Behaviour-driven dynamic pricing modelling via hidden Markov model

Qinfu Qiu and Xiong Chen*

Electronic Engineering Department,
Fudan University,
Shanghai, China
Email: qiuqinfu@gmail.com
Email: chenxiong@fudan.edu.cn
*Corresponding author

Abstract: The dynamic pricing strategy in airline tickets has gained a lot of concern from both air companies and customers. As it has been proved that putting out discount in airline tickets in sometime may increase a company’s total revenue. In this paper, after making analysis in both sides of airline and passengers’ behaviour, we found that all the different choices made by passengers, whether purchase or keep waiting, come from an invisible logical chain which contains several key driving elements. Thus, we implement hidden Markov model here, trying to model a new pricing mechanism to raise the revenue. The simulation verified this model’s practicability.

Keywords: dynamic pricing strategy; revenue; invisible logical chain; hidden Markov model.


Biographical notes: Qinfu Qiu is a graduate student in the School of Information Science and Technology at Fudan University, Shanghai, China. He majors in Circuit and System, and received a Bachelor’s degree in Electronic Information Science and Technology.

Xiong Chen is a Professor in School of Information Science and Technology, Fudan University, Shanghai, China, since 2001. From 1998 to 2001 he was a Postdoctoral Fellow at School of Information Science and Technology, Tongji University, Shanghai, China. He received his PhD in Automatic Control Theory and Application from Northeastern University, Shenyang, China, in 1998. His research interests include wireless sensor network, intelligent control theory and system, and data engineering.

1 Introduction

In the year 1980s, the US airline industry first proposed a concept of revenue management. This dynamic pricing model became a popular issue. Later, researchers presented large numbers of dynamic pricing model (Li, 2001; Anjos et al., 2005).

For air ticket pricing, once the price changed, it would have an impact on overall earnings. And according to the basic economical common sense, different price will cause different market demand, enhancing the affection on the airline company earnings (Vowles, 2000). And many customers will have a rational expectation on the price in the future. Thus, they choose the best time to buy the tickets (Su, 2007).

From the airline perspective view, if the fare is too high, it will neither bring great benefits to the airline nor increase the market demands (Kang et al., 2016). It may even make the revenue decline. From the customer perspective view, the uncertainty of purchasing reduces the company’s expected return. Such as price-sensitive passengers may be willing to sacrifice their time flexibility but buying tickets at lower prices. It is exactly this feature, makes the implementation of an effective revenue management strategy in customer behaviour particularly important (Yin et al., 2009). Searching the best ticket pricing strategy is more complex.

In previous studies, some focused on the interest conflict between the airlines and passengers and found that game theory is a pretty good solution (Zhang, 2006). There are other kinds of strategies which totally focus on passengers’ decision. For example, some researches focus on the airline seat inventory control. And they made recommendation about when to release seats to market. Several researches pointed out that if the customers have rational expectations of future price, then the marginal cost will equal to its price. Thus, the resulting profits will be zero (Coase, 1972). Some pointed out that the best pricing process can bring about a 20% increase in profits (Besanko and Winston, 1990). One paper even pointed out the price discount shall be in a certain threshold, and only discounted for one time during the whole process seems to be the best
choice (Elmaghraby et al., 2008). In addition, more detailed analysis focused on airline pricing competition (Mo, 2011), the impact of the current market environment for airlines (Levin and McGill, 2009).

To sum up, there is a strong correlation between the individual decisions and ticket pricing. This association should focus on the price-sensitive and the concern about the ticket quantity. The former factor may encourage customers to consider sidelines or directly abandon the tickets, while the later was able to encourage customers to consider purchase. The customers’ choice will focus on purchasing or not purchasing. But its invisible main factor shall focus on price changing and whether they have such opportunity to buy ticket in such price. These finally helped customers to make choices. So, we believe that the introduction of hidden Markov model can help to analyse in airline ticket pricing to seek maximum profits. How can we finish that will be told in the following sections.

Section 2 of the article will focus primarily on several hidden Markov algorithms. And Section 3 will be introduction of modelling. Section 4 will analyse our results. Section 5 is the conclusion.

