_r \;\leq\; \lambda_r^v + \Delta_r^v & \text{for all } r \in \mathcal{R} \\ & z \in \mathbb{R},\; \lambda \in [0,\infty)^{|\mathcal{R}|} & \end{array} \tag{52}$$

Five important issues to be addressed are the following:

1. How the stability center λ^v should be chosen and updated.
2. How the trust region size parameter Δ^v should be chosen and updated.
3. How the bundle $\mathcal{B}^v$ of affine pieces should be updated.
4. How to perform an appropriate stopping test in the presence of trust region constraints.
5. How to obtain an ε-optimal primal solution from the results when the algorithm is stopped.

To update the stability center λ^v, a distinction is made between major iterations (serious steps), during which the stability center is changed, and minor iterations (null steps), during which the stability center is not changed. Let (z^{vw}, λ^{vw}) denote the optimal solution obtained from solving the approximating problem (52) during major iteration v and minor iteration w, with approximating objective $\hat{L}^{vw}$, stability center $\lambda^v = \lambda^{v0}$, and trust region size parameter Δ^{vw}. The stability center is moved to the new solution λ^{vw} if the actual improvement $L^*(\lambda^{v0}) - L^*(\lambda^{vw})$ is sufficiently good compared with the improvement $L^*(\lambda^{v0}) - z^{vw}$ predicted by the approximation $\hat{L}^{vw}$. (Recall that $L^*(\lambda^{v0}) - z^{vw} \geq 0$.) This is implemented by choosing a move test parameter $\eta_m \in (0,1)$, and moving the stability center to the new solution λ^{vw}, that is, $\lambda^{v+1,0} := \lambda^{vw}$, if

$$L^*(\lambda^{vw}) \quad < \quad L^*(\lambda^{v0}) - \eta_m\left(L^*(\lambda^{v0}) - z^{vw}\right) \tag{53}$$

The next question is how the trust region size parameter Δ^{vw} should be updated. Again the test is based on comparing the actual improvement $L^*(\lambda^{v0}) - L^*(\lambda^{vw})$ obtained from solving the approximating problem (52) with the predicted improvement $L^*(\lambda^{v0}) - z^{vw}$. This is implemented by choosing a contraction test parameter $\eta_c \in (0, \eta_m)$, an expansion test parameter $\eta_e \in (\eta_m, 1)$, a contraction multiplier $m_c \in (0, 1)$, an expansion multiplier $m_e > 1$, and a minimum trust region size parameter $\underline{\Delta} > 0$. If

$$L^*(\lambda^{vw}) \quad > \quad L^*(\lambda^{v0}) - \eta_c \left(L^*(\lambda^{v0}) - z^{vw}\right) \tag{54}$$

then the trust region is contracted:

$$\Delta^{v,w+1} \quad := \quad \max\{m_c \Delta^{vw},\ \underline{\Delta}\} \tag{55}$$

If

$$L^*(\lambda^{vw}) \quad < \quad L^*(\lambda^{v0}) - \eta_e \left(L^*(\lambda^{v0}) - z^{vw}\right) \tag{56}$$

then the trust region is expanded:

$$\Delta^{v+1,0} \quad := \quad m_e \Delta^{vw} \tag{57}$$

There is another situation in which the trust region is expanded, which is discussed as part of the stopping test.

