On Markov games with the expected average reward criterion (II)

By

Kensuke TANAKA and Kazuyoshi WAKUTA

(Received September 15, 1975)

1. Introduction

This paper is a continuation of our paper with the same title [2] in which we have showed, mainly, that under some assumptions the Markov games with the criterion of long-run average reward has a value and both players have optimal strategies.

Here, we shall show that the solution $(d, u(\cdot))$ of the functional equation (3. 1) of assumption 4 in [2] can be solved by the method of successive approximations under certain conditions. We can obtain this method as an application of the method introduced by D. J. White in Dynamic Programming [3].

2. Preliminaries

In order to state the method of successive approximations, we assume the same conditions as those in [2], that is, (1) S, A and B are compact metric spaces, (2) $r=r(s, a, b)$ is a continuous function on $S\times A\times B$, (3) whenever $s_n\longrightarrow s_0$, $a_n\longrightarrow a_0$ and $b_n\longrightarrow b_0$, $q(\cdot\,|\,s_n, a_n, b_n)$ converges weakly to $q(\cdot\,|\,s_0, a_0, b_0)$, (4) there exist a continuous function $u(s)$ on S and a constant d such that for each $s\in S$,

$$d+u(s)=\sup_{\lambda\in P_A}\inf_{\mu\in P_B}\{r(s, \mu, \lambda)+\int u(s')\,dq(s'\,|\,s, \mu, \lambda)\}, \tag{2. 1}$$

where P_A and P_B are the sets of all probability measures on $(A, \mathcal{B}(A))$ and $(B, \mathcal{B}(B))$, respectively,

$$r(s, \mu, \lambda)=\iint r(s, a, b)\,d\mu(a)\,d\lambda(b),$$

and

$$q(E\,|\,s, \mu, \lambda)=\iint q(E\,|\,s, a, b)\,d\mu(a)\,d\lambda(b).$$

Moreover, we need the same lemmas as those in [2], that is,