Optimal control of continuous–time Markov chains

Onésimo Hernández–Lerma
Departamento de Matemáticas
CINVESTAV–IPN

* Joint work with X.P. Guo (Zhongshan University, Guangzhou, P.R. China)

Abstract. In the control of continuous–time Markov chains (also known as continuous–time Markov decision processes) one distinguishes two cases depending on whether the control actions can be taken (1) only at certain (usually random) epochs or (2) continuously in time. In the former case, (1), the problem can be reduced to a discrete–time Markov decision process for which there is an extensive literature.

This work concerns case (2), for continuous–time controlled Markov chains with possibly unbounded transition rates, and unbounded reward rates that are allowed to take positive or negative values (as opposed to be bounded either from below or from above). We give conditions for the existence of $\varepsilon$–optimal policies ($\varepsilon \geq 0$) for a discounted reward criterion, as well as a “martingale characterization” of optimal stationary policies. We also discuss how to extend to our context the well known “vanishing discount approach” to analyze the average reward case.