Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 117 (2007) 1357–1371
stochastic processes and their applications
www.elsevier.com/locate/spa

On Gittins' index theorem in continuous time

Peter Bank[a,*], Christian Küchler[b]

[a] Department of Mathematics, Columbia University, 2990 Broadway, Mail Code 4433, New York, NY 10027, USA
[b] Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany

Received 12 June 2006; received in revised form 10 January 2007; accepted 22 January 2007
Available online 25 January 2007

Abstract

We give a new and comparably short proof of Gittins' index theorem for dynamic allocation problems of the multi-armed bandit type in continuous time under minimal assumptions. This proof gives a complete characterization of optimal allocation strategies as those policies which follow the current leader among the Gittins indices while ensuring that a Gittins index is at an all-time low whenever the associated project is not worked on exclusively. The main tool is a representation property of Gittins index processes which allows us to show that these processes can be chosen to be pathwise lower semi-continuous from the right and quasi-lower semi-continuous from the left. Both regularity properties turn out to be crucial for our characterization and the construction of optimal allocation policies.

MSC: 60G07; 60G40; 60H25; 91B16; 91B28

Keywords: Gittins index; Multi-armed bandits; Representation theorem

0. Introduction

Multi-armed bandit problems arise when a limited amount of resources, usually thought of as working time, has to be allocated to a number of different projects. These offer dynamically varying, random rewards when worked on, and so one has to trade-off those projects which currently offer the highest rewards against those projects which possibly will yield even better payoffs after they have been worked on for some time.

* Corresponding author. Tel.: +1 2128545545; fax: +1 2128548962.
E-mail addresses: pbank@math.columbia.edu (P. Bank), ckuechler@math.hu-berlin.de (C. Küchler).
URLs: http://www.math.columbia.edu/~pbank (P. Bank), http://www.math.hu-berlin.de/~ckuechler (C. Küchler).

doi:10.1016/j.spa.2007.01.006

Typically, these problems are rather difficult (if not even impossible) to solve due to their high dimensionality and the complex interdependencies between the projects. For independent projects, however, it was shown first by Gittins and Jones [6] that there exists a project-specific dynamic performance measure, later called the Gittins index of a project, such that optimal allocations are obtained from an index policy which (essentially) amounts to focussing at each point only on those projects which exhibit a maximal Gittins index. This celebrated result was subsequently extended from Gittins' and Jones' original discrete-time, Markovian framework to a completely general continuous-time setting; see, e.g., Whittle [15], Varaiya, Walrand and Buyukkoc [13], Mandelbaum [10], Weber [14], El Karoui and Karatzas [3,4], Kaspi and Mandelbaum [8,9].

These accounts are based on two equivalent definitions of the Gittins index. First, there is the definition as an indifference threshold with respect to early retirement of a project like, e.g., in Whittle [15]. Also El Karoui and Karatzas [3,4] use this for their martingale approach in continuous time. Second, there is Gittins' forward characterization of his index as the properly discounted maximal expected future rewards per expected time used for obtaining these. In continuous time, this characterization is the starting point for the excursion theoretic approach pursued by Kaspi and Mandelbaum [8,9].

By contrast, our approach to Gittins' index theorem is based on a characteristic representation property which relates the Gittins index to the accumulated future expected rewards from a given project; see Corollary 2.1. This property, in conjunction with a novel partial-integration argument, not only allows for a comparably short proof of Gittins' index theorem under minimal assumptions, see Theorem 2, it actually also allows us to give necessary and sufficient conditions for optimality, a result already conjectured by Kaspi and Mandelbaum [9]. Another important point in our analysis is the lower semi-continuity of Gittins index processes. This new regularity result is crucial not only for the continuous-time construction of Gittins indices as proper optional processes, but it actually turns out to be indispensable even for the existence of optimal allocation strategies. Our Theorem 1 shows how this regularity follows from the representation property of Gittins indices. This result also complements the representation theorem of Bank and El Karoui [1].

1. Problem formulation

We shall follow Mandelbaum [10] in order to formalize our allocation problem as a multi-parameter control problem. To this end, let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space on which we have N real-valued processes $h^p = (h^p_s,\ s \geq 0)$ satisfying

$$\mathbb{E} \int_0^\infty \mathrm{e}^{-\alpha s} |h^p_s| \mathrm{d}s < \infty \quad (p = 1, \ldots, N).$$

We will interpret h^p_s as the reward rate from project p after we have spent already $s \geq 0$ units of time on this particular project. When working on a project we typically will learn more about its future prospects. This gain of information is modeled by a project-specific filtration $\mathbb{F}^p = (\mathscr{F}^p_s)_{s \geq 0}$ which satisfies the usual conditions of right-continuity and completeness. We assume each h^p is progressively measurable with respect to the filtration $\mathbb{F}^p$ associated with the corresponding project.

An allocation strategy is given by a collection $\mathbf{S}$ of increasing processes $(S^p_t,\ t \geq 0)$ $(p = 1, \ldots, N)$ specifying how much time should be spent on each project up to calendar time t.