Gamkrelitze's and Boltyanskii's results appear to entail the same degree of difficulty in computing the optimal parameter vector in the general nonlinear case $\hat{f}(t,\hat{x},w)$. However, if $\hat{f}$ is linear with respect to w and of the form

$$\hat{f}(t,\hat{x},w) = \hat{h}_0(t,\hat{x})w + \hat{h}_1(t,\hat{x}) \tag{7}$$

where $\hat{h}_0$ is an $(n+1) \times m$ matrix valued function and $\hat{h}_1$ is an $(n+1)$ vector-valued function, Gamkrelitze's results have a definite advantage over Boltyanskii's. Indeed, Boltyanskii's conditions (3) or (4) in the linear case (7) result in a set of implicit equations for w^*, while Gamkrelitze's maximum principle (2) may give an explicit form for it. This is shown in the following proposition.

Proposition 2: Let Ω be the closed unit hypercube in R^m and consider the system

$$\dot{x}(t) = h_0(t,x)w + h_1(t,x), \qquad t \in I, \tag{8}$$

with $w \in \Omega$. If w^* is an optimal parameter vector in Ω, transferring the state of (8) from $x_0 \in R^n$ to $x_1 \in R^n$ in minimum time, then w^* (if it exists) will be of the form

$$w^* = \operatorname{sgn}\left\{\int_{t_0}^{t_1} h_0{}^T(t,x^*)\psi(t)\,dt\right\} \tag{9}$$

where ψ is a nontrivial solution of the adjoint system corresponding to (8), and system (8) is assumed to be *normal* in the sense that

$$\int_{t_0}^{t_1} h_0{}^T(t,x^*)\psi(t)\,dt \neq 0.$$

Proof: The proof follows from Proposition 1.

Remark: The two-point boundary-value problem that results from substitution of (9) into (8) and its adjoint appears to be of the anticipative type! However, the optimal parameter can be found by solving 2^m equations of type (8), one for each of the vertices of the hypercube.

REFERENCES

[1] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*. New York: Wiley, 1962.

[2] V. G. Boltyanskii, *Mathematical Methods of Optimal Control*. New York: Holt, Rinehart and Winston, 1971.

[3] R. V. Gamkrelitze, "On some extremal problems in the theory of differential equations with applications to the theory of optimal control," *SIAM J. Contr.*, ser. A, vol. 3, no. 1, pp. 106–128, 1965.

[4] E. Hofer and P. Sagirow, "Optimal systems depending on parameters," *AIAA J.*, vol. 6, no. 5, pp. 953–956, 1968.

Linear Convex Stochastic Control Problems Over an Infinite Horizon

DIMITRI P. BERTSEKAS

***Abstract*—A stochastic control problem over an infinite horizon which involves a linear system and a convex cost functional is analyzed. We prove the convergence of the dynamic programming algorithm associated with the problem, and we show the existence of a stationary Borel measurable optimal control law. The approach used illustrates how results on infinite time reachability [1] can be used for the analysis of dynamic programming algorithms over an infinite horizon subject to state constraints.**

I. Problem Formulation

Consider a constant linear discrete system

$$x_{k+1} = Ax_k + Bu_k + w_k, \qquad k = 0,1,\cdots \tag{1}$$

Manuscript received November 1, 1972; revised December 21, 1972. This work was supported by the NSF under Grant GK 29237.

The author is with the Department of Engineering-Economic Systems, Stanford University, Stanford, Calif. 94305.

where $x_k \in E^n$ and $u_k \in E^m$ denote for all k the state and control of the system, w_k denotes a stochastic disturbance vector, and the matrices A and B are given. The problem that we are concerned with is finding a feedback control law $\{\mu_0,\mu_1,\cdots\}$, $u_k = \mu_k(x_k)$, which minimizes the cost functional

$$J(x_0,\mu_0,\mu_1,\ldots) = \lim_{N\to\infty} \underset{\substack{w_k \\ 0\le k\le N}}{E}\left\{\sum_{k=1}^{N} a^k g[x_k,\mu_k(x_k)]\right\} \tag{2}$$

where a is a scalar discount factor $0 < a < 1$, g is a convex function of (x,u), and $E\{\cdot\}$ denotes the expectation operator.