More affine pieces in the bundle $\mathcal{B}^{vw}$ gives a more accurate approximation $\hat{L}^{vw}$. However, more affine pieces also makes the linear program (52) harder to solve. Also, the convergence proofs of cutting-plane algorithms do not require any affine pieces to be retained when the stability center is moved; that is, for the purposes of convergence it would be acceptable to choose $\mathcal{B}^{v1} = \{(A^{v0}, G^{v0})\}$. A more efficient approach may be to retain the active constraints, as well as some constraints that are likely to be active at a future iteration, and to discard the remaining constraints when the stability center is moved. This is implemented by choosing constraint elimination test parameters $\xi > 0$ and $n_1, n_2, n_3, n_4 \geq 0$, with $n_3 < n_1$. If the stability center is moved at iteration vw, and the total number of constraints in the bundle $\mathcal{B}^{vw}$ is more than n_1, and the number of constraints that have been added since the previous iteration during which constraints were eliminated is more than n_2, then the slack $z^{vw} - (A^j + G^{jT}\lambda^{vw})$ is calculated for each constraint $j \in \mathcal{B}^{vw}$. Let $\mathcal{B}_-^{vw} := \{j \in \mathcal{B}^{vw} \ : \ z^{vw} - (A^j + G^{jT}\lambda^{vw}) > \xi(L^*(\lambda^{v0}) - z^{vw})\}$. If $|\mathcal{B}^{vw}| - |\mathcal{B}_-^{vw}| \geq n_3$, then all the constraints in $\mathcal{B}_-^{vw}$ are eliminated from $\mathcal{B}^{vw}$ to give $\mathcal{B}^{v+1,0} := \mathcal{B}^{vw} \setminus \mathcal{B}_-^{vw}$; otherwise, the $|\mathcal{B}^{vw}| - n_3$ constraints in $\mathcal{B}_-^{vw}$ with the greatest slack values are eliminated from $\mathcal{B}^{vw}$ to give $\mathcal{B}^{v+1,0}$. A modification of this approach is to eliminate a constraint only after its slack has been greater than the threshold value of type $\xi(L^*(\lambda^{v0}) - z^{vw})$ at least n_4 times at which it was tested.

For the purpose of stopping tests, several tolerances are used: a tolerance $\varepsilon_p > 0$ on the optimality gap of a primal solution, a tolerance $\varepsilon_d > 0$ on the optimality gap of a dual solution, and a tolerance $\varepsilon_f > 0$ on the infeasibility of a primal solution. When the Lagrangean problem (15)–(17) is solved with dual multipliers λ^{vw}, a primal solution $(x^{\lambda^{vw}}, u^{\lambda^{vw}})$ is obtained. If

$$\int_0^T g(x^{\lambda^{vw}}(t), u^{\lambda^{vw}}(t), t)\, dt + h(x^{\lambda^{vw}}(T)) \quad > \quad L^*(\lambda^{vw}) - \varepsilon_p \tag{58}$$

that is, the objective value of $(x^{\lambda^{vw}}, u^{\lambda^{vw}})$ is within ε_p of the upper bound $L^*(\lambda^{vw})$ on the primal optimal value, and

$$\sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}} R_{ri} q_{ik} x_{ik}^{\lambda^{vw}}(T) \quad < \quad b_r + \varepsilon_f \qquad \text{for each } r \in \mathcal{R} \tag{59}$$

that is, the feasibility error of $(x^{\lambda^{vw}}, u^{\lambda^{vw}})$ is less than ε_f, and the final amount of resource $G(x^{\lambda^{vw}}, u^{\lambda^{vw}})_r > -\varepsilon_f$ for each $r \in \mathcal{R}$, then the algorithm stops with primal solution $(x^{\lambda^{vw}}, u^{\lambda^{vw}})$

and dual solution λ^{vw}. Note that if (58) holds, then

$$(G^{vw})^T \lambda^{vw} \quad = \quad L^*(\lambda^{vw}) - \int_0^T g(x^{\lambda^{vw}}(t), u^{\lambda^{vw}}(t), t)\ dt + h(x^{\lambda^{vw}}(T)) \quad < \quad \varepsilon_p$$

However, the primal feasibility test (59) is needed because positive and negative terms in $(G^{vw})^T \lambda^{vw}$ may cancel out. The primal feasibility test (59) together with $(G^{vw})^T \lambda^{vw} < \varepsilon_p$ imply that complementary slackness holds with some error. Also, if none of the trust region constraints $\lambda_r^{v0} - \Delta_r^{vw} \leq \lambda_r \leq \lambda_r^{v0} + \Delta_r^{vw}$ is active at an optimal solution (z^{vw}, λ^{vw}) of the approximating problem (52), except for those r for which $\lambda_r^{vw} = \lambda_r^{v0} - \Delta_r^{vw} = 0$, then z^{vw} is a lower bound on the optimal dual (and primal) value $\inf\{L^*(\lambda) : \lambda \geq 0\}$. Thus, if $L^*(\lambda^{vw}) < z^{vw} + \varepsilon_d$, and one or more of the trust region constraints are active at (z^{vw}, λ^{vw}) (besides $\lambda_r^{vw} = \lambda_r^{v0} - \Delta_r^{vw} = 0$), then the trust region is expanded as in (57). Next, suppose that $L^*(\lambda^{vw}) < z^{vw} + \varepsilon_d$, and none of the trust region constraints is active at (z^{vw}, λ^{vw}) (except if $\lambda_r^{vw} = \lambda_r^{v0} - \Delta_r^{vw} = 0$). Then the approximating problem (52) is equivalent to