2 Hidden Markov model

The hidden Markov model is a valuable statistical tool, which we always use it for modelling for a generative sequence. As it can be characterised by an underlying process, generate an observation sequence. Figure 1 illustrate a hidden Markov chain.

Figure 1 Hidden Markov chain

The hidden Markov model has been applied in many areas. At first, in the early 20th century, Markov (1913) gave his name to this mathematical theory as Markov processes. Later, in year 1960, Baum et al. (1970) developed the theory of hidden Markov model.

In basic Markov models, such as a Markov chain, all of its states are directly visible to the observer. Thus, the only parameter is the state transition probabilities.

In a hidden Markov model, the state is not directly visible to the observer, but the output, also called observation sequence. It is dependent on the state which is visible. Each state has a probability distribution. Therefore the sequence of tokens generated by a hidden Markov model provides some information about the sequence of states. Here, ‘hidden’ means the state sequence through which the model passes, not to the parameters of the model. The model is still referred to as a ‘hidden’ Markov model even if these parameters are known already.

A hidden Markov model can be considered as a generalisation of a mixture model. The hidden variables control the mixture component to be selected for each observation, are related through a Markov process rather than independent of each other. Recently, hidden Markov models have been generalised to pairwise Markov models and triplet Markov models which allow consideration of more complex data structures (Pieczynski and Pieczynski, 2007; Boudaren et al., 202) and the modelling of no stationary data (Lanchantin and Pieczynski, 2005).

2.1 Basic definition

The hidden Markov model’s formal definition is as follows:

$$\lambda = (A, B, \pi)$$

Here, $A$ is called a transition array, showing the probability of changing from $j$ to the following state $i$:

$$A = [a_{ij}], a_{ij} = P(s_j | s_i)$$

$B$ is called the observation array, showing the probability of observation $k$ from state $j$:

$$B = [b_i(k)], b_i(k) = P(v_k | s_i)$$

We would like to combine all $s$ and $v$ to be observation alphabet and state alphabet:

$$S = \{s_1, s_2, ..., s_n\}$$

$$V = \{v_1, v_2, ..., v_m\}$$

$\pi$ is the initial probability:

$$\pi_i = P(s_i)$$

There still exist two coefficients need to be defined:

$$Q = q_1, q_2, ..., q_T$$

$$O = o_1, o_2, ..., o_T$$

$Q$ is a state sequence of length $T$, $O$ is the corresponding observation. These two coefficients is just a symbol.

Finally, two assumptions shall be made here. One is called the Markov assumption, which means the current state is dependent only on the previous state. The other is called independence assumption, which means the observation at any time is dependent only on the current state.

2.2 Three main implementations

According to the basic definition, hidden Markov model is mainly used to solve three kinds of question, evaluation, decoding, and learning.
2.2.1 Evaluation

If we are given a sequence of observation, we need to compute the probability of sequence of observation. This issue can be viewed as evaluation, knowing how well a model predicts a given sequence of observation.

Thus, the probability of $O$ for $Q$ is:

$$P(O | Q, \lambda) = \prod_{t=1}^{T} P(o_t | q_t, \lambda)$$  \hspace{1cm} (9)

And it state sequence probability is:

$$P(Q | \lambda) = \pi_0 \prod_{t=1}^{T} a_{q_t-1, q_t}$$  \hspace{1cm} (10)

So the probability of the given observation sequence can be finally got:

$$P(O | \lambda) = \sum_{Q} P(O | Q, \lambda) P(Q | \lambda)$$

$$= \sum_{q_1 \rightarrow q_T} \pi_{q_0} b_{q_0}(o_1) a_{q_1,q_2} b_{q_2}(o_2) \cdots a_{q_{T-1}, q_T} b_{q_T}(o_T)$$  \hspace{1cm} (11)

2.2.2 Decoding

If we are given a sequence of observation, how to know which hidden state sequence is most likely to produce this given observation? So here we need decoding.

