Thompson Sampling with Discrete Prior

Wei Zheng
Department of Business Analytics and Statistics, University of Tennessee, USA

Abstract

Thompson sampling is a popular algorithm for multi-armed bandit problems, but its Bayesian posterior update can be computationally expensive for complex reward distributions. Recently, prior discretization has been proposed to address this issue. In this paper, we propose a new prior discretization method that guarantees the same regret rate without requiring the unreasonable assumption that the true value of the parameter is one of the discrete points. Additionally, we introduce a modified posterior update approach that further improves the performance of discrete prior Thompson sampling. We prove that the accumulated regret has $O(\log(T))$ convergence rate with high probability. In addition, we conduct numerical experiments to validate our theoretical analysis and demonstrate that the proposed algorithm outperforms both the standard discrete prior method and the Laplace approximation approach for the continuous prior.