ON THE SOLUTION OF SOME MINIMAX PROBLEMS

Dimitri P. Bertsekas
Engineering-Economic Systems Department
Stanford University, Stanford, Calif.

Abstract

In dynamic minimax and stochastic optimization problems frequently one is forced to use a suboptimal controller since the computation and implementation of the optimal controller based on dynamic programming is impractical in many cases. In this paper we study the performance of some suboptimal controllers in relation to the performance of the optimal feedback controller and the optimal open-loop controller. Attention is focused on some classes of, so called, open-loop-feedback controllers. It is shown under quite general assumptions that these open-loop-feedback controllers perform at least as well as the optimal open-loop controller. The results are developed for general minimax problems with perfect and imperfect state information. In the latter case the open-loop-feedback controller makes use of an estimator which is required to perform at least as well as a pure predictor in order for the results to hold. Some of the results presented have stochastic counterparts.

1. Introduction

Since the dynamic programming approach towards the optimization of dynamic uncertain systems is often computationally impractical, suboptimal controllers for such systems are in common usage. Such controllers include the optimal open-loop controller, the naive feedback controller, and the open-loop-feedback controller. The precise definition of each of these controllers is not as yet standard in the current literature. For this reason we shall define each of them in relation to a specific minimax problem which will be of continuing interest in this paper.

Problem 1: Given is the uncertain dynamic system

\[ x_{k+1} = f_k(x_k, u_k, w_k) \quad k = 0, 1, \ldots, N - 1 \]  

where \( x_k \) and \( u_k \) denote for all \( k \) the state and control of the system and \( w_k \) denotes some uncertain parameter. The quantities \( x_k, u_k \) and \( w_k \) are elements of spaces \( S_{x_k}, S_{u_k}, S_{w_k} \), respectively, and the functions \( f_k = f_k(x_k, u_k, w_k) \) are known. It is assumed that, for each \( k \), the control \( u_k \) is constrained to take values from a given subset \( U_k \) of \( S_{u_k} \). It is also assumed that the disturbance \( w_k \) can take values from a given subset \( W_k \) of \( S_{w_k} \). Given the initial state of the system \( x_0 \) find (if it exists) a control law \( \{u_0, u_1, \ldots, u_{N-1}\} \) with

\[ u_k = k = 0, 1, \ldots, N - 1 \]  

which minimizes the cost functional

\[ J(x_0, u_0, u_1, \ldots, u_{N-1}) = \sup_{w_k \in W_k, k=0,1,\ldots,N-1} g(x_N) \]  

subject to the system equation constraints and where \( g = S_{x_N} \) is a given real valued function.

It is worth noting that the above problem formulation is very general since the state, control and disturbance spaces are arbitrary and may differ from one time instant to another. Among other things this allows one to reduce a wide variety of cost functionals to the terminal state cost functional of Eq (2) by means of various state augmentation and transformation techniques.

We shall denote by \( J_f(x_0) \) the optimal value of the cost functional in Problem 1. This value corresponds to the optimal feedback (o.f.) controller. This controller makes optimal use of the information obtained during the operation of the system, namely the value of the state at each time. The controller which is optimal (assuming it exists) in the class of admissible controllers which ignore the information, i.e., the class of controllers for which

\[ u_k = \text{constant}, \forall x_k \in S_{x_k}, k = 0, 1, \ldots, N - 1, \]  

is called the optimal open-loop \( \xi \) (o.l.) controller and the corresponding value of the cost functional is denoted by \( J_0(x_0) \).

Following Witsenhausen [1] we shall call any admissible controller \( \{u_0, u_1, \ldots, u_{N-1}\} \) quasi-adaptive if

\[ J_f(x_0) \leq J(x_0, u_0, u_1, \ldots, u_{N-1}) \leq J_0(x_0) \]  

and adaptive if

\[ J_f(x_0) \leq J(x_0, u_0, u_1, \ldots, u_{N-1}) < J_0(x_0) \]  

In a practical situation the computation and implementation of the optimal feedback controller is often impractical while the more easily implementable optimal o.l. controller may perform rather poorly. Thus suboptimal adaptive controllers which can be practically implemented are of interest. A suboptimal controller often used in practice is the, so called, naive feedback controller. In order to calculate this controller the disturbance vectors are assumed to have some fixed value \( \bar{w}_k \) for each time \( k \) with \( \bar{w}_k \in W_k \) and the feedback controller which is optimal for the resulting "deterministic" problem is used. This controller need not be calculated by dynamic programming but rather can be implemented by solving an open-loop "deterministic" optimization problem at each time \( k \) starting at the observed state \( x_k \).

