A numerical example in Markov decision process with interval-valued transition probability

Shunya Matsumori[a], Mitsuhiro Hoshino[b], Yousuke Araya[c]
Graduate School of Systems Science and Technology,
Akita Prefectural University, 015-0055, Japan[a]
Faculty of Systems Science and Technology,
Akita Prefectural University, 015-0055, Japan[b,c]

Abstract

In the standard Markov decision process, we treat the case that the transition probability of the state varies in some given domain at each time and its variation is unknown or unobservable. In order to analyze such a case, Kurano et al has introduced a new decision model, called a controlled Markov set-chain, based on Markov set-chains, and discussed the optimization problem of the discounted expected rewards under some set-order relations.

In this presentation, we will give a new numerical example which illustrates an average optimal policy maximizes the long-run expected average reward per unit time under some set-order relations.