The Sample Complexity of Online Contract Design

BANGHUA ZHU and STEPHEN BATES, Department of EECS, UC Berkeley, USA
ZHUORAN YANG, Department of Statistics and Data Science, Yale University, USA
YIXIN WANG, Department of Statistics, University of Michigan, USA
JIANTAO JIAO and MICHAEL I. JORDAN, Department of EECS, UC Berkeley, USA

Contract theory studies the interactions between a principal and an agent when the two parties transact in the presence of private information [Bolton and Dewatripont, 2004, Faure-Grimaud et al., 2001, Salanié, 2005]. The principal would like to achieve her desired outcomes by hiring agents to work for her. The agent wishes to make money by working for the principal. They develop agreements in the form of a contract, which specifies how much the principal would pay under the different possible outcomes of the agent’s work.

We study the hidden-action principal-agent contract design problem in an online setting. In each round, the principal posts a contract that specifies the payment to the agent based on each outcome. The agent then makes a strategic choice of action that maximizes her own utility, but the action is not directly observable by the principal. The principal observes the outcome and receives utility from the agent’s choice of action. Based on past observations, the principal dynamically adjusts the contracts with the goal of maximizing her utility. We treat this problem as a continuum-armed bandit problem, where we think of each potential contract as an arm.

We introduce an online learning algorithm and provide an upper bound on its Stackelberg regret. In particular, we show that when the contract space is $[0, 1]^m$, the Stackelberg regret is upper bounded by $\tilde{O}(\sqrt{m \cdot T^{1-1/(2m+1)}})$, and lower bounded by $\Omega(T^{1-1/(m+2)})$, where $\tilde{O}$ omits logarithmic factors. This result shows that exponential-in-$m$ samples are both sufficient and necessary to learn a near-optimal contract, resolving an open problem in Ho et al. [2016] on the hardness of online contract design. Moreover, when contracts are restricted to some subset $\mathcal{F} \subset [0, 1]^m$, we define an intrinsic dimension of $\mathcal{F}$ that depends on the covering number of the spherical code in the space and bound the regret in terms of this intrinsic dimension. When $\mathcal{F}$ is the family of linear contracts, we show that the Stackelberg regret grows exactly as $\Theta(T^{2/3})$.

Technically, the contract design problem is challenging because the utility function is discontinuous. Indeed, bounding the discretization error in this setting has been an open problem [Ho et al., 2016] since simple uniform discretization does not work for the contract design problem. Although the utility function is a piecewise linear function of the contract, it is not continuous because the best-response mapping from the agent is not continuous. Thus the discretization error can be arbitrarily large for the naive uniform discretization.

To address this challenge, we design a novel discretization scheme for the continuous contract space. We identify a limited set of directions in which the utility function is continuous, allowing us to design a new discretization method and bound its error. Our discretization uses coding theory, specifically we exploit the maximum packing of a spherical code [Kabatiansky and Levenshtein, 1978, Manin and Marcilli, 2019]. This leads to a near-optimal regret bound for both general contracts and linear contracts, which is the first upper bound with no restrictions on the contract and action space, resolving several open problems from prior work. The techniques we introduce may extend more generally to other Stackelberg games and continuum-armed bandits.

A full version of this paper can be found at https://arxiv.org/abs/2211.05732.