Contrary to deep-rooted convictions among engineers it is known [2] that in general the naive feedback controller may perform strictly worse than the optimal o.l. controller, i.e. it may not be quasi-adaptive.

The open-loop-feedback (o.l.f.) controller [1] which is the main object of study of this paper, is similar to the naive feedback controller except that it takes uncertainty explicitly into account. It is denoted by \( \{u_0, u_1, \ldots, u_{N-1}\} \) and defined as follows:

At any time \( k \) and state \( x_k \) let

\[ \{u_k, u_{k+1}, \ldots, u_{N-1}\} \]  

be the sequence which minimizes (assuming it exists) the cost functional

\[ J_f(x_{k+1}) = \sup_{w_k \in W_k} g(x_{N+1}) \]  

subject to the system equation constraints and where \( g = S_{x_N} \) is a given real valued function.

The author is with the Dept. of Engineering-Economic Systems, Stanford University, Stanford, Calif.
among all sequences \(\{u_k, u_{k+1}, \ldots, u_{N-1}\}\) such that
\[ u_i \in U_i, i=k, \ldots, N-1. \]
The value of the open-loop controller at state \(x_k\) is given by
\[ J_0^C(x_k) = u_k. \]
Clearly, the open-loop controller is easier to implement than the optimal feedback controller and it is more difficult to implement than the optimal open-loop controller (since the optimal open-loop controller is calculated already at the first stage of implementation of the open-loop controller). It is a general belief that
\[ J_k(x) \text{ is quasi-adaptive.} \]

This fact, apparently not proven in the literature even for the stochastic case, is demonstrated in Section 2. In the same section we also consider a problem similar to Problem 1 where in addition there are state constraints. We next consider the case where the controller has imperfect state information. Under these circumstances the open-loop controller makes use of an estimator which calculates at each time the exact set of possible states or a set which bounds the set of possible states (presumably this bounding set can be calculated more easily, as is the case, for example, of linear systems with ellipsoidal constraints [3],[4],[5]). It is shown in Section 3 that if the estimator used performs, roughly speaking, better than a pure predictor, then the resulting open-loop controller is quasi-adaptive.

2. Performance of the Open-Loop-Feedback Controller for the Perfect State Information Case

Let us assume that the open-loop controller \(\{\hat{u}_0, \hat{u}_1, \ldots, \hat{u}_{N-1}\}\) exists for Problem 1 and let the corresponding value of the cost functional be denoted by \(J_0^C(x_0)\). Then
\[ J_0^C(x_0) = J_0(x_0) \]
where the function \(J_0: S_{X_0} \to [0, \infty]\) is given recursively by the algorithm
\[ J_{N-1}(x) = \sup_{w \in \tilde{W}_{N-1}} \min \{g[f_{N-1}(x, \hat{u}_{N-1}(x), w)] \} \]
(3)
\[ J_k(x) = \sup_{w \in \tilde{W}_k} \min_{k=0,1,\ldots,N-2} J_{k+1}(x) \]
(4)
\[ J_k(x) = \sup_{w \in \tilde{W}_k} \min_{k=0,1,\ldots,N-2} J_{k+1}(x) \]
(5)

Let us also consider the functions
\[ J_k^C: S_{X_k} \to [0, \infty], k=0,1,\ldots,N-1, \text{ defined by} \]
\[ J_k^C(x) = \sup_{u_i \in U_i, i=k,\ldots,N-1} g(x_N) \]
(6)

The minimization problem indicated in the above equation must be solved at time \(k\) and state \(x\) in order to calculate the open-loop optimal cost. Thus if \(\{\hat{u}(x), \hat{u}_{k+1}(x), \ldots, \hat{u}_{N-1}(x)\}\) solves the problem in Equ. (6) we have \(J_k^C(x) = \hat{u}_k(x)\). Clearly, \(J_k^C(x)\) can be interpreted as the calculated open-loop optimal cost from time \(k\) to time \(N\) and starting from the state \(x\). The optimal open-loop cost is given by
\[ J_0^C(x_0) = J_0^C(x_0) \]

We prove the following proposition:

**Proposition 1:** For every \(x\) and \(k\) we have
\[ J_k(x) \leq J_k^C(x) \]
(7)

In particular for \(k=0\) we have
\[ J_0^C(x_0) = J_0^C(x_0) \]
(8) and hence the open-loop-feedback controller is quasi-adaptive.

