Dynamic Pricing with Finitely Many Unknown Valuations

Tommaso CESARI, Université degli Studi di Milano.

Mots-clés : Posted prices Auctions, Bandits, Learning

We consider revenue maximization in stochastic dynamic pricing where the distribution of buyers' private values is supported on an unknown set of points in $[0,1]$ of unknown cardinality K. This choice of setting is well-suited to model scenarios in which buyers are grouped in an unknown number of latent types, characterized by their private values for the good on sale. As a concrete example, consider the buyer's game in repeated first price-auctions. This is formally equivalent to the seller's game in dynamic pricing, where valuations correspond to the highest bids among other bidders. In online advertising, other bidders are Demand-Side Platforms that bid according to their own user segmentation. Then, their higher bids will be values corresponding to the union of these unknown segments.

We show that no algorithm can achieve a regret significantly better than $\widetilde{\Theta}(\sqrt{KT})$, where T is the time horizon, and we present an algorithm achieving regret $\widetilde{\mathcal{O}}(\sqrt{KT})$ up to a constant independent of the distribution of buyers' valuations.

Our setting is strictly harder than K-armed stochastic bandits: we show that it is impossible to prove regret bounds of order $\log T$ even when the gap Δ between the revenue of the optimal valuation and that of the second-best valuation is constant. We then investigate distribution-dependent bounds that rely on additional information on the demand curve. We prove an upper bound of order $(1/\Delta + (\log\log T)/\gamma^2)K\log T$, where γ is a known lower bound on the smallest probability of a valuation (i.e., the smallest drop in the demand curve). As $(K/\Delta)\log T$ is the regret in K-armed stochastic bandits, this shows that the price for identifying each one of the K valuations is at most $(\log T)(\log\log T)/\gamma^2$, which corresponds (up to log log factors) to the known upper bound on noisy binary search. This is a significant improvement on previously known regret bounds for discontinuous demand curves, that are at best of order $(K^{12}/\gamma^8)\sqrt{T}$. We also prove that when $K = 2$, a regret bound of order $\log(T)/\Delta$ (up to log log factors) can be achieved with no additional information, which is the same bound that one would have if all the valuations were known in advance.

Finally, we show a $\mathcal{O}(\sqrt{T})$ upper bound on the regret for the setting in which the buyers' decisions are nonstochastic, and the regret is measured with respect to the best between two fixed valuations one of which is known to the seller.

Références