On incentive pricing algorithms under the lack of information about agent utilities

Rokhlin D.B.
Southern Federal University, Rostov-on-Don

We consider a leader, who prices a resource or good, and tries to change the behavior of selfish agents in a desired way. Usually, leader's aim is to stimulate a broadly understood socially optimal behavior. In this case the desired price can be approximated by using dual gradient-based algorithms, which require the information only about agent reactions. We discuss two such cases: resource pricing in communication networks [1], and transfer pricing within a corporation [2].

Let us consider the last case in more detail. Assume that a firm consists from n production and m sales divisions. There are d commodities produced by each production division. The same commodities are saled by each sales division. Denote by $f_i : X_i \mapsto \mathbb{R}_+$, $i = 1, \dots, m$ the revenue functions of the sales divisions, and by $g_i : Y_i \mapsto \mathbb{R}_+$, $i = 1, \dots, n$ the cost functions of the production divisions. A vector $x_i \in X_i$ describes the amounts of commodities to be sold by i-th sales division, and $y_i \in Y_i$ describes the amounts of commodities to be produced by i-th production division.

Assumption 1. The sets X_i, Y_i are convex, compact, and contain $[0, \varepsilon]^d$.

Assumption 2. The functions $f_i : X_i \mapsto \mathbb{R}_+$ (resp., $g_i : Y_i \mapsto \mathbb{R}_+$) are Lipschitz, non-decreasing in each argument, and $f_i(0) = g_i(0) = 0$.

Assumption 3. The functions f_i (resp., g_i) are strongly concave (resp., strongly convex).

The firm announces the commodity *transfer price* vector $\lambda_t \in \mathbb{R}^d_+$ with the obligation to buy the commodities at these prices from the production divisions, and sell them to the sales divisions. Put $\langle a, b\rangle = \sum_{i=1}^d a_i b_i$. Optimal division (agent) reactions are defined by

$$\widetilde{x}_i(\lambda) \in \arg\max_{x_i \in X_i} (f_i(x_i) - \langle \lambda, x_i\rangle), \quad i = 1, \dots, m,$$
$$\widetilde{y}_i(\lambda) \in \arg\max_{y_i \in Y_i} (\langle \lambda, y_i\rangle - g_i(y_i)), \quad i = 1, \dots, n,$$

We will say that the plan $\widetilde{z}(\lambda) = (\widetilde{x}(\lambda), \widetilde{y}(\lambda))$ is stimulated by the transfer price vector λ. The goal of the firm manager is to stimulate the agent reactions coinciding with the optimal solution $z^* = (x^*, y^*)$ of the total profit maximization problem:

$$F(x, y) = \sum_{i=1}^m f_i(x_i) - \sum_{i=1}^n g_i(y_i) \to \max_{(x,y)\in S},$$
$$S = \left\{ (x, y) \in Z : \sum_{i=1}^m x_i = \sum_{j=1}^n y_j \right\}, \quad Z = \prod_{i=1}^m X_i \times \prod_{j=1}^n Y_j.$$

Applying the SOLO FTRL algorithm [3] to the dual problem, we get the recurrence relation

$$\lambda_t = -\frac{\sum_{j=1}^{t-1} \Delta\widetilde{z}(\lambda_j)}{\sqrt{\sum_{j=1}^{t-1} \|\Delta\widetilde{z}(\lambda_j)\|^2}}, \quad \lambda_0 = 0; \quad \Delta\widetilde{z}(\lambda) := \sum_{i=1}^n \widetilde{y}_i(\lambda) - \sum_{i=1}^m \widetilde{x}_i(\lambda).$$

The obtained algorithm uses only the information on division reactions to current prices. It does not depend on any parameters and requires no information on the production and cost functions (from the manager point of view). The next result shows that the optimality gap and feasibility residuals are of order $T^{-1/4}$ in the number T of iterations.

Research was supported by the Regional Mathematical Center of the Southern Federal University with the Agreement № 075-02-2022-893 of the Ministry of Science and Higher Education of Russia.

Theorem 1 *For the average transfer price vector* $\overline{\lambda}_T = \frac{1}{T}\sum_{t=1}^{T}\lambda_t$ *we have*

$$F(z^*) - F(\widetilde{z}(\overline{\lambda}_T)) \le \frac{C}{T^{1/4}}, \quad \|\Delta\widetilde{z}(\overline{\lambda}_T)\| \le \frac{C}{T^{1/4}},$$

where the constant C *depends of* f_i, g_i, X_i, Y_i, $i = 1, \ldots, d$.

Similar results were obtained for a dynamic problem, where the functions f_i g_i depend on a sequence of i.i.d. random variables.

The leader also can be selfish. Consider, for example, the product revenue management problem with unknown demand. The main difference with the previous case is that the leader objective function need not be convex, and its gradient is unknown. To overcome the first difficulty we use price discretization and its probabilistic interpretation. Then we use zero-order algorithms for price tuning.

REFERENCES

1. *Rokhlin D.B.* Resource allocation in communication networks with large number of users: the dual stochastic gradient method *Theory of Probability & Its Applications*, **66**:1, (2021), 105–120.
2. *Rokhlin D.B., Ougolnitsky G.A.* SOLO FTRL algorithm for production management with transfer prices *Journal of Mathematical Sciences* (to appear)
3. *Orabona F., Pál D.* Scale-free online learning *Theoretical Computer Science*, **716**, (2018), 50–69.