Such problems are often encountered in practice [4]–[6] and, in lack of special structure, they must be solved numerically by dynamic programming (DP). A question of practical importance is whether the DP algorithm converges to a steady-state solution. It is also of interest to examine the question of existence of an optimal control law. A recent interesting paper [4] has considered the problem in a framework which required somewhat restrictive assumptions in order to assert existence. The aim of this note is to offer an alternative problem formulation which is natural, suitable to a DP solution, and which provides a framework within which both the convergence and the existence questions can be answered in the affirmative.

In the formulation that we consider we assume that the random vector w_k takes values for all k in a Borel subset W of E^n and is characterized by a probability measure P defined on the σ-algebra of Borel subsets of W. The class of admissible control laws consists of all sequences $\{\mu_0,\mu_1,\cdots\}$ of Borel measurable functions μ_k defined on a set $X \subset E^n$ such that $u_k = \mu_k(x) \in U(x) \subset E^n$ for all k and $x \in X$. The set X is a state constraint set and the set $U(x)$ is a state dependent control constraint set. It is assumed that the sets X and $M = \{(x,u) | u \in U(x), x \in X\}$ are convex and compact. It is further demanded that the control law $\{\mu_0,\mu_1,\cdots\}$ be such that the state x_k of the closed-loop system $x_{k+1} = Ax_k + B\mu_k(x_k) + w_k$ belongs to X for all k and for all values $w_k \in W$; i.e., it is demanded that the control law $\{\mu_0,\mu_1,\cdots\}$ achieve infinite time reachability of the set X according to the terminology of [1]. In the next section we prove the existence of a stationary control law within the admissible class which minimizes the cost functional (2), and we show the uniform convergence of the associated DP algorithm.

II. The Dynamic Programming Algorithm

Prior to the formulation of the DP algorithm it is necessary that the issues associated with the domain of definition of this algorithm be clarified. The presence of the state constraint $x \in X$ complicates matters since given any point $x_k \in X$ there may not exist any control $u_k \in U(x_k)$ such that the state of the system at the next stage $x_{k+1} = Ax_k + Bu_k + w_k$ is guaranteed to belong to X for all values $w_k \in W$. Thus the state-space region of infinite time feasibility $R^*(X)$ (the set of initial states starting from which there exists an admissible control law resulting in satisfaction of all the constraints of the problem) may be (and usually will be) only a proper subset of the set X. Since the region of feasibility $R^*(X)$ is the domain of definition of the DP algorithm, it is necessary to demonstrate that this region can be characterized effectively.

Using the results of [1] and [2], we have that for the problem of this note

$$R^*(X) = \bigcap_{n=1}^{\infty} R^n(x) \tag{3}$$

where the function R^n mapping subsets of E^n into subsets of E^n is the composition $R \cdot R \cdot \cdots \cdot R(n \text{ times})$ of the function R defined by

$$R(Z) = \{x | \exists u \in U(x) \ \text{ s.t. } \ Ax + Bu + W \subset Z\} \cap Z. \tag{4}$$

It is shown in [1] and [2] that $R^n(X)$ converges to the convex, compact set $R^*(X)$ in a well-defined sense. Furthermore, it is shown in [1] and [2] that the class Ω of all admissible control laws which achieve infinite time reachability of the set X is the class of control laws $\{\mu_0,\mu_1,\cdots\}$ for which μ_i is Borel measurable and

$$[x,\mu_i(x)] \in \bigcap_{n=1}^{\infty} C_n(X), \qquad \forall x \in R^*(X) \tag{5}$$

where the convex and compact sets $C_n(X)$ are defined by

$$C_n(X) = \{(x,u) | x \in R^{n-1}(X),\ u \in U(x),\ Ax + Bu + W \subset R^{n-1}(X)\}. \tag{6}$$

