Markov Decision Process in Dynamic Optimization of Fare Price

Jan Kislinger

jan.kislinger@gmail.com

Abstract

Fare price optimization is a modification of well-known *Newsboy problem.* In this case, the inventory level (number of seats in the train) is fixed and the decision variables are prices for individual routes. It is a problem with endogenous uncertainty because the demand for train tickets depends on their price. Another distinction of this problem is that there is a rivalry between customers that are not necessarily traveling on the same route. Actual capacity for each route is, therefore, random and dependent with demand for tickets for other routes. We assume that the occupancy of the train follows inhomogeneous Markov process for which the transition intensities depend on chosen prices. We find the optimum using Markov decision process where the price is the decision.

Problem Setup

Sets describing the environment:

Stations	$\mathcal{K} = \{1, ..., K\}$
Routes	$\mathcal{R} = \left\{(k,l) \in \mathcal{K}^2, k < l\right\}$
Seats	$\mathcal{I} = \{1, ..., I\}$
Passenger types	$\mathcal{M} = \{1, ..., M\}$
Ticket types	$\mathcal{C} = \bigcup_{i=1}^{\infty} \mathcal{M}^i$

States indicating actual seat occupancy:

$$\mathcal{S} = \left\{ \boldsymbol{s} \in \mathbb{N}_0^{\binom{K}{2}} : \sum_{k=1}^{h} \sum_{l=h+1}^{K} s_{k,l} \leq I, 1 \leq h \leq K-1 \right\}$$

Optimization Problem

Process $\boldsymbol{X} = \{X_t, \, t \geq 0\}$ follows a Markov process with intensity $\boldsymbol{Q}(t, \boldsymbol{p})$. There exists a reward function $\varrho : \mathcal{S} \to \mathbb{R}$ given by

$$\varrho(\boldsymbol{0}) = 0,$$
$$\varrho(\boldsymbol{s}') - \varrho(\boldsymbol{s}) = p_{\boldsymbol{s},\boldsymbol{s}'},$$

for states $\boldsymbol{s}$ and $\boldsymbol{s}'$ that differ in one ticket with price $p_{\boldsymbol{s},\boldsymbol{s}'}$. The multidimensional price is

$$\boldsymbol{p} = \{p_{k,l,c} : (k,l) \in \mathcal{R}, c \in \mathcal{C}\}$$

The optimization problem is

$$\max_{\varphi} \quad \mathbb{E}_{\varphi}[\varrho(X_T)],$$

where $\varphi : \mathcal{S} \times [0, \infty) \to$ is a decision policy for the prices.

Modeling demand

Let us assume that the demand for ticket of type $c \in \mathcal{C}$ for route $(k,l) \in \mathcal{R}$ follows inhomogeneous Poisson process and the intensity follows log-linear model with

$$\log\left(\lambda_{k,l,c}(p, t; \boldsymbol{\beta})\right) = \beta_{1,k,l,c} + \beta_{2,k,l,c} \log(p) + \beta_{3,k,l,c} t + \beta_{4,k,l,c} t \log(p).$$

The price elasticity of the demand for such model is constant in price and linear in time, i.e.

$$\frac{\partial \lambda_{k,l,c}(p, t; \boldsymbol{\beta}_{k,l,c})}{\partial p} \frac{p}{\lambda_{k,l,c}(p,t)} = \beta_{2,k,l,c} + \beta_{4,k,l,c} t.$$

Cumulative intensity of demand (and expected number of tickets sold at time t with unlimited seat availability) is given by

$$\Lambda_{k,l,c}(p, t; \boldsymbol{\beta}) = \frac{\lambda_{k,l,c}(p, t; \boldsymbol{\beta})}{\beta_{1,k,l,c} + \beta_{2,k,l,c} \log(p)}.$$

Simplifications

Discrete decision times. The original problem is to find an optimum policy, a function that returns for each state of occupation and time optimum price. We simplify the problem by setting fixed times of decision. The problem is reduced to multistage stochastic optimization. Denote these decision times by $T_0, ..., T_{n-1}$.
Single ticket type. The other simplification reduces passengers to single type and allows only one passenger per ticket. The reduction of number of passenger types does not violate our assumptions if the prices for different passenger types are fixed at the same ratio to the full-price passenger type (e.g. senior ticket is always worth 60% of adult ticket for the same route). The simplification for number of passengers is used so that we do not need to estimate the distribution of number of passengers per one ticket. The distribution would be dependent on both time and price. These simplifications imply that $\mathcal{C} = \mathcal{M} = \{1\}$.

