Revenue Management

Tool: Optimization (prescriptive analytics).

The Analytics Edge: There are several attributes of revenue management systems that make them widely used:

1. They find ways to increase revenue without necessarily changing products;
2. Prices are better aligned to what customers are willing to pay;
3. Computational power allows to solve large-scale optimization models.

Overview

In the late 1970s, deregulation in the airlines industries provided opportunities (and challenges) for many of the airline carriers to increase profit and stay ahead of the competition. Charter airlines started selling much cheaper tickets than American Airlines. Managing the seats sold effectively was realized to be important, so revenue management techniques were used to increase revenue. Today, revenue management and the use of analytics for this purpose is affecting many industries, including hotels and online retailers. We will use the example of American Airlines to see how techniques for revenue management can be used to generate revenue. We will also discuss extensions of these ideas to pricing decisions in revenue management. The key tool that is used in these applications is optimization (prescriptive analytics).

Revenue management increases company profits by using analytics *to sell the right product to the right customer at the right time for the very right price.* After deregulation in the airlines industry in the U.S., companies had the flexibility to explore different pricing and routing options and the best mix of passengers to maximize revenue. Prior to deregulation, airlines could only fly certain routes, and fares were determines by the federal Civil Aeronautics Board based on mileage and costs. Among some of the features of the airlines industry that made it suitable for revenue management are: 1) fixed capacity (i.e., number of seats), 2) high fixed costs but low variable costs, 3) variety of customer types (e.g., price sensitive, luxury travellers), and 4) ability to sell tickets to customers without seeing what other customers paid. American Airlines is considered to be one of the first airlines that successfully implemented a revenue management system.

The scale of this problem can be huge, since reservations are made often months in advance to a day before and there are multiple departures for an airline carrier; so, the system needs to be computerized just to handle the volume of business. Here, we discuss some of these challenges taken from the influential paper "Yield management at American Airlines" (Smith et al., 1992). One of the problems described in the paper is the discount allocation, which is the problem of determining the number of discount fares to offer for the flights. The tradeoff is between offering discounts so as to fill all seats or not offering discounts, but with the risk of having some empty seats (which can be taken by late-arriving high-revenue passengers).

Discount allocation

One important hub was Dallas/Fort Worth International Airport (DFW), where passengers used to travel from Portland (PDX) and Los Angeles (LAX) or connect to travel to New York (JFK) and Miami (MIA) (see Figure 0.1). Thus, there are four types of passenger routes in this network. Furthermore, there are two classes of passengers, one who books in advance and wants to pay less and a second class who books late but is willing to pay a lot more. This inspired the airlines two offer two products: 1) more expensive tickets that can be purchased anytime, with no restrictions, and fully refundable, and 2) cheaper tickets, to be bought at least 3 weeks in advance, with penalties for changes.

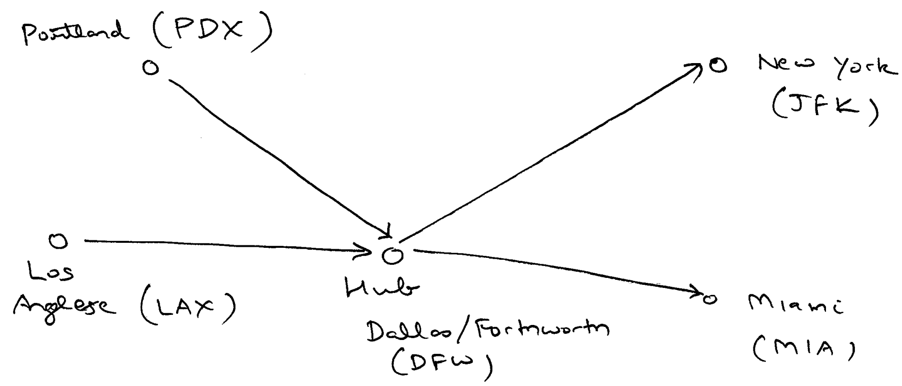

Figure 0.1: Four spoke network centred in Dallas/Fort Worth International Airport (DFW).

