Learning to Dynamically Optimize from Individual Experience

Nathan Palmer

February 9, 2016

Abstract

Sims (1980) famously stated that non-rational behavior is a "wilderness:" there is only one way for economic agents to be "rational" but infinitely many ways to be non-rational. While this may be strictly true, some paths through the wilderness are more familiar and well-trod than others. This paper introduces a framework for implementing bounded rationality in the canonical household intertemporal consumption-savings problem, which converges to rational behavior in the limit of agent experience. I expand upon a growing economic literature that uses tools from reinforcement learning and approximate dynamic programming to impose bounded rationality in intertemporal choice problems. My work contributes to the literature by applying these tools to the canonical household portfolio choice under uncertainty problem. This problem is central to macroeconomics and macro-finance, including consumption-based asset pricing. I employ a novel learning-to-optimize algorithm which functions at the individual level: agents use experience to approximate their value function, and then use this value function to find a regret-minimizing policy function. This approach is new in the literature; it is closest to Q-learning but avoids the difficulties associated with applying Q-learning to continuous-valued choice problems. Numerical results demonstrate convergence to a stable neighborhood around the optimal policy. Importantly, usage of the canonical optimal control framework allows this model to be directly compared to optimal solutions. Methods of estimation, a key future extension, are discussed as well.

JEL Classification: C61, C63, D14, D83, D91, E21