BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 34, Number 2, April 1997, Pages 205–206
S 0273-0979(97)00708-8

Discrete-time Markov control processes: Basic optimality criteria, by O. Hernández-Lerma and J. B. Lassere, Appl. Math., vol. 30, Springer-Verlag, Berlin and New York, 1996, xiv+216 pp., $34.95, ISBN 0-387-94579-2

Markov control processes (also called Markov decision processes) arise in stochastic optimization models drawn from a wide variety of applications in engineering, economics, management science, biology and medicine. Only discrete-time models are considered in this book, thus avoiding various difficult technical issues encountered in continuous-time stochastic control theory. Let x_t denote the state at time $t = 0, 1, 2, \cdots$ and a_t an action (or control) at time t. The actions influence the stochastic dynamics of the state process x_t, via the one-step transition probabilities $Q(x_{t+1} \in B | x_t, a_t)$ for any Borel subset B of the state space X. Alternatively, the dynamics are often expressed through a difference equation of the form

$$x_{t+1} = F(x_t, a_t, \xi_t) \tag{1}$$

with $\{\xi_t\}$ an IID sequence of exogenous random inputs. The goal is to choose a control sequence $\{a_t\}$ to optimize some performance criterion J on either a finite or infinite time horizon. On an infinite horizon a discounted cost criterion

$$J = E\left[\sum_{t=0}^{\infty} \alpha^t c(x_t, a_t)\right] \tag{2}$$

can be considered, where $c(\cdot, \cdot)$ is a running cost function and $0 < \alpha < 1$ a discount factor. Another frequently used criterion is average cost per unit time, which arises naturally by taking a limit $\alpha \to 1^-$. The information available to the controller must also be specified. This book is concerned with complete state information, in which a_t can be chosen as a function of x_t, namely $a_t = f_t(x_t)$. The sequence $\{f_t\}$ is a Markov control policy, which is stationary if $f_t = f$ does not depend on t.

The method of dynamic programming is very often used to study Markov control problems. A formal description of dynamic programming considers the value function, which is the minimum $v(x)$ of the criterion J considered as a function of the initial state $x_0 = x$. Under suitable assumptions the value function satisfies a nonlinear equation (the dynamic programming equation), which for the criterion (2) becomes

$$v(x) = \min_{A(x)} \left[c(x, a) + \alpha \int_X v(y) Q(dy|x, a)\right], \tag{3}$$

with $A(x)$ the set of possible control actions a if x is the state. To determine an optimal stationary Markov control policy f, arg min is taken on the right side of (3) for each $x \in X$. However, a measurable selection theorem is generally needed to insure Borel measurability of f.

A main goal of the book is to put dynamic programming on a mathematically rigorous basis in a general setting which avoids unwanted boundedness, compactness or continuity assumptions which may be violated in applications. However, by avoiding utmost generality the authors have succeeded in making the book accessible to

1991 *Mathematics Subject Classification.* Primary 90C40, 93E20.

readers with a graduate-level background in analysis and a basic understanding of probability. The authors call their model *semicontinuous - semicompact*. It lies in generality between the so-called *semicontinuous* model considered in *Controlled Markov Processes*, by E. B. Dynkin and A. A. Yushkevich, Springer-Verlag, 1979, and the more mathematically demanding *Borel model* [see e.g. M. Schäl and W. Sudderth, *Stationary policies and Markov policies in Borel dynamic programming*, Probab. Theory Rel. Fields **74** (1987), 91–111].

An alternative to dynamic programming is an infinite-dimensional linear programming formulation of Markov control problems. This approach is nicely introduced in the final chapter. The linear programming formulation has the advantage that it can handle Markov control problems with constraints.

The book is well written and provides a good entree to the subject for nonexperts. It is self-contained, except for various technical results which are summarized with references (for example, measurable selection theorems.) The mathematical developments are illustrated by a few examples, including consumption-investment and inventory-production models. A broader perspective on the role of Markov control processes in applications can be found in the following references: D. B. Bertsekas, *Dynamic Programming*: *Deterministic and Stochastic Models*, Prentice-Hall, 1987; D. J. White, *A survey of applications of Markov decision processes*, J. Opl. Res. Soc. **44** (1993), 1073–1096; P. Whittle, *Optimal Control: Basics and Beyond*, John Wiley and Sons, 1996.

Wendell H. Fleming
Brown University
E-mail address: whf@cfm.brown.edu