It has been shown also [1], [2] that, for the problem of this note, $R^*(X)$ is the projection on the state space of the set $\bigcap_{n=1}^{\infty} C_n(X)$, i.e.,

$$R^*(X) = P_x\left[\bigcap_{n=1}^{\infty} C_n(X)\right].$$

We assume that $\bigcap_{n=1}^{\infty} C_n(X)$ is nonempty. We can prove now that the class Ω of admissible control laws is nonempty. To show this, consider the multivalued mapping $A: R^*(X) \rightarrow E^n$ defined by

$$A(x) = \{u | (x,u) \in \bigcap_{n=1}^{\infty} C_n(X)\}, \tag{7}$$

the graph of which is $\bigcap_{n=1}^{\infty} C_n(X)$. Now for every closed set $S \in E^m$ the set

$$A^{-1}(S) = \{x \in R^*(X) | A(x) \cap S \neq \phi\} = P_x\left[\bigcap_{n=1}^{\infty} C_n(X) \cap (R^*(X) \times S)\right]$$

is compact by the compactness of $\bigcap_{n=1}^{\infty} C_n(X)$. Hence the mapping A is Borel measurable according to the definition of [7]. By using a theorem of Kuratowski and Rull–Nardzewski (see [7, corollary 1.1]) it follows that there exists a Borel measurable function $\mu: R^*(X) \rightarrow E^n$ such that $[x,\mu(x)] \in \bigcap_{n=1}^{\infty} C_n(X)$ for all $x \in R^*(X)$. Hence the stationary control law $\{\mu,\mu,\cdots\}$ is Borel measurable and achieves reachability of X. The conclusion is that the class Ω of admissible control laws is nonempty and that, by using any control law in this class, the state of the closed-loop system will belong to the set $R^*(X)$ for all times. Hence any DP algorithm used for solving the problem need only be defined over the set $R^*(X)$.

Consider now the cost functional (2). We have, by the continuity of g,

$$r_1 \leq g(x,u) \leq r_2, \qquad \forall (x,u) \in \bigcap_{n=1}^{\infty} C_n(X)$$

for some scalars r_1 and r_2. It follows that, for every fixed control law in Ω, the series in (2) converges absolutely and the limit indicated exists.

Furthermore, it can be easily shown that

$$J_\infty(x_0) = \inf_{\{\mu_0,\mu_1,\cdots\}\in\Omega} J(x_0,\mu_0,\mu_1,\cdots) = \lim_{N\rightarrow\infty} J_N(x_0)$$

where $\lim_{\rightarrow\infty} J_N(x_0)$ exists for every $x_0 \in R^*(X)$, and where the function $J_N: R^*(X) \rightarrow R$ is defined recursively by

$$J_0(x) = \min_{u\in A(x)} g(x,u) \tag{8}$$

$$J_{k+1}(x) = \min_{u\in A(x)} [g(x,u) + a\mathop{E}_{w}\{J_k(Ax + Bu + w)\}], \qquad k = 0,1,\cdots \tag{9}$$

where $A(x)$ is the compact set defined for every x by (7).

In order to show the existence of a Borel measurable control law, we prove the following lemma.

Lemma: Let $R \subset E^n$ be a nonempty convex and compact set and $A: R \rightarrow E^m$ be a multivalued mapping with $A(x) \neq \emptyset$, for all $x \in R$, such that its graph $C = \{(x,u) \in E^n \times E^m | x \in R,\ u \in A(x)\}$ is convex and compact. Also let $h: C \rightarrow E$ be a lower semicontinuous convex function. Then the function

$$H(x) = \min_{u\in A(x)} h(x,u)$$

is convex and lower semicontinuous over R. Furthermore, there exists a Borel measurable function $\mu: R \rightarrow E^m$ with $\mu(x) \in A(x)$, for all $x \in R$, such that

$$H(x) = h[x,\mu(x)]. \tag{10}$$

Proof: By direct use of Theorem 9.2 in [8], H can be shown to be convex and lower semicontinuous. Now consider the epigraph of h

$$\text{epi}\ (h) = \{(x,u,\alpha) | (x,u) \in C, \quad \alpha \in E, \quad h(x,u) \leq \alpha\}.$$

