Linear Privacy Pricing

Kenny Peng and Bhaskar Krishnamachari

November 2021

1 Introduction

2 Optimization Problem Formulation

Given a set of epsilon values, $E = \{\epsilon_1, \epsilon_2, \ldots, \epsilon_L\}$ where $0 < \epsilon_1 < \epsilon_2 < \cdots < \epsilon_L$, the platform broadcasts a payment vector $\vec{p} = (p_1, p_2, \ldots, p_L) \in \mathbb{R}^L$ where p_i corresponds to ϵ_i.

Users are given the utility function

$$u(\epsilon, p) = p - \epsilon c$$

and are expected to choose an ϵ_i which maximizes this (a greedy strategy). If the maximal utility is negative for a user, they will not participate. Furthermore, let each user individually draw c from a known distribution.

The platform is given the utility function

$$U_p(\vec{n}, \vec{p}) = R(\vec{n}) - \vec{p} \cdot \vec{n}$$

where $\vec{n} = (n_1, n_2, \ldots, n_L)$ and n_i is the number of users that have chosen ϵ_i.

For ease of exposition, we extend the domain E with the value $\epsilon_0 = 0$ in order to contain the decision to not participate. We presume the utility of this decision is zero, so this adds the constraint $p_0 = 0$.

3 Solution when restricted to strictly increasing pricing

The marginal utility in going from ϵ_i to ϵ_{i+1} can be expressed as

$$u(\epsilon_{i+1}, p_{\epsilon,i+1}) - u(\epsilon_i, p_{\epsilon,i}) = (p_{i+1} - p_i) - c(\epsilon_{i+1} - \epsilon_i)$$

so a user would do so if the following is true

$$c < \frac{p_{i+1} - p_i}{\epsilon_{i+1} - \epsilon_i}$$

Now consider the set of all payment vectors satisfying this super-linear rule

$$\frac{p_{i+1}-p_i}{\epsilon_{i+1}-\epsilon_i} \geq \frac{p_i-p_{i-1}}{\epsilon_i-\epsilon_{i-1}} \geq 0$$

It can be shown that, if a user would have gone from ϵ_{i-1} to ϵ_i, then they also would go from ϵ_i to ϵ_{i+1} because

$$\frac{p_{i+1}-p_i}{\epsilon_{i+1}-\epsilon_i} \geq \frac{p_i-p_{i-1}}{\epsilon_i-\epsilon_{i-1}} \geq c$$

Therefore, if the base case is true (the user goes from ϵ_0 to ϵ_1), then the user would *always* choose a greater epsilon, ending at ϵ_L. Therefore, such payment vectors would split the distribution $f(c)$ into two parts? One part does not participate, and the other part participates at ϵ_L and gets paid p_L.

However, a subset of payment vectors would also accomplish this: that satisfying the linear rule

$$\frac{p_{i+1}-p_i}{\epsilon_{i+1}-\epsilon_i} = \frac{p_i-p_{i-1}}{\epsilon_i-\epsilon_{i-1}} = k$$

for some arbitrary $k \geq 0$. For any payment vector in the super-linear set, the payment vector in the linear set that splits $f(c)$ in the same way has $k = p_L/\epsilon_L$.

Now consider the set of all payment vectors satisfying this sub-linear rule, also including the linear rule as a subset:

$$0 \leq \frac{p_{i+1}-p_i}{\epsilon_{i+1}-\epsilon_i} \leq \frac{p_i-p_{i-1}}{\epsilon_i-\epsilon_{i-1}}$$

in other words p_i is strictly increasing, linearly or sub-linearly, over i. We will refer to this set, $A \subset \mathbb{R}^L$, as "diminishing-returns pricing".

It can be shown that, if the user would do have gone from ϵ_i to ϵ_{i+1}, they also would have gone from ϵ_{i-1} to ϵ_i because

$$c < \frac{p_{i+1}-p_i}{\epsilon_{i+1}-\epsilon_i} \leq \frac{p_i-p_{i-1}}{\epsilon_i-\epsilon_{i-1}}$$

This result can be used to conclude that if the marginal utility in going from ϵ_i to ϵ_{i+1} was positive for some i, it was also positive for all previous to i.