$$\begin{array}{lll} \min & z & \\ \text{subject to} & z \ \geq \ A^j + G^{jT} \lambda & \text{for all } (A^j, G^j) \in \mathcal{B}^{vw} \\ & z \in \mathbb{R},\ \lambda \in [0, \infty)^{|\mathcal{R}|} & \end{array} \tag{60}$$

with dual problem

$$\begin{array}{lll} \max & \displaystyle\sum_{(A^j,G^j)\in\mathcal{B}^{vw}} A^j \theta_j & \\ \text{subject to} & \displaystyle\sum_{(A^j,G^j)\in\mathcal{B}^{vw}} \theta_j \ = \ 1 & \\ & \displaystyle\sum_{(A^j,G^j)\in\mathcal{B}^{vw}} G^j \theta_j \ \geq \ 0 & \\ & \theta_j \geq 0 & \text{for all } (A^j, G^j) \in \mathcal{B}^{vw} \end{array} \tag{61}$$

Note that feasible dual solutions θ of (61) are convex multipliers. Let θ^{vw} denote an optimal solution of (61), and for each $(A^j, G^j) \in \mathcal{B}^{vw}$, let λ^j denote the corresponding dual multipliers, and let (x^j, u^j) denote the corresponding optimal solution of the Lagrangean problem with dual multipliers λ^j that generated $G^j \in \partial L^*(\lambda^j)$ and $A^j := L^*(\lambda^j) - G^{jT} \lambda^j$. Then, it follows from linear programming duality that

$$\begin{aligned} z^{vw} \quad &= \sum_{(A^j,G^j)\in\mathcal{B}^{vw}} A^j \theta_j^{vw} \\ &= \sum_{(A^j,G^j)\in\mathcal{B}^{vw}} \theta_j^{vw} \left[L^*(\lambda^j) - G^{jT} \lambda^j\right] \\ &= \sum_{(A^j,G^j)\in\mathcal{B}^{vw}} \theta_j^{vw} \left[\int_0^T g(x^{\lambda^j}(t), u^{\lambda^j}(t), t)\ dt + h(x^{\lambda^j}(T))\right] \end{aligned}$$

Let

$$(x^\theta, u^\theta) \quad := \sum_{(A^j,G^j)\in\mathcal{B}^{vw}} \theta_j^{vw} (x^j, u^j) \tag{62}$$

(Note that if θ^{vw} is chosen to be a basic solution, then at most $n+1$ components θ_j^{vw} will be positive, and thus (x^θ, u^θ) will be a convex combination of at most $n+1$ solutions (x^j, u^j).) Recall from Section 6.1 that $\mathcal{Y} \subset \mathcal{X} \times \mathcal{U}$ is convex, and that $(x^j, u^j) \in \mathcal{Y}$ for all j, and thus $(x^\theta, u^\theta) \in \mathcal{Y}$. It follows from the second constraint of the dual linear program (61) and the linearity of G in (x, u) that

$$\begin{aligned} b - RQx^\theta(T) &= \sum_{(A^j, G^j) \in \mathcal{B}^{vw}} \theta_j^{vw} \left(b - RQx^j(T)\right) \\ &= \sum_{(A^j, G^j) \in \mathcal{B}^{vw}} \theta_j^{vw} G^j \quad \geq \quad 0 \end{aligned}$$

and thus (x^θ, u^θ) is feasible for the primal problem. (Concavity of G in (x, u) would have been sufficient.) Furthermore, from the linearity of g and h in (x, u),

$$\begin{aligned} \int_0^T g(x^\theta(t), u^\theta(t), t)\, dt + h(x^\theta(T)) &= \sum_{(A^j, G^j) \in \mathcal{B}^{vw}} \theta_j^{vw} \left[\int_0^T g(x^{\lambda^j}(t), u^{\lambda^j}(t), t)\, dt + h(x^{\lambda^j}(T))\right] \\ &= z^{vw} \quad > \quad L^*(\lambda^{vw}) - \varepsilon_d \end{aligned}$$

and thus the objective value of (x^θ, u^θ) is within ε_d of the primal optimal value. (Again, concavity of g and h in (x, u) would have been sufficient.) Then the algorithm stops with primal solution (x^θ, u^θ) and dual solution λ^{vw}.