The set epi (h) is a closed convex set, and it is also the graph of the multivalued epigraph mapping $K: R \rightarrow E^{m+1}$

$$K(x) = \{(u,\alpha) | (x,u) \in C, \quad \alpha \in E, \quad h(x,u) \leq \alpha\}.$$

For any compact set $S \subset E^{m+1}$ we have that

$$K^{-1}(S) = \{x \in R | K(x) \cap S \neq \emptyset\} = P_x[\text{epi}\ (h) \cap (R \times S)]$$

is compact by the compactness of $(R \times S)$. Hence, K is Borel measurable and $h(x,u) + \delta(x,u|C)$ (where $\delta(\cdot|C)$ is the indicator function of C) is a normal convex integrand on $R \times E^m$, [7, theorem 4]. The existence of the Borel measurable function μ satisfying (10) now follows from Corollary 4.3 of [7]. Q.E.D.

By using the lemma, we have that the functions J_k of (8) and (9) are convex and lower semicontinuous over $R^*(X)$. Furthermore, it can be easily shown by using an argument similar to the one in [4] that the pointwise limit J_∞ of the sequence $\{J_k\}$ is a convex and lower semicontinuous function over $R^*(X)$ satisfying

$$J_\infty(x) = \min_{u\in A(x)} [g(x,u) + a\mathop{E}_{w}\{J_\infty(Ax + Bu + w)\}]. \tag{11}$$

In addition, the convergence is uniform on any compact subset of the relative interior of $R^*(X)$ [8, theorem 10.8]. The existence of a stationary Borel measurable optimal control law follows by using the lemma in conjunction with (11).

The actual solution of the problem can be effected in two ways. One method is to use the dynamic programming algorithm (8), (9) and simultaneously determine the region of k-step reachability $R^k(X)$ which is the domain of the function J_k. Since $R^k(X)$ converges to $R^*(X)$ and J_k converges to J_∞ on $R^*(X)$, this procedure will determine both J_∞ and its domain $R^*(X)$. The other method is to determine $R^*(X)$ first, and then to use a generalization of Howard's policy iteration algorithm [3] for finding J_∞. Which of the two methods is more efficient should depend on the particular problem at hand.

Acknowledgment

The author wishes to thank Prof. Paul R. Kleindorfer for his helpful comments.

References

[1] D. P. Bertsekas, "Infinite-time reachability of state-space regions by using feedback control," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 604–613, Oct. 1972.
[2] ——, "Convergence of the feasible region in infinite horizon optimization problems," in *Proc. 1972 Joint Automatic Control Conf.* (Stanford Univ., Stanford, Calif.), 1972.
[3] A. Kaufmann and R. Cruon, *Dynamic Programming.* New York: Academic, 1967.
[4] P. R. Kleindorfer and K. Glover, "Linear convex stochastic optimal control with applications in production planning," in *Proc. 1971 IEEE Decision and Control Conf.* (Miami, Fla.), 1971; also *IEEE Trans. Automat. Contr.* (Short Papers), vol. AC-18, pp. 56–59, Feb. 1973.
[5] M. J. Sobel, "Production smoothing with stochastic demand I: Finite horizon case," *Management Sci.*, vol. 16, Nov. 1969.
[6] ——, "Production smoothing with stochastic demand II: Infinite horizon case," *Management Sci.*, vol. 17, July 1971.
[7] R. T. Rockafellar, "Measurable dependence of convex sets and functions on parameters," *J. Math. Anal. Appl.*, vol. 28, pp. 4–25, 1969.
[8] ——, *Convex Analysis.* Princeton, N.J.: Princeton Univ. Press, 1970.