**Proof:** We shall prove (7) by induction. Since by the definition of the open-loop controller we have
\[ J_{k+1}(x) = J_{k+1}^C(x), \]
(9)
then for all \(x \in S_{X_k}\) we have
\[ J_{k+1}[f_k(x, \hat{u}_k(x), w)] \leq J_{k+1}^C[f_k(x, \hat{u}_k(x), w)] \]
and
\[ J_k(x) = \sup_{w \in \tilde{W}_k} J_{k+1}^C[f_k(x, \hat{u}_k(x), w)] \]
(10)

It is to be noted that by (7) the calculated open-loop optimal cost from state \(x_k\), \(J_k^C(x_k)\), provides a readily obtainable performance bound for the open-loop controller. Two generalizations of the above proposition should also be noted. The first is concerned with the case where the control constraints are state dependent and the disturbance constraints are state and control dependent. In this case the result of Proposition 1 follows by using a similar proof. The second generalization concerns the stochastic version of Problem 1 where the disturbances \(W_k\) are random vectors with given probability distribution and the sup in the cost functional (2) is replaced by an expectation. Again a result analogous to the one of Proposition 1 follows by using an entirely similar argument.

The case where in Problem 1 there are additional state constraints \(x_k \in X_k\) where \(X_k\) is a given subset of \(S_{X_k}\) requires separate discussion since one can envisage two different ways of defining the open-loop controller. The first method is similar to the one used earlier whereby at state \(x_k\) the open-loop controller solves the problem
\[ \min_{u_i \in U_i, i=k,\ldots,N-1} g(x_N) \]
(11)
subject to the additional constraint on the sequence \(\{u_k, u_{k+1}, \ldots, u_{N-1}\}\) such that \(x_i \in X_i, i=k,\ldots,N-1\) for all
The control constraints are \( u_0, u_1 \) on problem (9). Problem 1 (without state constraints) namely that if the control constraint sets are enlarged it is possible that the performance of the o.l.f. controller deteriorates. This is due, of course, to the fact that the o.l.f. controller does not have access to the exact system state but rather receives (possible noise-corrupted) measurements providing information about the system state. We consider the following problem.

Problem 2: Consider Problem 1 where the controller has access to measurements of the form

\[
z_k = h_k(x_k, v_k) \quad k=0,1,...,N-1
\]

where \( v_k \) is an uncertain "disturbance" known to belong to a given subset \( V_k \) of a space \( S_k \) and the function \( h_k: S_k \times X_k \rightarrow S_k \) is given. Find (if it exists) a control law \( \{u_0, u_1, ..., u_{N-1}\} \) with \( u_k: S_k \times S_{k-1} \times ... \times S_0 \rightarrow S_0 \times ... \times S_0 \) such that \( u_k = \mu_k(z_0, ..., z_k, u_0, u_1, ..., u_{k-1}) \), \( k=0,1,...,N-1 \) which minimizes the cost functional

\[
J(x_0, u_0, u_1, ..., u_{N-1}) = \sup_{w_k \in W_k} g(z_N) \quad w_k \in W_k \quad v_k \in V_k \quad k=0,1,...,N-1
\]

Given a particular information sequence \( z_k = (z_0, z_1, ..., z_k, u_0, u_1, ..., u_{k-1}) \) there exists a corresponding set of possible states \( x_k \) given that \( z_k \) has occurred. We denote this set by \( X_k(z_k) \). It is the set of all states \( x_k \) consistent with the measurements \( z_0, ..., z_k \), the controls \( u_0, ..., u_{k-1} \), the system equation and the constraints \( w_i \in W_i, i=0, ..., k-1 \), \( v_i \in V_i, i=0, ..., k \). It is known that the set \( X_k(z_k) \) constitutes a sufficient information [1], [9], [10] for the controller relative to Problem 2, i.e. the optimal controller need only be a function of \( X_k(z_k) \) rather than \( z_k \). In a sense that can be well defined the set \( X_k(z_k) \) can be viewed as the state of a new system. This system evolves in time according to the equations

\[
X_{k+1}(z_{k+1}) = f_k[X_k(z_k), u_k, w_k, v_{k+1}] \quad k=0,1,...,N-2
\]

\[
X_N(z_{N-1}) = f_N[X_{N-1}(z_{N-1}), u_{N-1}, w_{N-1}]
\]

where

\[
f(2,0) = f(0,-1) = 0
\]

\[
f(2,-1) = f(0,0) = 2
\]

\[
f(1,0) = f(1,-1) = f(-1,0) = f(-1,-1) = 0.5
\]