Problem formulation

Assume the marketing department has already decided the prices, and we are interested in deciding how many seats to reserve in each fare class and itinerary. For simplicity, we consider only three flights: LAX $\rightarrow$ DFW, LAX $\rightarrow$ JFK, and DFW $\rightarrow$ JFK. Only two legs are used here in terms of flights. The data regarding flight capacity are given in Table 0.1, while the data on revenue and demand are reported in Table 0.2.

Table 0.1: Capacity of LAX $\rightarrow$ DFW and DFW $\rightarrow$ JFK flights.

Flight	Capacity
LAX $\rightarrow$ DFW	300
DFW $\rightarrow$ JFK	200

The problem of maximizing revenue can be formulated through the following Linear Program (LP):

Table 0.2: Revenue and demand for each fare of LAX $\rightarrow$ DFW, LAX $\rightarrow$ JFK, and DFW $\rightarrow$ JFK flights. The last column reports the decision variables for each leg.

Flight	Fare	Revenue ($)	Demand	Decision variable
	Regular	100	20	x_1
LAX $\rightarrow$ DFW	Discount	90	40	x_2
	Saver	80	60	x_3
	Regular	215	80	y_1
LAX $\rightarrow$ JFK	Discount	185	60	y_2
	Saver	145	70	y_3
	Regular	140	20	z_1
DFW $\rightarrow$ JFK	Discount	120	20	z_2
	Saver	100	30	z_3

$$
\begin{array}{ll}
\max & 100x_1 + 90x_2 + 80x_3 + 215y_1 + 185y_2 + 145y_3 + 140z_1 + 120z_2 + 100z_3 \\
\\
\text{s.t.} & x_1 + x_2 + x_3 + y_1 + y_2 + y_3 \leq 300 \\
& y_1 + y_2 + y_3 + z_1 + z_2 + z_3 \leq 200 \\
& 0 \leq x_1 \leq 20 \\
& 0 \leq x_2 \leq 40 \\
& 0 \leq x_3 \leq 60 \\
& 0 \leq y_1 \leq 80 \\
& 0 \leq y_2 \leq 60 \\
& 0 \leq y_3 \leq 70 \\
& 0 \leq z_1 \leq 20 \\
& 0 \leq z_2 \leq 20 \\
& 0 \leq z_3 \leq 30
\end{array}
$$

The optimal solution is: $x_1^* = 20$, $x_2^* = 40$, $x_3^* = 60$, $y_1^* = 80$, $y_2^* = 60$, $y_3^* = 40$, $z_1^* = 20$, $z_2^* = 0$, $z_3^* = 0$, with a total revenue of $ 47,300.

What would a greedy method do?

Give priority to the highest revenue customers and greedily allocate them. In the problem describe above, we could set $y_1 = 80$, $y_2 = 60$, and $y_3 = 60$ (since the capacity of the DFW $\rightarrow$ JFK is only 200). We would still have 100 seats on the LAX $\rightarrow$ DFW flight, so we could set $x_1 = 20$, $x_2 = 40$, and $x_3 = 40$. The total revenue in this case would be $ 46,000, which is about 3% less than the optimal revenue. Aggregating such loss over all flights and all fare classes would lead to a significant amount of money.

Practical implementation

There are some potential challenges with this model:

1. As soon as regular priced ticket are sold, only discount tickets are available though it is possible that some customers might still be willing to buy at a regular price. Airlines use "nesting control", where seats reserved for discount classes are made available to more expensive classes. In other words,

companies make all tickets available to regular customers, discount and super saver tickets to discount customers and only super saver tickets are available to super saver customers.

2. The advantage of an LP model is that it provides a shadow price for each constraints (dual variable). This reflects the revenue value of the last seat on each leg of the flight for the capacity constraints. These prices—also referred to as bid prices—are used in practice to decide whether or not an additional seat should be offered for a particular flight and fare class. For example, if leg 1 dual variable $= p_1$ and leg 2 dual variable $= p_2$, then, for any new fare request on the leg 1 + leg 2 itinerary, if fare willing to pay $\geq p_1 + p_2$, the request is accepted.

