

Adaptive maximization of social welfare

Abstract

We consider the problem of repeatedly choosing policy parameters, such as tax or transfer rates, in order to maximize social welfare, defined as the weighted sum of private utility and public revenue. The outcomes of earlier policy choices inform later choices. In contrast to multi-armed bandit models, utility is not observed, but needs to be indirectly inferred from the integral of the response function. In contrast to standard optimal tax theory, response functions need to be learned through policy choices. We derive a lower bound on regret for this problem, and a matching adversarial upper bound on regret for a variant of the Exp3 algorithm. In both cases, cumulative regret grows at a rate of $T^{2/3}$. This implies that (i) the social welfare maximization problem is harder than the multi-armed bandit problem (with a rate of $T^{1/2}$), and (ii) that our proposed algorithm achieves the optimal rate. Simulations confirm these results, as well as the viability of the proposed algorithm. We also compare the social welfare maximization problem to two related learning problems, monopoly pricing (which is easier), and price setting for bilateral trade (which is harder). We lastly discuss extensions to nonlinear income taxation, and to commodity taxation.

This paper is joint work with Nicolò Cesa-Bianchi and Roberto Colomboni (Dipartimento di Informatica, Università degli Studi di Milano).

Maximilian Kasy

Professor of Economics at the University of Oxford