To summarize, the cutting-plane algorithm for the Lagrangean dual problem is stated:

Cutting-Plane Algorithm with Trust Region for Lagrangean Dual Problem

Step 0 (Initialization): Choose algorithm parameters: move test parameter $\eta_m \in (0,1)$, contraction test parameter $\eta_c \in (0, \eta_m)$, expansion test parameter $\eta_e \in (\eta_m, 1)$, contraction multiplier $m_c \in (0,1)$, expansion multiplier $m_e > 1$, minimum trust region size parameter $\underline{\Delta} > 0$, constraint elimination test parameters $\xi > 0$, $n_1, n_2, n_3, n_4 \geq 0$, primal optimality gap tolerance $\varepsilon_p > 0$, dual optimality gap tolerance $\varepsilon_d > 0$, and primal infeasibility tolerance $\varepsilon_f > 0$. Choose an initial solution $\lambda^{1,0} \geq 0$, and an initial trust region size parameter $\Delta^{1,1} > 0$. Solve the Lagrangean problem with dual multipliers $\lambda^{1,0}$, producing objective value $L^*(\lambda^{1,0})$, subgradient $G^{1,0} \in \partial L^*(\lambda^{1,0})$, and $A^{1,0} := L^*(\lambda^{1,0}) - (G^{1,0})^T \lambda^{1,0}$. Let $\mathcal{B}^{1,1} := \{(A^{1,0}, G^{1,0})\}$, let $n_4^{1,0} = 0$, and let $\hat{L}^{1,1}(\lambda) := A^{1,0} + (G^{1,0})^T \lambda$. Set $v = 1$, $w = 1$, and $d = 1$.

Step 1 (Approximating Problem): Solve the approximating problem (52), producing optimal solution (z^{vw}, λ^{vw}).

Step 2 (Lagrangean Problem): Solve the Lagrangean problem (15)–(17) with dual multipliers λ^{vw}, producing solution $(x^{\lambda^{vw}}, u^{\lambda^{vw}})$, objective value $L^*(\lambda^{vw}) = L(x^{\lambda^{vw}}, u^{\lambda^{vw}}, \lambda^{vw})$, subgradient $G^{vw} \in \partial L^*(\lambda^{vw})$, and $A^{vw} := L^*(\lambda^{vw}) - (G^{vw})^T \lambda^{vw} = F(x^{\lambda^{vw}}, u^{\lambda^{vw}})$.

Step 3 (Primal Stopping Test): If $F(x^{\lambda^{vw}}, u^{\lambda^{vw}}) > L^*(\lambda^{vw}) - \varepsilon_p$ and $G(x^{\lambda^{vw}}, u^{\lambda^{vw}})_r > -\varepsilon_f$ for each $r \in \mathcal{R}$, then stop with primal solution $(x^{\lambda^{vw}}, u^{\lambda^{vw}})$ and dual solution λ^{vw}.

Step 4 (Dual Stopping Test): If $L^*(\lambda^{vw}) < z^{vw} + \varepsilon_d$ and none of the trust region constraints is active at (z^{vw}, λ^{vw}) (except if $\lambda_r^{vw} = \lambda_r^{v0} - \Delta_r^{vw} = 0$), then let θ^{vw} denote an optimal solution of (61), and let $(x^\theta, u^\theta) := \sum_{(A^j, G^j) \in \mathcal{B}^{vw}} \theta_j^{vw} (x^j, u^j)$, and stop with ε_d-optimal primal solution (x^θ, u^θ) and dual solution λ^{vw}. If $L^*(\lambda^{vw}) < z^{vw} + \varepsilon_d$ and one or more of the trust region constraints are active at (z^{vw}, λ^{vw}) (besides $\lambda_r^{vw} = \lambda_r^{v0} - \Delta_r^{vw} = 0$), then the trust region is expanded: $\Delta^{v,w+1} := m_e \Delta^{vw}$, and go to step 6.