Usually, we like to use Viterbi algorithm to solve this problem. And we need to define:

$$\varphi(i) = \max_{q_{i-1} \rightarrow q_i} P(q_i | q_{i-1} = s_i)$$  \hspace{1cm} (12)

So that we can find the most probable path for the given observation sequence. And Viterbi algorithm can be broken down into several steps.

$$\varphi(i) = \pi_i b_i(o_i), 1 \leq i \leq N$$  \hspace{1cm} (13)

$$\varphi(j) = \max_{1 \leq i \leq N} \left[ \varphi_{i-1}(i) a_{i,j} b_j(o_j) \right]$$  \hspace{1cm} (14)

$$\Psi_t(j) = \arg \max_{1 \leq i \leq N} \left[ \varphi_{i-1}(i) a_{i,j} \right]$$  \hspace{1cm} (15)

So we got the final path and its probability:

$$P = \max_{1 \leq i \leq N} \left[ \varphi_T(i) \right], q_T = \arg \max_{1 \leq i \leq N} \left[ \varphi_T(i) \right]$$  \hspace{1cm} (16)

Here, one thing needs to be mentioned. In evaluation process, it contains two basic algorithms which called forward-algorithm and backward-algorithm. While somehow, the formation of forward-algorithm equals to that of Viterbi algorithm. But in the first step shown above, we use maximising is in forward-algorithm, rather than summing and storing the state which was chosen as the maximum for a back pointer’s use.

2.2.3 Learning

If we are given a real example, usually we need to estimate the model’s parameters at first. And there are two standard ways to get the parameters: supervised training and unsupervised training. If inputs and outputs of a Markov process are provided, we can perform supervised training. If only the inputs are provided, thus we need to use unsupervised training.

In order to get $\lambda$, we are supposed to collect large quantity of training examples to make an accurate answer.

3 Modelling

3.1 Fundamental analysis

When we are making analysis on the pricing issue, the key players are airline companies and passengers. Although airlines may present different prices at different time, all prices will be limited in an appropriate benchmark. Therefore, we will set the price $P_i$ ($i = 1, 2, 3 \ldots T$) as the price of each period. But here is a premise that the airline will make stepped pricing gradually increasing ticket prices. In fact, this rule is the core foundation of all pricing strategies.

Another problem is that based on current experience, airline fares and departure time has an equivalent increase in their relationships. That is, when the departure time comes more recent, the fares will be higher.
In addition, from the passengers’ view, they choose airline but no other kinds of transportation is just because the total cost of air ticket is lower than others. Thus, the total cost mentioned here equals the combination of fare cost and time cost. We would like to separate the passengers to two kinds. One does not care about time but care about the price. The other one is less price-sensitive, but they care about time. Thus, the second kind of people may be more willing to accept a higher price ticket.

3.2 Definition of coefficients and algorithm

$S$ means the quantity of the seats in the airplane. $T$ means the total sales cycle. $\eta$ means a customer possibility of purchasing in the first day. $\beta$ means transaction arrow between observation sequence and underling sequence. We also need to assume that there will be $r$ persons book the ticket in the open day.

Thus, to facilitate the analysis, we use a discrete time model. The total sales cycle will be separated into days. And at one time, there will be one customer buy one ticket. Here are two core visible causes that made people to decide whether to buy an air ticket: the increasing price and the less opportunity of buying tickets. Here, we put purchase or not purchase into invisible elements. And we predict the number of purchasing when we start arranging the price.

As purchasing and not purchasing become the invisible elements, there transition possibly will always change. And we would like to use pricing change (increased) and stay still into two observation elements.

Here, the total increase ratio equals to dividing full price by initial price. And this ratio will be discrete during the selling process. Once the model detects that even we increase the price, the passenger will still buy the ticket, and it will raise the price in this round. According to the situation, we prefer Viterbi algorithm as our base model of pricing. And the model shall be with discrete data so that all the tickets selling process can be divided in to several rounds (days).

The aim of Viterbi algorithm is to maximise an objective function $G(s)$, where $s$ is a state sequence and $G(s)$ has a special property. The property of $G(s)$:

$$G(s) = g_1(s_1) + g_2(s_2, s_1) + ... + g_T(s_T, s_{T-1})$$

Every state sequence $s$ corresponds to a path from $t = 1$ to $t = T$. And we define that $g(a, b)$ on the link as the weight from state $a$ at $t = 1$ to state $b$ at $t$.