Simulated Optimization

The problem is computationally unsolvable because of the number of states. The distribution of the process at given time cannot be calculated and stored in memory of a PC. One solution is to estimate the optimum using simulated optimization. We propose two methods for finding the optimum – response surface and cross entropy for noisy optimization. Both of them are described for single-stage optimization. The reader can easily generalize these algorithms for multistage optimization.

Removing Exogeneity

It is possible to separate the dependence on price from the intensity of the Poisson process (see Proposition 1). With this we can simulate the demand for single route as compound Poisson process where intensity of arrivals of passengers (passenger show interest to buy a ticket) depends only on time and the alternative distribution (indicating whether the passenger actually buys the ticket) depends on both time and price. We can simulate for each passenger the maximum price that is willing to pay. Once we want to evaluate the outcome for different price we distinguish passengers that buy the ticket or not. Therefore, we do not need to simulate the demand for each price separately.

Proposition 1. *Let $N = \{N_t, t \geq 0\}$ be an inhomogeneous Poisson process with intensity $\lambda(t)$ and jump times $T_1, T_2, ...$ Let $\pi : [0, \infty) \to (0,1)$ be a right-continuous function and $Z_i \sim Alt(\pi(T_i))$. Then the family of random variables*

$$M_t = \sum_{i=1}^{\infty} Z_i \times \mathbb{I}[t \geq T_i].$$

is an inhomogeneous Poisson process with intensity $\lambda(t)\pi(t)$.

Algorithms

Denote by $\boldsymbol{p}$ the vector of all prices and by ϱ the simulated outcome (price of all sold tickets).
Algorithm 1 (Response Surface).

1. Set initial value $\boldsymbol{p}^{(0)}$ and put $k := 0$.
2. Generate data (multiple independent observations) from distribution $(\boldsymbol{p}, \varrho)$, where the predictors $\boldsymbol{p}$ have the distribution centered at $\boldsymbol{p}^k$. Join the data with previously generated data (if any).
3. Fit the local polynomial regression at point $\boldsymbol{p}^{(k)}$. If the estimated first-order derivatives of μ are sufficiently close to zero then continue with step 5.
4. Update $\boldsymbol{p}^{(k+1)}$ as in gradient method using estimated derivatives, increase $k := k+1$ and repeat steps 2–4.
5. Return value $\boldsymbol{p}^{(k)}$.

Algorithm 2 (Cross Entropy for Noisy Optimization).

1. Set initial parameter $\boldsymbol{\psi}^{(0)}$ and counter $k := 1$.
2. Generate data (multiple independent observations) of size n from distribution $(\boldsymbol{p}, \varrho)$, where the predictors have the distribution $\boldsymbol{p} \sim f(\,\cdot\,; \boldsymbol{\psi}^{(k-1)})$. Let γ_k be the $(1-r)$-quantile of $\varrho_1, ..., \varrho_n$.
3. Estimate the parameter $\boldsymbol{\psi}^{(k)}$ from the best performing observations using maximum likelihood
$$\boldsymbol{\psi}^{(k)} = \max_{\psi} \sum_{\boldsymbol{p}_i \geq \gamma_k} \log(f(\boldsymbol{p}_i; \boldsymbol{\psi})).$$
4. If the distribution $f(\,\cdot\,; \boldsymbol{\psi}^{(k)})$ is almost degenerated then continue with step 6, otherwise increase $k := k+1$ and repeat steps 2–4.
5. Return value $\boldsymbol{p}$ such that the distribution $f(\,\cdot\,; \boldsymbol{\psi}^{(k)})$ is almost degenerated at point $\boldsymbol{p}$.

Numerical results

Each part (route between two consequent stations) of the path of the train contains different number of routes that goes through this part. For example, if the path has 15 stations the number of routes going through part between the first two stations is 14 and this number in the middle of the way is 56. Therefore, the prices should be higher (even if the demand is the same) in the middle routes then the edge ones. This idea is supported by numerical result with six stations. The table shows estimated optimum ticket prices for pairs of boarding (rows) and exiting (columns) stations.

	2	3	4	5	6
1	171.0	215.8	336.2	441.1	487.2
2		190.8	305.1	385.2	448.8
3			253.0	310.8	342.9
4				137.1	179.6
5					114.9