

Rothblum, Uriel G.

Multiplicative Markov decision chains. (English) Zbl 0535.90097
Math. Oper. Res. 9, 6-24 (1984).

This paper considers multiplicative Markov decision chains in which the transition matrices are nonnegative, but not necessarily substochastic, and the associated reward gained by the end of period $N + 1$ is the product of the one-period transition matrices in periods $1, \dots, N$ times a positive terminal reward.

Previous treatments of this model restricted attention to stationary policies and assumed that all transition matrices are irreducible and aperiodic. They also used a "first term" optimality criterion, namely maximizing the spectral radius of the associated transition matrix.

In this paper a constructive proof is given for the existence of optimal policies among all policies under new cumulative average optimality criteria which are more sensitive than the maximization of the spectral radius. The algorithm for finding an optimal policy, first searches for a stationary policy with a nonnilpotent transition matrix, provided such a rule exists. Otherwise, the method still finds an optimal policy; though in this case the set of optimal policies usually does not contain a stationary policy! If a stationary policy with a nonnilpotent transition matrix exists, then a policy improvement algorithm which finds a stationary optimal policy is followed.

MSC:

90C40 Markov and semi-Markov decision processes
49L20 Dynamic programming in optimal control and differential games
90C39 Dynamic programming

Cited in 14 Documents

Keywords:

[multiplicative Markov decision chains](#); [constructive proof](#); [existence of optimal policies](#); [nonnilpotent transition matrix](#); [policy improvement algorithm](#)

Full Text: [DOI](#)