The control constraints are \( u_0, u_1 \in \{0, 1\} \) and the state constraint is \( x_2 \in (0, 0.5) \). The cost functional is

\[
J(x_0, u_0, u_1) = \max_{w, x_2} x_2 \quad w \in \{0,-1\}
\]

It follows by straightforward calculation that the cost corresponding to the o.l.f. controller based on problem (9) is 0 while the cost corresponding to the o.l.f. controller based on problem (10) is 0.5.

The above example demonstrates also a counterintuitive property of the o.l.f. controller for Problem 1 (without state constraints) namely that if the control constraint sets are enlarged it is possible that the performance of the o.l.f. controller deteriorates. This is due, of course, to the fact that the o.l.f. controller does not have access to the exact system state but rather receives (possible noise-corrupted) measurements providing information about the system state. We consider the following problem.

3. Performance of the Open-Loop-Feedback Controller for the Imperfect Information Case

We turn now to the case where the o.l.f. controller
controller for Problem 2' is exactly what we shall call the o.l.f. controller for Problem 2. Notice that this o.l.f. controller can be realized by solving at time $k$ the problem

$$\min_{u_i \in U_i} \sup_{x_i \in X_k(\hat{x}_k)} g(x_N)$$

$$u_i \in U_i \quad x_i \in X_k(\hat{x}_k)$$

$$i = k, \ldots, N-1$$

$$w_i \in W_i$$

$$x_i+1 = f_1(x_i, u_i, w_i)$$

$$i = k, \ldots, N-1$$

and by taking as the current control the first element of the minimizing sequence. Given that the optimal open-loop controller for Problem 2' is the same as the optimal open-loop controller for Problem 2 and using Proposition 1 we have:

**Proposition 2:** Let $V_{o.f.}(x_0)$, and $V_{o.l.}(x_0)$ be the values of the cost functional (12) corresponding to the o.l.f. controller and the optimal o.l controller respectively. Then

$$V_{o.f.}(x_0) \leq V_{o.l.}(x_0)$$

While the calculation of the set $X_k(\hat{x}_k)$ by the controller is possible in principle, in practice this calculation can be very difficult or impossible, i.e. it may be difficult to construct a realization of the corresponding estimator given by Equ.(13) through (16). For this reason it is of interest to examine the performance of open-loop feedback controllers based on estimators that can be more easily implemented. We shall consider a class of estimators which we shall call recursive bounding estimators. Such estimators provide estimate sets $\hat{X}_k(\hat{x}_k)$ which contain the set of possible states $X_k(\hat{x}_k)$ and are realized by a recursive algorithm of the general form

$$\hat{X}_k(\hat{x}_k) = E_k[\hat{X}_{k+1}(\hat{x}_{k+1})]$$

$$\hat{X}_0(\hat{x}_0) = \{x_0\}$$

(21)

where $E_k$ is some function. Examples of such recursive bounding estimators are the estimators of [3], [4], [5]. We shall say that a recursive bounding estimator is uncertainty reducing if for each admissible $x_k, u_k, w_k, v_k, \Sigma_k$, we have

$$\hat{X}_k(\hat{x}_k) \subset E_k[\hat{X}_{k+1}(\hat{x}_{k+1})]$$

Thus the estimator is uncertainty reducing if it provides a set $\hat{X}_k(\hat{x}_k)$ which is contained in the set which would be obtained by pure prediction given $X_k(\hat{x}_k), u_k$. In other words an uncertainty reducing estimator uses the new measurement $z_k$ with advantage.

We can define now an o.l.f. controller using a recursive bounding estimator (21) as follows. Given the set $X_k(\hat{x}_k)$ solve the problem

$$\min_{u_i \in U_i} \sup_{x_i \in X_k(\hat{x}_k)} g(x_N)$$

$$u_i \in U_i \quad x_i \in X_k(\hat{x}_k)$$

$$i = k, \ldots, N-1$$

$$w_i \in W_i$$

$$x_i+1 = f_1(x_i, u_i, w_i)$$

$$i = k, \ldots, N-1$$

and takes as the current control the first element of the minimizing sequence.

