
DAVID SAUNDERS, University of Waterloo

Exploratory Investment-Consumption with Non-Exponential Discounting

We extend the classic Merton optimal investment-consumption problem to the reinforcement learning (RL) framework. Additionally, we incorporate a general non-exponential discounting function to capture an investor's risk preferences, which leads to time inconsistency in the exploratory control problem. Under entropy regularization and logarithmic utility, we obtain closed-form equilibrium investment-consumption policies. Specifically, the optimal investment policy follows a Gaussian distribution, while the optimal consumption policy follows a Gamma distribution. Our results show that uncertainty about the discount rate leads the investor to adopt more conservative policies, with the Gaussian-distributed investment policy retaining the same mean but lower variance, and the Gamma-distributed consumption policy having both a lower mean and variance. We further develop and implement two RL algorithms- one based on the policy evaluation approach and the other on the q-learning approach- demonstrating their effectiveness through simulation studies. This is joint work with Y. Chen and Y. Li from the University of Waterloo.