3. The data are often changing with time, so dynamic models with randomness in demand must be incorporated in such implementation.

Prescribing the right price

One of the key aspects of revenue management is pricing. Consider the plot illustrated in Figure 0.2, which illustrates the demand curve for a single product (with price reported on the y axis). The shaded area illustrates the revenue lost for setting one price P^* for all customers (to sell the inventory Q^*).

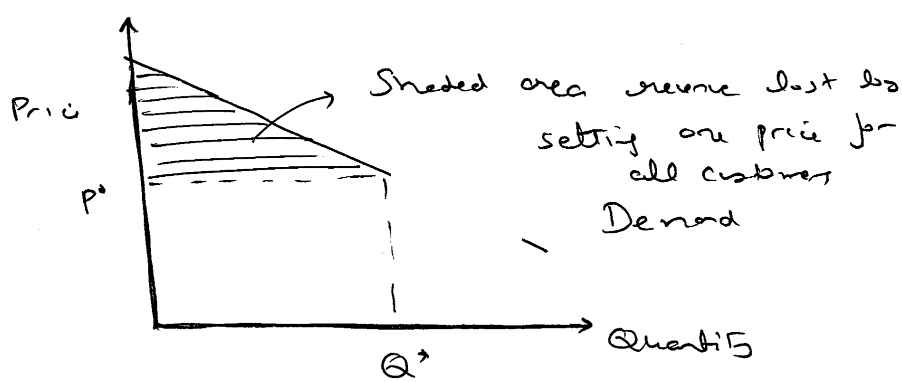

Figure 0.2: Example of demand curve for a single product.

Suppose we can set three prices for the same product. Then, we can reduce the amount of revenue lost by setting a single price. This concept is illustrated in Figure 0.3, which considers the case of customers willing to pay more than P_1 but less P_2, and customers willing to pay more than P_2 but less than P_3.

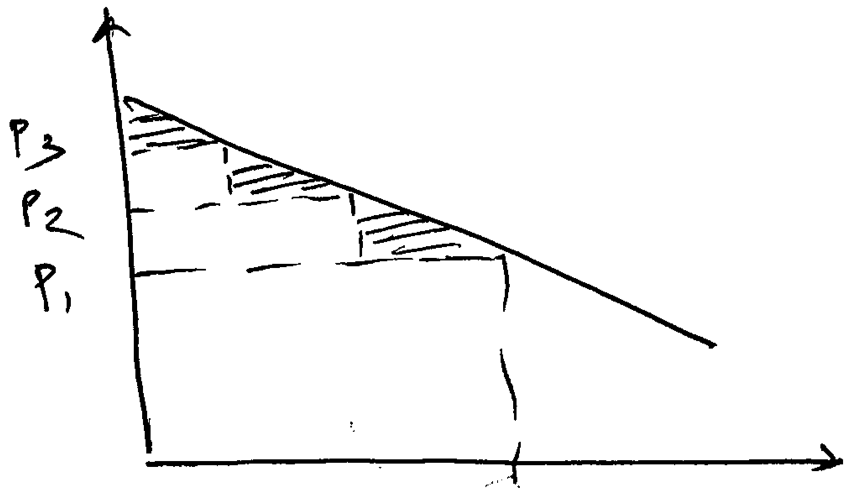

Figure 0.3: Lost revenue with three prices for the same product.

Pricing products with a Multinomial Logit model

Discrete choice models are also used to determine prices of products since they help describe consumers' demand. Consider a set of products denoted as $\{1, \ldots, n\}$ and an outside option denoted by 0. The goal of the decision-maker is to price products $1, \ldots, n$. These products are substitutes and this is accounted for by choice models:

$$\max_{p_1 \geq 0, \ldots, p_n \geq 0} \sum_{i=1}^{n} (\underbrace{p_i}_{\text{price}} - \underbrace{c_i}_{\text{cost}}) \underbrace{\mathbb{P}(\text{choosing product } i \text{ given prices } \bar{P})}_{\substack{\text{Obtained from} \\ \text{logit model}}},$$

where

$$\mathbb{P}(\text{choosing product } i \text{ given prices } \bar{P}) = \frac{e^{v_i - p_i}}{\sum_{j=0}^{n} e^{v_j - p_j}}.$$