$g_k(k)$ is the weight of state $k$ at the starting node. And $G(s)$ is the sum of the weights on the links in path $s$. Our job is to find of the best possible sequence for the total revenue of selling tickets.

Because the price always changes, all the possibility of these arrows will not stay constant but according to a linear increasing or decreasing process.

![Figure 4](image)

The initial arrow to purchase and non-purchase in day 1 will be:

$$\pi_{Pur} = r / S; \pi_{Non} = (M - r) / S$$

The transaction array and purchase to itself, non-purchase to itself will be as follows:

$$A_{P-N} = \eta (1 - t / T)$$
$$A_{N-P} = t / T$$
$$A_{P-P} = \eta t / T$$
$$A_{N-N} = 1 - t / T$$

For the first day, we would like to calculate the following equation to detect whether there will be someone who will buy the ticket:

$$\max \left\{ \begin{array}{ll}
(M - r) \cdot \beta _{N-1} \cdot \beta _{B-P}
\end{array} \right\}$$

$$\max \left\{ \begin{array}{ll}
(M - r) \cdot \beta _{P-1} \cdot \beta _{B-N}
\end{array} \right\}$$

Here, we may say that there surely will not have change in the fare in day 2, so the choice will be:

$$\max \left\{ \begin{array}{ll}
\eta (1 - t / T) \cdot \beta _{B-P} \cdot (1 - t / T) \cdot \beta _{B-N}
\end{array} \right\}$$

The following paths are all the same. What we can do is according to the pricing change can track the invisible driven cause. So that we may know whether there is someone who will buy the ticket in this day.

Finally, our job is to seek the highest expectation of the total revenue of the airline company. We will calculate its expectation of its revenue. And the final result can help to make the best pricing system.

If there will be $C$ purchaser in day $t$, so at that day, the company’s revenue will be:

$$\text{Revenue}(t) = C \cdot f(P)$$

So the equation of total revenue:

$$E [\text{Revenue}(f, C)] = \sum_{t=1}^{T} C f(P)$$

Here is the equation’s restriction:

$$\sum_{t=1}^{T} C \leq S$$
In this paper, the price we want to optimise is not continuous. It belongs to a discrete optimisation. In order to compare the actual effect of our hidden Markov model, we use MATLAB 7.0 to solve this problem.

4 Effects and analysis

In this section, we will choose a standard linear demand curve. As in common, the demand curve shows to be positive. And the non-purchase probability shows to be negative. Next, suppose the hidden transition probability will also float through a fixed rule.

In this section, we choose an airline as an example, HU7608. It uses Boeing 737-800 for flight. 737-800 contains 150 seats. This flight is from Shanghai to Beijing with one flight per-day. The full price is 1,130 Yuan. Its sales process extends to 105 days (15 weeks). The price is from 540 Yuan to 850 Yuan.

In order to avoid extra variables, we would like just choose the Monday flight as an example. That means, the sales process is simplified into 15 terms.

Here, we list the pricing strategy which is being used now.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>HU7608 pricing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>540</td>
</tr>
<tr>
<td>Week 2</td>
<td>540</td>
</tr>
<tr>
<td>Week 3</td>
<td>540</td>
</tr>
<tr>
<td>Week 4</td>
<td>540</td>
</tr>
<tr>
<td>Week 5</td>
<td>540</td>
</tr>
<tr>
<td>Week 6</td>
<td>570</td>
</tr>
<tr>
<td>Week 7</td>
<td>570</td>
</tr>
<tr>
<td>Week 8</td>
<td>570</td>
</tr>
<tr>
<td>Week 9</td>
<td>570</td>
</tr>
<tr>
<td>Week 10</td>
<td>570</td>
</tr>
<tr>
<td>Week 11</td>
<td>570</td>
</tr>
<tr>
<td>Week 12</td>
<td>620</td>
</tr>
<tr>
<td>Week 13</td>
<td>620</td>
</tr>
<tr>
<td>Week 14</td>
<td>850</td>
</tr>
<tr>
<td>Week 15</td>
<td>900</td>
</tr>
</tbody>
</table>

Suppose the passenger demand curve will follow the Poisson distribution, which their coefficient \( \lambda = [2, \ldots, 17] \) in 15 weeks.