In other words, a utility-maximizing user would progressively choose greater ϵ_i until the following is satisfied?

$$\frac{p_{i+1}-p_i}{\epsilon_{i+1}-\epsilon_i} < c < \frac{p_i-p_{i-1}}{\epsilon_i-\epsilon_{i-1}}$$

Given each user draws their c from $f(c)$, the platform is aware that the probability of a user choosing some ϵ_i is therefore

$$g(\epsilon_i) = \int_{a_i}^{b_i} f(c)\,dc \text{ where } a_i = \frac{p_{i+1}-p_i}{\epsilon_{i+1}-\epsilon_i} \text{ and } b_i = \frac{p_i-p_{i-1}}{\epsilon_i-\epsilon_{i-1}}$$

with the boundary conditions $a_L = -\infty$ and $b_0 = \infty$, and these probabilities, dependent on $f(c)$, defines the multinomial distribution

$$h(\vec{n}) = \binom{N}{n_0, n_1, \ldots, n_L} \prod_{i=0}^{L} g(\epsilon_i)^{n_i}$$

from which $\vec{n}$ will be drawn (N is the total number of users).

So, when considering "diminishing-return" pricing, the platform may maximize the expected value of their utility with

$$\vec{p}_{max} = \underset{\vec{p} \in A}{\operatorname{argmax}} \{E\{U_p\}\}$$

4 Transformation into combinatorial optimization problem after discretization

Let the space $\mathbb{R}^L$ be discretized into the space

$$P = \{\Delta p, 2\Delta p, \ldots, N_P \Delta p = p_U\}^L \subset \mathbb{R}^L$$

where p_U is an arbitrary upper limit. (Subjectively, p_U should be $k\epsilon_L$ where k is a value above which an insignificant number of users would draw their c, according to the known $f(c)$.)

The intersection $A_P = A \cap P$ is the set of all "diminishing-returns" payment vectors that can exist after discretization.

Further, let us require that $\epsilon_i - \epsilon_{i-1}$ be fixed to a constant $\Delta\epsilon$.

From here, further proof requires representing each $\vec{p} \in A_P$ as L draws from $\{\Delta p, 2\Delta p, \ldots, N_P \Delta p = p_U\}$ without order and with replacement. Since $p_0 = 0$, we propose a procedure that maps $\vec{p} \in A_P$ to the set of all permutations for some L draws, $X \in B$, that establishes the necessary one-to-one correspondence. These procedures relate each individual draw to a *difference* between components of the $\vec{p}$, not one of the components themselves. Informally, they are

- $f : A_P \mapsto B$
  1. $d \leftarrow (p_1 - p_0, p_2 - p_1, \ldots, p_L - p_{L-1})$
  2. Return the set of all permutations of d
- $f^{-1} : B \mapsto A_P$
  1. $d \leftarrow x \in X$ s.t. x_i is increasing over i
  2. Return $p = (x_1 + p_0, x_2 + p_1, \ldots, x_L + p_{L-1})$

Draws without order and with replacement are justified because all permutations of L draws together represent a single $\vec{p}$ by the "diminishing-returns" condition:

$$0 < \frac{p_{i+1} - p_i}{\epsilon_{i+1} - \epsilon_i} \leq \frac{p_i - p_{i-1}}{\epsilon_i - \epsilon_{i-1}}$$

$$0 < \frac{p_{i+1} - p_i}{\Delta\epsilon} \leq \frac{p_i - p_{i-1}}{\Delta\epsilon}$$

$$0 < p_{i+1} - p_i \leq p_i - p_{i-1}$$

That is, the order of the L draws does not matter because we need to use them in decreasing order, and replacement allows the equality case. Furthermore, these differences cannot be negative, and the consequence of the upper limit in this new perspective is the condition $\sum_{i=1}^{L} x_i \leq p_U$.

The equality case of this particular condition, combined with the previous conditions, is directly comparable with another problem: enumeration of all possible integer partitions of N_P into exactly L parts.

If X_n is the set of all pricing vectors derived from all partitions of some number n into L parts, then all possible pricing vectors is the set

$$A_P = \bigcup_{n=1}^{N_P} X_n$$

5 Proof of search space boundedness over L

If $p_k(n)$ is a function that outputs the number of partitions of n into exactly k parts, we can derive from the above that the size of A_P is exactly

$$\sum_{n=1}^{N_P} p_L(n)$$

If any positive integer cannot be split in an infinite number of ways, barring the use of zero and negative numbers, then let $p(n)$ be the *finite* number of ways n can be split, in general. Therefore, it can be shown by definition that

$$p_L(n) \leq \sum_{k=0}^{n} p_k(n) = p(n) < \infty$$

Further, we can conclude from this that an approximation of the global argmax, the global argmax in A_P, can at least be found via a brute-force search.

6 Case Study

7 Conclusion