Step 5 (Expansion or Contraction of Trust Region): If $L^*(\lambda^{vw}) > L^*(\lambda^{v0}) - \eta_c \left(L^*(\lambda^{v0}) - z^{vw}\right)$, then the trust region is contracted: $\Delta^{v,w+1} := \max\{m_c\Delta^{vw}, \underline{\Delta}\}$. If $L^*(\lambda^{vw}) < L^*(\lambda^{v0}) - \eta_e \left(L^*(\lambda^{v0}) - z^{vw}\right)$, then the trust region is expanded: $\Delta^{v,w+1} := m_e\Delta^{vw}$. Otherwise, set $\Delta^{v,w+1} := \Delta^{vw}$.

Step 6 (Serious Step or Null Step): If $L^*(\lambda^{vw}) > L^*(\lambda^{v0}) - \eta_m \left(L^*(\lambda^{v0}) - z^{vw}\right)$, then go to step 9. Move the stability center to λ^{vw} and begin a new major iteration; that is, $\lambda^{v+1,0} := \lambda^{vw}$, $L^*(\lambda^{v+1,0}) = L^*(\lambda^{vw})$, $A^{v+1,0} := A^{vw}$, $G^{v+1,0} := G^{vw}$, and $\Delta^{v+1,1} := \Delta^{v,w+1}$.

Step 7 (Trim Bundle): If $|\mathcal{B}^{vw}| < n_1$, or $d < n_2$, then go to step 8. For each $j \in \mathcal{B}^{vw}$ such that $z^{vw} - (A^j + G^{jT}\lambda^{vw}) > \xi(L^*(\lambda^{v0}) - z^{vw})$, set $n_4^j \leftarrow n_4^j + 1$. Let $\mathcal{B}_-^{vw} := \{j \in \mathcal{B}^{vw} : z^{vw} - (A^j + G^{jT}\lambda^{vw}) > \xi(L^*(\lambda^{v0}) - z^{vw})$ and $n_4^j > n_4\}$. If $|\mathcal{B}^{vw}| - |\mathcal{B}_-^{vw}| \geq n_3$, then let $\mathcal{B}^{v+1,0} := \mathcal{B}^{vw} \setminus \mathcal{B}_-^{vw}$; otherwise, the $|\mathcal{B}^{vw}| - n_3$ constraints in $\mathcal{B}_-^{vw}$ with the greatest slack values are eliminated from $\mathcal{B}^{vw}$ to give $\mathcal{B}^{v+1,0}$. Set $d = 0$.

Step 8 (Update): Set $v \leftarrow v + 1$ and $w = 0$.

Step 9 (Continue): Let $\mathcal{B}^{v,w+1} := \mathcal{B}^{vw} \cup \{(A^{vw}, G^{vw})\}$, let $n_4^{vw} = 0$, and let $\hat{L}^{v,w+1}(\lambda) := \max\{A^j + G^{jT}\lambda : (A^j, G^j) \in \mathcal{B}^{v,w+1}\}$. Set $w \leftarrow w + 1$ and $d \leftarrow d + 1$. Go to step 1.

7 Numerical Results

In this section some numerical results are briefly described. The purpose of the numerical tests was to investigate the ability of the method to solve problems of different sizes.

Revenue management problems were randomly generated for hub-and-spoke airline networks. For each location, a weight was generated randomly, and the demand between two locations was generated randomly and proportional to the product of the weights of the two locations (a gravity model). The capacities of the flights between each location and the hub were also generated randomly and proportional to the weight of the location. For each origin-destination pair there were a number of prices and a number of customer classes interested in travelling from the origin to the destination. The arrival and cancellation rates of all customer classes increased over time, but more so for the higher customer classes. Also, for all customer classes the probability of a customer purchasing the product decreased as a function of price, but less so for the higher customer classes.

