LEGO: Optimal Online Learning under Sequential Price Competition

We consider price competition among multiple sellers over a selling horizon of T periods. In each period, sellers simultaneously offer their prices and subsequently observe their respective demand that is unobservable to competitors. The realized demand of each seller depends on the prices of all sellers following a private unknown linear model. We propose a least-squares estimation then gradient optimization (LEGO) policy, which does not require sellers to communicate demand information or coordinate price experiments throughout the selling horizon. We show that our policy, when employed by all sellers, leads at a fast convergence rate $O\left(\frac{1}{\sqrt{T}}\right)$ to the Nash equilibrium prices that sellers would reach if they were fully informed. Meanwhile, each seller achieves an optimal order-of-$\sqrt{T}$ regret relative to a dynamic benchmark policy. Our analysis further shows that the unknown individual price sensitivity contributes to the major difficulty of dynamic pricing in sequential competition and forces regret to the order of $\sqrt{T}$ in the worst case. If each seller knows their individual price sensitivity coefficient, then a gradient optimization policy can achieve an optimal order-of-$\frac{1}{T}$ convergence rate to Nash equilibrium as well as an optimal order-of-$\log T$ regret.