ON DYNAMIC PROGRAMMING FOR MULTISTAGE DECISION PROBLEMS UNDER UNCERTAINTY

Paolo Dai Pra, Cristina Rudari and Wolfgang J. Runggaldier, Padova

Summary : We investigate the possibility of applying the Dynamic Programming (DP) algorithm to obtain an optimal solution for a class of multistage decision problems under uncertainty. This class corresponds to the discrete time control problems, where the uncertainty is due to disturbances in the evolution of the state process. The disturbances are assumed to belong to random sets, which is related to the so-called Mathematical Theory of Hints.

1.Introduction

A class of multistage decision problems under uncertainty is the class of discrete time control problems under uncertainty. The characteristic feature of these problems is that they concern a system whose state $x_k \in X_k$ evolves according to a recursive relation of the form

$$x_{k+1} = f_k(x_k, a_k, w_k)\,; \qquad k = 0, 1, \ldots \tag{1}$$

where $a_k \in C_k$ is the decision/control in period k and $w_k \in D_k$ is the value in period k of a disturbance that, together with x_0, accounts for the uncertainty in the system. The disturbances are usually unobserved. On the other hand in this paper we assume that, for each period k, the value of the state x_k is being observed; one may however consider also more general models allowing for a "partial observation" of the state. The decision variables $a_0, a_1, \cdots$, that are chosen in each period k on the basis of the current history of observed states and controls, are evaluated according to a cost criterion that, in the finite horizon version, takes the form

$$G(x^N, a^{N-1}) = \sum_{k=0}^{N-1} g_k(x_k, a_k) + g_N(x_N) \tag{2}$$

where $x^N := (x_0, \cdots, x_N)$ is a realization of the system states according to (1).

The peculiarity of the given class of control problems with respect to general multistage decision problems is (see also [VW]) that the decisions/controls affect the cost not only directly, but also indirectly via the state, the "scenarios" being determined by the disturbances.

A crucial feature in formulating and solving control problems under uncertainty is the modeling of the disturbances. The case mostly studied is when the disturbances form a stochastic process (the "stochastic control" case) and then the optimality criterion is the minimization of the expected value of the cost (2).

In some cases the a-priori informations do not allow any reasonable assumptions on the statistics of the disturbances. In such a context one chooses an optimality criterion that guarantees robustness of the optimal control : optimization is under the assumption that the disturbances behave in the worst possible way. This leads to a min-max type of problem, reminiscent of dynamic games.

In this paper we show that the stochastic and "robust" theories are limit cases of a more general description of a control problem under uncertainty. It consists in assuming that the disturbances take values in certain random sets with a-priori known distributions. If therefore the random sets consist of a single point, we recover the stochastic case; if however these sets are deterministically given, we are back to the "robust" case.

U. Derigs et al. (eds.), *Operations Research Proceedings 1994*
© Springer-Verlag Berlin Heidelberg 1995

In our general setting there are various choices of optimality criteria that in the stochastic or "robust" limits behave consistently with the two limit theories. While they both appear to be meaningful, except for some special cases, only one of them allows for a solution approach via Dynamic Programming (DP).

The idea of modeling uncertainty via random sets is inherent in the so-called Mathematical Theory of Hints (see [K1]), which is related to the Mathematical Theory of Evidence (see [S]) and it can be shown (see [DPRR]) that dynamical Hints, as described in [K2], are equivalent to the model we consider in this paper. What we add here is the formulation and study of optimal control problems in the above framework.

2. The model

To model the disturbances in (1), let $(\Omega, \mathcal{A}, P)$ be a probability space and, for $k = 0, \cdots, N$, let $(\Omega_k, \mathcal{A}_k)$ be measurable spaces and

$$\begin{aligned} \eta_0 &: \Omega \longrightarrow \Omega_0 \\ \eta_k &: \Omega \times X_{k-1} \times C_{k-1} \longrightarrow \Omega_k \end{aligned} \tag{3}$$

be measurable functions. For fixed $x_0, \cdots, x_{k-1}, a_0, \cdots, a_{k-1}$, let $\eta_k(\cdot, x_{k-1}, a_{k-1})$ be independent of $\eta_0(\cdot), \cdots, \eta_{k-1}(\cdot, x_{k-2}, a_{k-2})$. Finally, let

$$A_0 \; : \; \Omega_0 \longrightarrow \mathcal{P}(X_0)\backslash\emptyset \quad , \qquad A_k \; : \; \Omega_k \longrightarrow \mathcal{P}(D_k)\backslash\emptyset \quad (k > 0) \tag{4}$$

be set-valued functions. The uncertainty in the system is due to the fact that we assume

$$x_0 \in A_0(\eta_0(\omega)) \, ; \; w_k \in A_{k+1}(\eta_{k+1}(\omega, x_k, a_k)) \quad (0 \le k < N) \tag{5}$$

Optimality criterion A)

Assuming that a_k is a function of the current history of states and controls is equivalent to assuming that $a_k = \mu_k(x^k)$ for some measurable function μ_k. Due to (1), the cost $G(x^N, a^{N-1})$ in (2) is completely determined once x_0 and $\{w_k\}_{k=0,\cdots,N-1}$ are specified. Given a control policy $u = \{\mu_k(\cdot)\}_{k=0}^{N-1}$, we say that u is A-admissible if the function of $\omega \in \Omega$

$$\sup_{x_0 \in A_o \, ; \, w_k \in A_{k+1}} G(x^N, \mu^{N-1}(\cdot)) \tag{6}$$

is a P-integrable r.v. Randomness in (6) comes from A_k. We denote by U_A the set of *A-admissible controls.*

Definition 2.1. The functional $J_A \; : \; U_A \longrightarrow \mathcal{R}$

$$J_A(u) = E \left\{ \sup_{x_0 \in A_o \, ; \, w_k \in A_{k+1}} G(x^N, \mu^{N-1}(\cdot)) \right\} \tag{7}$$

is called *Optimality criterion A*. The associated optimal control problem consists in finding $J_A^* = \inf_{u \in U_A} J_A(u)$ and, whenever it exists, a control policy $u^* \in U_A$ such that $J_A(u^*) = J_A^*$. The control u^* is then called *A-optimal.*

Optimality criterion B)

The control $u = \{\mu_k(\cdot)\}_{k=0}^{N-1}$ is said to be B-admissible if for each $k = 0, \cdots, N-1$

$$\sup_{w_k \in A_{k+1}} E \cdots \sup_{w_{N-2} \in A_{N-1}} E \sup_{w_{N-1} \in A_N} G(x^N, \mu^{N-1}) \tag{8}$$