A EQUIVALENCE BETWEEN TREATMENT EFFECTS AND VALUE OPTIMIZATION

Here, we show that the Knapsack formulation in Equation (1), written with treatment effect terms τ^r and τ^c, can be reformulated using the counterfactual terms $R(t)$ and $C(t)$, with an updated budget B':

$$\begin{aligned} \max \quad & \sum_{i=1}^{n}\sum_{t=0}^{m} z_{it} \cdot R_i(t) \\ \text{s.t.} \quad & \sum_{i=1}^{n}\sum_{t=0}^{m} z_{it} \cdot C_i(t) \leq B' \\ & \sum_{t=0}^{m} z_{it} = 1 \\ & z_{ij} \in \{0,1\}, \forall i. \end{aligned} \tag{11}$$

Since the instances are assumed to be *i.i.d.*, it suffices to demonstrate equivalence for a single instance i:

$$\begin{aligned} t^* &= \arg\max_{t\in\mathcal{T}} \tau^r_{it} - \tau^c_{it} \\ &= \arg\max_{t\in\mathcal{T}} R_i(t) - R_i(0) - \lambda\left(C_i(t) - C_i(0)\right) \\ &= \arg\max_{t\in\mathcal{T}} R_i(t) - \lambda C_i(t) + \left(\lambda C_i(0) - R_i(0)\right) \\ &= \arg\max_{t\in\mathcal{T}} R_i(t) - \lambda C_i(t) + Const_i \\ &= \arg\max_{t\in\mathcal{T}} R_i(t) - \lambda C_i(t), \end{aligned} \tag{12}$$

where $Const_i$ is just a constant value for each instance. For the constraint, we observe that the budget limit B must be updated with the value of the current treatment, specifically:

$$B' = B + \sum_i C_i(0). \tag{13}$$

B PROOF THAT MCKP WITH TRIGGER REWARDS AND COSTS ONLY REQUIRES CONVERTED DATA

Given a MCKP marketing problem formulated as as equations (1) or (11), the solution can be approximated by (4). Here, we will prove that in the case of trigger rewards and costs as defined in (5), only converted quantities $\Pr(T=t|x,\phi=1)$, $E[R|x,\phi=1]$, and $E[C|x,\phi=1]$ and propensity scores $\Pr(T=t|x)$ are required to solve the problem. In the additional case that the data comes from RCTs or policy data, the propensity scores are known, and only converted quantities are necessary.

To prove that that is so, let us focus on the decision boundary of the MCKP problem in equation (4), in whether an instance should be treated or not :

$$t = \arg\max_{k\in\mathcal{T}} \left(\tau^r_{ik} - \lambda^* \cdot \tau^c_{ik}\right)$$

To show that only converted quantities are necessary, we just need to show that the decision boundary can be estimated with those quantities. For that, it is necessary to roll-out the expected value of the quantities involved using the conditional expectation over a partition and the Bayesian theorem:

$$\begin{aligned} E[R|x,T=t] &= E[R|x,T=t,\Phi=1]\Pr(\Phi=1|x,T=t) \\ &\quad + E[R|x,T=t,\Phi=0]\Pr(\Phi=0|x,T=t) \\ &= E[R|x,T=t,\Phi=1]\Pr(\Phi=1|x,T=1) \\ &= E[R|x,T=1,\Phi=1]\,\frac{\Pr(T=t|x,\Phi=1)\Pr(\Phi=1|x)}{\Pr(T=t|x)}, \end{aligned} \tag{14}$$

where $E[R|x,T=t,Y=1]$ represents the converted reward, $\Pr(T=t|x,\Phi=1)$ the ratio of conversions in treatment $T=t$ versus the other treatments $T=\neg t$, $\Pr(\Phi=1|x)$ the average conversion for context x (over all treatment groups), and $Pr(T=t|x)$ the propensity of a context x being in treatment group $T=t$. An equivalent derivation can be done for cost attribute $E[C|x,T=t]$.

Plug-in the Bayes transformed $E[R|x,T=t]$ of the previous equation into the CATE definition of (2), to obtain:

$$\begin{aligned} \tau^r(x,t) &= \Bigg[\frac{E[R|x,T=t,\Phi=1]\Pr(T=1|x,\Phi=1)}{\Pr(T=t|x)} \\ &\quad - \frac{E[R|x,T=0,\Phi=1]\Pr(T=0|x,\Phi=1)}{\Pr(T=0|x)}\Bigg]\cdot\Pr(\Phi=1|x), \\ &= \tau^{r,\Phi=1}(x,t)\cdot\Pr(\Phi=1|x) \end{aligned} \tag{15}$$

where $\tau^{r,\Phi=1}(x,t)$ is the treatment effect of reward per conversion (also called incremental Reward per Conversion in Proença and Moraes [38]) that only depends on converted quantities and the propensity score, and the average conversion $\Pr(\Phi=1|x)$ was isolated outside. The incremental cost per conversion or $\tau^{c,\Phi=1}(x,t)$ can be derived for the cost attribute $\tau^c(x,t)$.

Going back to the decision boundary, we see that to prove that one of the terms $\tau^r_{it} - \lambda^*\cdot\tau^c_{it}$ is the maximum only requires to show that it is larger than the terms with other treatments and above zero (control group).

$$\begin{aligned} &\tau^r_{it} - \lambda^*\cdot\tau^c_{it} \geq \tau^r_{it'} - \lambda^*\cdot\tau^c_{it'} \Leftrightarrow \\ &\left[\tau^{r,\Phi=1}_{it} - \lambda^*\cdot\tau^{c,\Phi=1}_{it}\right]\cdot\Pr(\Phi=1|x) \geq \left[\tau^{r,\Phi=1}_{it'} - \lambda^*\cdot\tau^{c,\Phi=1}_{it'}\right]\cdot\Pr(\Phi=1|x) \Leftrightarrow \\ &\tau^{r,\Phi=1}_{it} - \lambda^*\cdot\tau^{c,\Phi=1}_{it} \geq \tau^{r,\Phi=1}_{it'} - \lambda^*\cdot\tau^{c,\Phi=1}_{it'}, \end{aligned} \tag{16}$$

where the last expression only depends on converted terms (and the propensity score). To prove that a treatment is larger than the control effect group, we only need to show that it is above zero, which is trivially shown by setting $\tau^r_{it'} - \lambda^*\cdot\tau^c_{it'} = 0$ for the control group.

C TRAINING TIME COMPLEXITY

The time complexity of a gradient-descent learning algorithm, such as linear regression, gradient-boosting machines, or neural networks, is proportional to the number of iterations k and the number of instances in the training data n. Formally, the time complexity can be expressed as:

$$O(k\cdot n\cdot K_m), \tag{17}$$

where K_m is a constant associated with the model architecture, such as the number of features for linear regression or the complexity associated with the type of layers and number of hidden layers for neural networks.

Thus, it is straightforward to show that reducing the *necessary training data* by 50% results in an asymptotic reduction of the training time by approximately 50%.