We have the following proposition:

**Proposition 3:** Let $J_{o.f.}(x_0)$ and $J_{o.l.}(x_0)$ be the values of the cost functional (12) corresponding to an o.l.f controller using a recursive bounding estimator (21), and to the optimal o.l controller respectively. If the estimator is uncertainty reducing we have

$$J_{o.f.}(x_0) \leq J_{o.l.}(x_0)$$

Proof: Let $\{\Sigma_0, \Sigma_1, \ldots, \Sigma_{k+1}\}$, $u_k = \sigma_k(\hat{x}_k)$ be the o.l.f. controller. Then $J_{o.f.}(x_0) = v_0(x_0)$ where the function $J_0$ is given recursively by the algorithm

$$J_N = \max_{x \in X_N(\hat{x}_N)} \{g(x_N)\}$$

and

$$J_{k+1} = \max_{x \in X_{k+1}(\hat{x}_{k+1})} \{g(x_{k+1})\}$$

Consider also (c.f. (6)) the calculated cost

$$J_{o.l.}(x_0) = \min_{u \in U} \sup_{x \in X} g(x_N)$$

By using the fact that the estimator is bounding we have that

$$J_{o.l.}(x_0) = J_{o.l.}(x_0)$$

Thus in order to prove the proposition it will be sufficient to prove that

$$J_{k} = J_{o.f.}(x_0) \leq J_{o.l.}(x_0)$$

for all $k$ and all admissible $\Sigma_k$.

By using the fact that the estimator is bounding we have that

$$J_{o.l.}(x_0) = J_{o.l.}(x_0)$$

where $J_{o.l.}(x_0)$ is defined for all $k$ by

$$J_{o.l.}(x_0) = \sup_{x \in X_{k+1}(\hat{x}_{k+1})} \{g(x_{k+1})\}$$

By using the fact that the estimator is uncertainty reducing it follows easily that

$$J_{o.l.}(x_0) \leq J_{o.l.}(x_0)$$

We have $J_{o.l.}(x_0) = J_{o.l.}(x_0)$. Thus in order to prove the proposition it will be sufficient to prove that

$$J_{k} = J_{o.f.}(x_0) \leq J_{o.l.}(x_0)$$

for all $k$ and all admissible $\Sigma_k$.

By using the fact that the estimator is bounding we have that

$$J_{k} = J_{o.l.}(x_0) \leq J_{o.l.}(x_0)$$

where $J_{o.l.}(x_0)$ is defined for all $k$ by

$$J_{o.l.}(x_0) = \sup_{x \in X_{k+1}(\hat{x}_{k+1})} \{g(x_{k+1})\}$$

By using the fact that the estimator is uncertainty reducing it follows easily that

$$J_{o.l.}(x_0) \leq J_{o.l.}(x_0)$$

We have $J_{o.l.}(x_0) = J_{o.l.}(x_0)$.
for all \( k, r \). Hence (23) holds and the proposition is proved.

An easily obtained generalization of the above proposition concerns the possibility of proving stochastic counterparts to Propositions 2 and 3. The stochastic counterpart of Prop. 2 can be proved similarly with no difficulty. The role of the set \( X_k(\mathcal{C}_k) \) is played by the conditional probability \( p(X_k, \mathcal{C}_k) \). However for stochastic problems it is not clear how one is to define the analog of a recursive bounding and uncertainty reducing estimator except for some special cases. One such special case is the well known linear quadratic Gaussian problem. For this case linear estimators will produce Gaussian state estimates which can be partially ordered by means of their error covariance matrix. Thus an uncertainty reducing linear estimator is one for which the corresponding error covariance matrix is smaller (in the pos. definite sense) than the error covariance matrix corresponding to pure prediction. Using this definition a proposition similar to Proposition 3 can be proved.

4. Conclusions

In this paper it was shown under general assumptions that open-loop-feedback controllers perform at least as well as optimal open-loop controllers in dynamic minimax problems. The classes of problems considered include the perfect state information case and the imperfect state information case. In the latter case the open-loop-feedback makes use of an estimator computing either the exact set of possible states of the system or an estimate set that bounds the set of possible states. The estimator is required to perform better than a pure predictor in order for the results to hold. Some of the results in this paper can also be proved in a stochastic control framework.

5. References