VALUE GAP DUE TO MODEL ERROR

ASHWIN RAO AND AJAY NERURKAR

1. Introduction

This is a brief article to set the foundation for exploring the gap in Returns due to Model Error within the context of a Markov Decision Process. We consider a very simple example of Price Optimization to illustrate these concepts.

2. Notation and Abstract Setting

Let us say that we have estimated a model M_1. Solving (optimizing) the MDP with this model M_1 gives us the policy π_1^* and the corresponding Value Function V_1^*. Now let us say that afterwards we come to know the true model is M_2. Applying policy π_1^* on the true MDP (with true model M_2) gives us the Value Function V_2. Had we known the true model M_2 beforehand, we would have instead arrived at the optimal policy π_2^* and the optimal Value Function V_2^*. We want to understand the gaps between the three value functions: $[V_1^*, V_2, V_2^*]$.

- We can conceptualize the gap between V_1^* and V_2^* as that due to "Model Estimation Inadequacy"
- We can conceptualize the gap between V_1^* and V_2 as that due to "Outcome Surprise"
- We can conceptualize the gap between V_2 and V_2^* as the slippage due to "Lack of Prior Information"

We also want to understand the gap between the optimal policy π_1^* and π_2^*.

3. Price Optimization

For ease of illustration of the above concepts, consider a very simple case of single time-step Price Optimization, an exponential model of Price Elasticity of Demand, and little chance of lost sales.

Let the model M_1 be given by the demand function $f_1(p) = \alpha_1 \cdot e^{-\beta_1 p}$. The single time-step revenue for a choice of price p is: $p \cdot f_1(p) = \alpha_1 \cdot p \cdot e^{-\beta_1 p}$. The optimal price (for maximum revenue) is given by:

$$p_1^* = \frac{1}{\beta_1}$$

and the maximum revenue (optimal value function) is given by:

$$V_1^* = \frac{\alpha_1}{\beta_1 \cdot e}$$

Now let us say that the true model M_2 is given by the demand function $f_2(p) = \alpha_2 \cdot e^{-\beta_2 p}$. Executing the previously identified optimal policy for M_1 (i.e., p_1^*) on the true MDP (with true model $f_2(p)$) give us the value function:

$$V_2 = p_1^* \cdot (\alpha_2 \cdot e^{-\beta_2 p_1^*}) = \frac{\alpha_2}{\beta_1} \cdot e^{-\frac{\beta_2}{\beta_1}}$$

We also know that:

$$p_2^* = \frac{1}{\beta_2}$$

$$V_2^* = \frac{\alpha_2}{\beta_2 \cdot e}$$

Now let us look at the percentage changes between $[V_1^*, V_2, V_2^*]$ and the percentage change between p_1^* and p_2^*.

- $1 - \frac{V_2}{V_2^*} = 1 - \frac{\beta_2}{\beta_1} \cdot e^{1-\frac{\beta_2}{\beta_1}}$
- $1 - \frac{V_2}{V_1^*} = 1 - \frac{\alpha_2}{\alpha_1} \cdot e^{1-\frac{\beta_2}{\beta_1}}$
- $1 - \frac{V_2^*}{V_1^*} = 1 - \frac{\alpha_2 \beta_1}{\alpha_1 \beta_2}$
- $1 - \frac{p_2^*}{p_1^*} = 1 - \frac{\beta_1}{\beta_2}$

Now let us define α_2 as a small perturbation of α_1 and β_2 as a small perturbation of β_1. Let $\frac{\alpha_2}{\alpha_1} = 1 + \delta_\alpha$ and let $\frac{\beta_2}{\beta_1} = 1 + \delta_\beta$. Then,

- $1 - \frac{V_2}{V_2^*} = 1 - (1 + \delta_\beta) \cdot e^{-\delta_\beta} \approx 1 - (1 + \delta_\beta)(1 - \delta_\beta + \frac{\delta_\beta^2}{2}) = \frac{\delta_\beta^2}{2}$
- $1 - \frac{V_2}{V_1^*} = 1 - (1 + \delta_\alpha) \cdot e^{-\delta_\beta} \approx 1 - (1 + \delta_\alpha)(1 - \delta_\beta) = \delta_\beta - \delta_\alpha$
- $1 - \frac{V_2^*}{V_1^*} = 1 - \frac{1+\delta_\alpha}{1+\delta_\beta} \approx 1 - (1 + \delta_\alpha)(1 - \delta_\beta) \approx \delta_\beta - \delta_\alpha$
- $1 - \frac{p_2^*}{p_1^*} = 1 - \frac{1}{1+\delta_\beta} \approx 1 - (1 - \delta_\beta) = \delta_\beta$