Prior-Independent Dynamic Auctions for a Value-Maximizing buyer

Authors: Yuan Deng, Hanrui Zhang
Speaker: Zhuming Shi
Mentor: Zhaohua Chen

April 6, 2022

Content

1 Introduction

2 Notations

3 Mechanism

4 Proof

5 Open Problems

Introduction

- 传统的 setting 都是已知买家的价值分布下的，我们的工作中买家的价值分布未知，靠学习出来
- 以往的工作集中在最大化效用（价值-支付价格）的的买家，我们现在考虑最大化所得价值的 1 个买家，他只要满足本轮（第 t 轮）回报率 $\geq \tau_t$，每轮回报率要求可以不同
- 相对已知买家价值分布的机制，我们比它差 $\tilde{O}(T^{2/3})$，T 是总轮数
- 作者采用了传统的先探索再利用 (exploration-exploitation）的思路，首先在探索阶段结束后估计买家的价值分布，然后给出一个鲁棒的收益最优的（revenue-optimal）的机制，最后在利用阶段的每一轮中反复使用

动态机制

- b_t 为买家的第 t 轮出价
- $x_t \in [0, 1]$ 是在第 t 轮中卖家分配多少比例的拍品给买家，x_t 由历史出价 $b_1 \cdots b_t$（包括本次出价）决定
- p_t 是同样由历史出价 $b_1 \cdots b_t$（包括本次出价）决定的本次的买家支付价格
- 动态机制被定义为 $M = \{M_t\}_{t\in[T]} = \{(x_t, p_t)\}_{t\in[T]}$

买家的约束

- 考虑 1 个买家，买家最大化所得的价值
- **私有的**买家价值分布 $\mathcal{D} \in \Delta([0,1])$
- **公开的**第 t 轮 ROI 阈值 $\tau_t \in [1,+\infty)$，买家本轮期望回报率至少 τ_t，也就是要满足

$$\mathbb{E}\left[x_t v_t - p_t \tau_t\right] \geq 0$$

- **公开的**贴现因子 $\lambda \in (0,1)$，买家是不耐心的，拖得越久价值越低，第 t 轮所得价值要乘上 λ^t
- 买家是理性的，总是选择以下的最优解作为他们的购买策略

买家的优化目标

给定一个动态机制 M，则买家面对的优化问题是

$$\max_{b_1,\ldots,b_T} \mathop{\mathbb{E}}_{(v_1,\ldots,v_T)\sim\mathcal{D}^T} \left[\sum_{t\in[T]} \lambda^t \cdot x_t \cdot v_t \right]$$

subject to $\mathop{\mathbb{E}}_{v_t\sim\mathcal{D}} [x_t \cdot v_t - \tau_t \cdot p_t \mid v_1, \ldots, v_{t-1}] \geq 0,$
$\forall t \in [T], v_1, \ldots, v_{t-1} \in [0, 1]$

卖家的 first-best revenue

- 第 t 轮中卖家生产 1 单位拍品的成本是 $c_t \in [0,1]$
- 在卖家已知买家价值分布 $\mathcal{D}$ 的条件下，卖家的第 t 轮最优期望收益为

$$\mathrm{OPT}_t(\mathcal{D}) = \mathop{\mathbb{E}}_{v\sim\mathcal{D}}[\max\{v/\tau_t - c_t, 0\}]$$

- 卖家在 T 轮中总的最优收益 first-best revenue 为 $\mathrm{OPT}(\mathcal{D}) = \sum_{t\in[T]} \mathrm{OPT}_t(\mathcal{D})$

卖家的目标

找到一种拍卖机制 M，使得

- 对于买家的任意一种价值分布 $\mathcal{D}$，该买家都存在一种满足 ROI 限制的购买策略
- 在此策略下，卖家的总收益和卖家在已知买家价值分布 $\mathcal{D}$ 时可以取得的最优总收益相比，差别在 $o(T)$ 级别

$$\sup_{\mathcal{D}\in\Delta(V)}\left(\mathrm{OPT}(\mathcal{D})-\sum_{t\in[T]}(p_t-x_t\cdot c_t)\right)=o(T)$$

探索部分

1. 计算参数 $\varepsilon = T^{-1/3}$, $T_1 = \Theta\left(\log T/\varepsilon^2\right)$, $T_2 = \Theta\left(\frac{\log\left((1-\lambda)\cdot\varepsilon^2\right)}{\log\lambda}\right)$
2. 前 $T_1 + T_2$ 轮是探索轮，对轮数 $t = 1, \ldots, T_1 + T_2$, 买家出价 $b_t \in \mathbb{R}_+$ 则分配 $x_t = \min\{b_t, 1\}$ 部分拍品给买家，收取的价格为，假设买家心中拍品价值为 $\min\{b_t, 1\}$ 时，当前轮 ROI 约束下的最大支付价格 $p_t = x_t \cdot \frac{\min\{b_t, 1\}}{\tau_t}$.