First, suppose the transition probability will stay still during the selling period, \( \alpha_{k,l,1} = \alpha_{k,l,T} \). These kind of characteristics always exist in some remote district flights, where other transportation cannot easily reach there. During this process, one key point is the initial probability and the last probability: \( \pi_{k,1}, \pi_{k,15} \). So it is needed to show the effect of using the hidden Markov model under all different situations.

Here will list some conclusions. When the \( \pi_{k,1} = 0.3 \), \( \pi_{k,15} = 0.9 \). The optimal price is shown in Table 2.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>HU7608 optimal pricing 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>540</td>
</tr>
<tr>
<td>Week 2</td>
<td>540</td>
</tr>
<tr>
<td>Week 3</td>
<td>540</td>
</tr>
<tr>
<td>Week 4</td>
<td>540</td>
</tr>
<tr>
<td>Week 5</td>
<td>540</td>
</tr>
<tr>
<td>Week 6</td>
<td>540</td>
</tr>
<tr>
<td>Week 7</td>
<td>540</td>
</tr>
<tr>
<td>Week 8</td>
<td>540</td>
</tr>
<tr>
<td>Week 9</td>
<td>540</td>
</tr>
<tr>
<td>Week 10</td>
<td>540</td>
</tr>
<tr>
<td>Week 11</td>
<td>540</td>
</tr>
<tr>
<td>Week 12</td>
<td>600</td>
</tr>
<tr>
<td>Week 13</td>
<td>633</td>
</tr>
<tr>
<td>Week 14</td>
<td>667</td>
</tr>
<tr>
<td>Week 15</td>
<td>703</td>
</tr>
</tbody>
</table>

The optimised total revenue equals to 88,415. The initial total revenue is 74,630 with 112 tickets sold. About 18.5% increase in total revenue.

If the initial probability is fixed, different last status initial probability will cause different working efforts. One of the key-points is the initial probability difference \( \Delta \).

When \( \Delta = 0.4 \), hidden Markov model’s effects will float because of the different initial probability.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>HU7608 optimal pricing with different ( \Delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pi_{k,1} )</td>
<td>0.1</td>
</tr>
<tr>
<td>Optimal</td>
<td>85,950</td>
</tr>
<tr>
<td>Increase ratio</td>
<td>2.26%</td>
</tr>
</tbody>
</table>

As it is shown, the hidden Markov model pricing strategy did not change the price in the first several rounds because the expectation of raising price was still lower than keeping the price. If we raise the price rashly, the single income increased, but fewer people would buy the ticket at that time. And in the end, total revenue will still decrease. Also, from the calculation, when the demand curve has a greater slope, it will trigger the price-increase earlier. And the optimisation is more close to the needs of customers, so that its benefit maximisation will come true. When become larger, its total increase will become larger as well. The reason why the second column’s ratio is higher than the first column’s ratio is because the growth in week 12 in the initial price list increases the revenue base. And actually, the revenue difference is still increasing.

In sum, under the circumstance that the transition probability stay still, which usually shows up in some irreplaceable but not much necessary flights, hidden Markov model is much close to the demand change and change the price once the demand floated.

In the process of simulation calculation, we found that if all the \( \pi_{k} \) cannot be higher than 0.5, the whole flight prices cannot tend to rise during the sales. Airlines should be increase the initial price as high as possible and ensure that the initial price will not affect the first round decision in the hidden Markov model. That is also the reason why train ticket price nearly change. In this case, we suppose all the transition stay still and only if \( 2\pi_{k,1} + \Delta > 1 \), the airlines can tend to increase its sales price during the sales.