So, the optimization problem can be rewritten as:

$$\max_{p_1 \geq 0, \ldots, p_n \geq 0} \sum_{i=1}^{n} \frac{(p_i - c_i) e^{v_i - p_i}}{\sum_{j=0}^{n} e^{v_j - p_j}},$$

where $v_o = p_o = 0$. This is not an easy objective function to deal with. Alternatively, we can solve the problem using choice probability variables. Let:

$$q_i = \frac{e^{v_i - p_i}}{\sum_{j=0}^{n} e^{v_j - p_j}} \ \forall j = 1, \ldots, n,$$

and

$$q_o = \frac{1}{\sum_{j=0}^{n} e^{v_j - p_j}},$$

where $v_o = p_o = 0$. Then:

$$v_i - p_i = \log(q_i) - \log(q_o) \tag{0.1}$$
$$p_i = v_i - \log(q_i) + \log(q_o). \tag{0.2}$$

So, we rewrite the problem as follows:

$$\begin{array}{ll} \max & \sum_{i=1}^{n}(v_i - \log(q_i) + \log(q_o) - c_i)q_i \\ & \\ \text{s.t.} & \sum_{i=0}^{n} q_i = 1 \\ & q_i \geq 0 \ \forall i = 0, \ldots, n. \end{array}$$

Note that the problem simplifies to:

$$\begin{array}{ll} \max & \sum_{i=1}^{n}(v_i - c_i - \log(q_i))q_i + (1 - q_o)\log(q_o) \\ & \\ \text{s.t.} & \sum_{i=0}^{n} q_i = 1 \\ & q_i \geq 0 \ \forall i = 0, \ldots, n. \end{array}$$

Note that $-q\log(q)$ is concave in q and that $(1 - q_o)\log(q_o)$ is concave in q_o, so the problem can be solved efficiently using convex optimization:

$$\frac{d}{dq}(-q\log(q)) = -1 - \log(q),$$

$$\frac{d^2}{dq^2}(-q\log(q)) = -\frac{1}{q} \leq 0.$$

$$\frac{d}{dq}((1-q)\log(q)) = \frac{1-q}{q} - \log(q),$$

$$\frac{d^2}{dq^2}((1-q)\log(q)) = \frac{q(-1)-(1-q)}{q^2} - \frac{1}{q} = -\frac{1}{q^2} - \frac{1}{q} \leq 0.$$

In practice, how might companies choose the value v? For example, online retailers can easily change the prices and observe the market share of the products. By doing this many times, they can try to estimate v (as illustrated in Figure 0.4).

We can then do a simple regression to estimate the evaluation of products. Also in some cases, companies such as GM use complex simulators to predict market shares. How to optimize prices is a problem to solve!

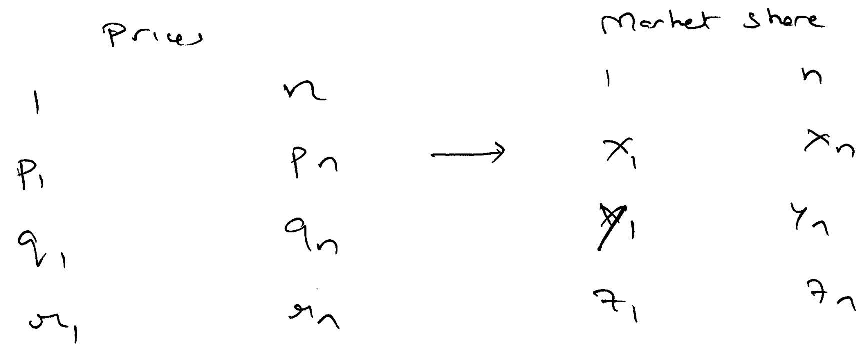

Figure 0.4: Setting the value v by trial-and-error.

References

Smith, B. C., J. F. Leimkuhler, and R. M. Darrow (1992). Yield management at american airlines. *interfaces 22*(1), 8–31.