Table 1 shows statistics of the problems solved and algorithm performance. The following values were used for the algorithm parameters: move test parameter $\eta_m = 0.1$, contraction test parameter $\eta_c = 0.001$, expansion test parameter $\eta_e = 0.7$, contraction multiplier $m_c = 0.67$, expansion multiplier $m_e = 1.5$, minimum trust region size parameter $\underline{\Delta} = 0.1$, constraint elimination test parameters $\xi = 0.1$, $n_1 = |\mathcal{R}| + 50$, $n_2 = 50$, $n_3 = |\mathcal{R}| + 10$, $n_4 = 2$, primal optimality gap tolerance $\varepsilon_p = 0.001$, dual optimality gap tolerance $\varepsilon_d = 0.001$, and primal infeasibility tolerance $\varepsilon_f = 0.5$. The values of the algorithm parameters were not tuned. Computations were performed on a PC with a 133MHz processor.

It appears that the number of major iterations does not depend on the size of the problem. However, the number of minor iterations, and the computational effort per iteration, grow with the size of the problem. Because minor iterations are required to construct a sufficiently accurate approximation to the dual objective function L^* in a neighborhood of the current stability center, it seems that it becomes harder to construct such an approximation as the dimension of the domain of L^* increases. Improvements in the efficiency of constructing accurate approximations to convex functions will contribute much to the ability to solve large dynamic pricing problems.

Instance Number	Number of Product-Price Pairs $\vert\mathcal{I}\vert$	Number of Customer Classes $\vert\mathcal{K}\vert$	Number of Resources $\vert\mathcal{R}\vert$	Number of Major Iterations	Number of Minor Iterations	CPU time [seconds]
1	550	220	20	57	168	3
2	2100	840	40	44	256	20
3	4650	1860	60	44	408	76
4	8200	3280	80	58	1075	408
5	12750	5100	100	41	641	343
6	18300	7320	120	80	2293	2712
7	24850	9940	140	53	2151	3753
8	32400	12960	160	66	2568	5970
9	40950	16380	180	53	2926	9188
10	50500	20200	200	56	5529	55562

Table 1: Computational results for randomly generated dynamic pricing problems for hub-and-spoke airline networks.

References

Bertsekas, D. P. 1996. *Constrained Optimization and Lagrange Multiplier Methods.* Athena Scientific, Belmont, MA.

Bonnans, J. F. and Shapiro, A. 2000. *Perturbation Analysis of Optimization Problems.* Springer-Verlag, New York.

Fleming, W. H. and Rishel, R. W. 1975. *Deterministic and Stochastic Optimal Control.* Springer-Verlag, New York, NY.

Geraghty, M. K. and Johnson, E. 1997. Revenue Management Saves National Car Rental. *Interfaces*, **27**, 107–127.

Hiriart-Urruty, J. B. and Lemaréchal, C. 1993. *Convex Analysis and Minimization Algorithms.* Springer-Verlag, Berlin, Germany.

Kelley, J. E. 1960. The Cutting Plane Method for Solving Convex Programs. *SIAM Journal*, **8**, 703–712.

Kiwiel, K. C. 1990. Proximity Control in Bundle Methods for Convex Nondifferentiable Minimization. *Mathematical Programming*, **46**, 105–122.

Linderoth, J. and Wright, S. 2001. Decomposition Algorithms for Stochastic Programming on a Computational Grid, *Technical Report ANL/MCS-P875-0401*, Mathematics and Computer Science Division, Argonne National Laboratory.

McGill, J. I. and Van Ryzin, G. J. 1999. Revenue Management: Research Overview and Prospects. *Transportation Science*, **33**, 233–256.

Moré, J. J. 1997. Recent Developments in Algorithms and Software for Trust Region Methods. In *Mathematical Programming, the State of the Art.* A. Bachem, M. Grötschel and B. Korte (editors). Springer Verlag, Berlin, Germany, 258–287.

Rockafellar, R. T. 1970. *Convex Analysis.* Princeton University Press, Princeton, NJ.

Ruszczyński, A. 1986. A Regularized Decomposition for Minimizing a Sum of Polyhedral Functions. *Mathematical Programming*, **35**, 309–333.

Smith, B. C., Leimkuhler, J. F. and Darrow, R. M. 1992. Yield Management at American Airlines. *Interfaces*, **22**, 8–31.