拟合买家价值分布

1. 利用前 T_1 轮数据估计买家的价值分布 $\hat{\mathcal{D}}$，在这个价值分布 $\hat{\mathcal{D}}$ 中，任意一个价值 $v \in [0,1]$ 按照 $\hat{\mathcal{D}}$ 被抽中的概率正比于 v 在前 T_1 轮的 $\min\{b_t, 1\}$ 中出现的频率

$$\hat{\mathcal{D}}(v) = \Pr_{v' \sim \hat{\mathcal{D}}}[v' = v] = \frac{\sum_{t \in [T_1]} \mathbb{I}[\min\{b_t, 1\} = v]}{T_1}$$

应用部分

1 从 $T_1 + T_2 + 1$ 轮开始直到结束的每一轮 $t = T_1 + T_2 + 1, \ldots, T$, 计算

$$q_t = \mathop{\mathbb{E}}_{v \sim \hat{\mathcal{D}}}\left[v/\tau_t \mid v/\tau_t \geq c_t\right] - \frac{\varepsilon}{\Pr_{v \sim \hat{\mathcal{D}}}\left[v/\tau_t \geq c_t\right]}$$

这个 q_t 是买家价值分布 $\hat{\mathcal{D}}$ 的情况下，卖家不亏本 $\frac{b_t}{\tau_t} \geq c_t$ 时买家在满足 ROI 时可以接受的最高支付价格的期望，再减去一个鲁棒性项。如果本轮买家报价 b_t 能使售出部分拍品后，卖家回本 $\frac{b_t}{\tau_t} \geq c_t$，则把本轮拍品全部卖出 $x_t = 1$ 并索要支付价格 $p_t = \max\{q_t, c_t\}$，否则不出售拍品 $x_t = 0$ 同时不收取支付 $p_t = 0$.

证明思路

loss 分两部分：探索阶段损失较大，试探较多，但控制住处于该阶段时间，就能控制住探索阶段的误差；应用阶段只要控制 $\hat{\mathcal{D}}$ 和真实的买家价值分布很接近，就能控制住应用阶段的误差。

探索 (exploration) 阶段

对于探索 (exploration) 阶段，作者给出了一个衡量两个价值的概率分布 $\mathcal{D}, \hat{\mathcal{D}}$ 的差别的函数，是它们的联合分布中的分布 D_j 中两个点的期望距离的下确界。$\mathcal{D}, \hat{\mathcal{D}}$ 越接近，$W_1(\mathcal{D}, \hat{\mathcal{D}})$ 就越小

$$W_1\left(\mathcal{D}_1, \mathcal{D}_2\right) := \inf_{\mathcal{D}_j \in \Gamma(\mathcal{D}_1, \mathcal{D}_2)} \left(\mathop{\mathbb{E}}_{(x,y)\sim\mathcal{D}_j} [\|x - y\|] \right)$$

其中 $\Gamma(\mathcal{D}_1, \mathcal{D}_2)$ 是一族 $\mathbb{R}^2$ 上的分布，在两个维度上分别有 $\mathcal{D}$ 和 $\hat{\mathcal{D}}$ 的边缘分布。作者证明了在上述 T_1 轮次下，有 $1 - 1/T$ 的高概率 $\mathcal{D}, \hat{\mathcal{D}}$ 足够接近，$W_1(\mathcal{D}, \hat{\mathcal{D}}) \leq \epsilon$

利用 (exploitation) 阶段

在利用 (exploitation) 阶段，作者证明了只要 $\mathcal{D}, \hat{\mathcal{D}}$ 足够接近，$T_1 + T_2 + 1$ 轮之后的每轮，卖家收益和已知 $\mathcal{D}$ 时的最优收益只差 2ϵ，$W_1(\mathcal{D}, \hat{\mathcal{D}}) \leq \epsilon$，则 $\forall t \in \{T_1 + T_2 + 1, \cdots, T\}$，有期望收益

$$\mathrm{REV}_t(\mathcal{D}, \hat{\mathcal{D}}) \geq \mathrm{OPT}_t(\mathcal{D}) - 2\epsilon$$

把 bound 合起来

最终把两结论组合起来，根据之前调节的 $\epsilon = T^{-1/3}$, $T_1 = \Theta\left(\log T/\varepsilon^2\right)$, and $T_2 = \Theta\left(\frac{\log\left((1-\lambda)\cdot\varepsilon^2\right)}{\log\lambda}\right)$ 这些参数，做加和，可以确定有 $1 - 1/T$ 的高概率总收益至少有 $\mathrm{OPT}(\mathcal{D}) - O(T^{2/3}\log T)$

Open Problems

- 从 $O(T^{2/3}\log T)$ 继续改进，去掉 $\log T$？
- 目前的分布 $\hat{\mathcal{D}}$ 只在有限个值上有概率取得，但实际买家价值分布更可能是连续的，是否有改进空间？