Recursive adaptation of stepsize parameter for non-stationary environments.

Noda, Itsuki


Summary: In this article, we propose a method to adapt stepsize parameters used in reinforcement learning for non-stationary environments. In general reinforcement learning situations, a stepsize parameter is decreased to zero during learning, because the environment is generally supposed to be noisy but stationary, such that the true expected rewards are fixed. On the other hand, we assume that in the real world, the true expected reward changes over time and hence, the learning agent must adapt the change through continuous learning. We derive the higher-order derivatives of exponential moving average (which is used to estimate the expected values of states or actions in major reinforcement learning methods) using stepsize parameters. We also illustrate a mechanism to calculate these derivatives in a recursive manner. Using the mechanism, we construct a precise and flexible adaptation method for the stepsize parameter in order to optimize a certain criterion, for example, to minimize square errors. The proposed method is validated both theoretically and experimentally.

doi:10.1007/978-3-642-11814-2_5