Second, suppose the transition probability will also float during the selling period, as in common, according to a linear change. The slope of ‘purchase’ state to raising price is \( k_{p} \). The slope of ‘non-purchase’ to raising price is \( k_{n} \).

After the calculation, most price raising will happened after the time that the expectation of purchase under the situation of raising is higher than the expectation of non-purchase. Accurately, once the \( s \) comes to \( t \):

\[
t > \frac{k_{n} - \pi_{k,1}(k_{p} + k_{n})}{\Delta \times k_{p}} T
\]

Price shall be raised in the next round.
In order to simulate a change which is closer to the actual situation, during the sales, customer need has three positive float and negative float.

\[ \pi_{3,1} = 0.4, \pi_{4,15} = 0.5, \min \pi_k = 0.3, \max \pi_k = 0.9 \]

Thus, the following three tables will show the simulation of hidden Markov model and compare it with the current strategy.

### Table 4  HU7608 optimal pricing 2

<table>
<thead>
<tr>
<th>Week</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 4</th>
<th>Week 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>540</td>
<td>540</td>
<td>540</td>
<td>569</td>
<td>600</td>
</tr>
<tr>
<td>633</td>
<td>667</td>
<td>667</td>
<td>667</td>
<td>667</td>
</tr>
<tr>
<td>667</td>
<td>667</td>
<td>703</td>
<td>741</td>
<td>741</td>
</tr>
</tbody>
</table>

In this case, the total revenue of hidden Markov model is 92,842, while the existing pricing can only provide 80,990. Hidden Markov model has 14.6\% more in revenue. During this process, the price has been raised for the first time and then it did not change in the next round. It is just the feedback of the simulation because we setup a decrease in demand just two weeks ago, and hidden Markov model immediate reacted. Here is only one case. Actually, no matter how to adjust the initial condition, even add some emergency to influent the customer choices, hidden Markov model will always quickly approximate to the maximum revenue.

All the simulation above has shown the power of hidden Markov model. Compare with the traditional pricing, hidden Markov model pricing provides a new strategy. It is clear that if the price increases rapidly while later to be stable at one price after conducting a hidden Markov model. We believe the core reason is the customer volume. Because once the customer volume is coming to decline, it will give negative influence in the transition array. All the hidden Markov model would like to do is to keep a highest attractiveness to the customer and increase their purchasing possibility. While if the volume decreases, especially the continuously decline in standard normal distribution, it will cause the price never to increase once it detect the decline. On the other hand, this may also explain why under hidden Markov model, there will be less total customer volume in the end.

Thus, the key learning from this letter is that the hidden Markov model can be more preferred when the market thrive. Above all, an increasing customer volume seems to be the most important key driver. If in the future, we will get the chance to input the real data of ticket booking to detect our model further.

### 5 Conclusions

In this paper, we presented a new airline pricing strategy based on hidden Markov model. It is clear that draw there is a strong correlation between the individual decisions and ticket pricing.

In order to build a hidden Markov chain, we focus on both observation sequence and invisible sequence. While the invisible sequence, we believe, shall be separated into two basic elements which are purchasing and non-purchasing. And the observation sequence shall be reflected in the change of the price. Tracking back from the observation sequence, we will find out the purchase rate in each discrete time period. Also, we added another key point, the ticket quantity. With lower ticket quantity, per unit of the price discount will get more customers.

In sum, it is true that the introduction of hidden Markov model can help to analyse in airline ticket prices to seek maximum profits. And management implications also showed up in this letter:

1. To those low competition flights, airlines can continuously increase their price with fixed cycle after the first raise.
2. In those high competition flights, airlines should shorten the raise cycle.

Actually, this new method provides a new thinking of discount process. Its advantages can be concluded as follows: First, hidden Markov model is a highly automatic model. According to the selling data, it will figure out the customer need and its variation trend and then it will quickly simulate the best-selling path with the highest total revenue. Second, we can use discrete data in the hidden Markov model so that we can split out the selling process into day-by-day, which must be more accurate.

In the future, our researches will focus on adding in the competitive factors between airline companies.

### Acknowledgements

This paper is supported by Shanghai science and Technology Committee, No.15DZ1201803.

### References


