Annals of Economic and Social Measurement, 6/3, 1977

LOW PROFILE ECONOMIC POLICY WITH GUARANTEED RETURN

By D. D. Sworder*

When the system to be controlled is stochastic, it is usually impossible to select a control policy which maximizes the utility function for all possible system realizations. There do exist algorithms for synthesizing regulators with maximum mean utility and these are often used even when other measures of system performance are more natural. Exclusive concern with mean performance often masks the inherent variability found in the response characteristics of stochastic systems.

This paper takes a different approach and for a restricted class of random parameter linear systems provides a control policy which minimizes the worst possible cost of operation. If a system must operate satisfactorily in a variety of different modes, such a policy gives guaranteed performance without regard to the actual evolution of the system dynamics.

1. Introduction

The consequences of a specific economic program are often difficult to predict when the program is initiated. It is common knowledge that selecting a policy based on the continuation of extant conditions may yield untoward results when there are exogenous changes in the dynamic structure of the system. Through failure to modify policy to match system environment, conditions that the analyst seeks to ameliorate may instead be made worse.

One way in which the uncertainty that surrounds econometric models may be made quantitative is to employ a stochastic model of the economic entity of interest. The sample functions of the model parameters are selected to display the salient features of the uncertainties actually encountered in depicting the system. A difficulty immediately encountered with this approach is that the utility function, which is sample function dependent, no longer serves to order the decision policies. To pose a meaningful problem, expected value of the utility is customarily used performance index. Although this may yield analytically tractable algorithms for finding the best policy, such an approach suffers from the philosophical objection that concern actually centers on the performance attained during a single period of evolution of the system rather than the average performance over an ensemble of possible evolutions. In some cases a more conservative, worst-case, approach to policy synthesis is thus appropriate.

*This research was supported by the National Science Foundation under Grant GK-39817.

The use of "linear-quadratic" synthesis techniques creates subtle difficulties of a different sort. Policy derived using a quadratic utility function depends parametrically on the weights used in the performance index. Choice of these weights is often subjectively based and particularly the weights associated with control are hard to justify. Commonly, the weights are made proportional to one over the square of the maximum permissible value of the associated control input. Although this technique enables the analyst to deduce a policy which yields acceptable levels of performance with permissible values of control, in many cases the political realities surrounding the use of the control variables may produce a desire for less active policy when the performance degradation is not too severe.

In this paper a linear-quadratic regulation problem is considered in which the system model is subject to a stochastic exogenous influence. A control policy is derived which minimizes the worst possible cost of system operation. Of the many policies which provide minimax risk, that one is chosen which requires least control effort.

2. Problem Description

The system to be controlled will be assumed to be adequately described by a linear, differential equation

$$\dot{x} = Ax + Bv \qquad t_0 \leq t \leq \dot{T} \tag{1}$$
$$x(t_0) = x_0.$$

Equation (1) typically represents the equation of evolution of the deviations from a desired trajectory of a nonlinear system of equations which delineate the behavior of the dynamic economic model. It will be assumed that the analyst seeks to maintain x and v as close to zero as possible. The specific measure of closeness will be quadratic. Define

$$J = \int_{t_0}^{T} (x'\tilde{Q}x + v'\tilde{R}v)\, dt. \tag{2}$$

The best feedback policy will be that choice of $v(t)$ as a function of $(x(t), t)$ which minimizes J.

It may happen that the analyst is uncertain about how the system dynamics will vary with time. For example, if (1) represents the perturbation variables of a macroeconomic model of the national economy, the state of the economy and hence the matrix $[A, B]$, cannot be known with certainty for times very far into the future. Selection of the endogenous variables must be made with an appropriate discounting of future events of uncertain occurrence. Suppose

$$[A(t), B(t)] = [A_i, B_i] \quad \text{if} \quad r(t) = i; \quad i = 1, \ldots, N$$

where $r(t)$ is an indicator variable describing the dynamic state of the system given in (1). To express the initial uncertainty about the behavior of the process, r will be assumed to be a random process. In fact it will be assumed that r is a Markov jump process with finite state space and transition matrix Q;

$$\text{(3)} \qquad \text{Prob}(r(t + \Delta) = j \mid r(t) = i) = \begin{cases} 1 + q_{ii}\Delta + 0(\Delta) & i = j \\ q_{ij}\Delta + 0(\Delta) & i \neq j \end{cases}$$

If the elements of the Q matrix are constant, the residence times of r in each state are exponentially distributed.

The stochastic model complicates the selection of good policy. The observation will be assumed to be $(x(t), r(t), t)$. The selection of the endogenous variable at time t is contingent on the value of r at time t, but only probabilistic information is available on how r will behave in the interval $[t, T]$. Let r be defined on the probability space Ω. Then the cost function given in (2) is a random variable, $J(\omega)$.

Since no feedback policy will minimize J uniformly, an analyst must often content himself with minimizing $E(J)$. While this may be a rational course of action in some situations, such a policy does lead to sudden changes in the closed loop dynamic behavior of the system. If $r(t)$ is in a favorable dynamic state in (1), the closed loop performance is very good. If on the contrary $r(t)$ is such that (1) is difficult to control, the closed loop system may have very poor dynamic response. This is particularly true when the expected lifetime in the disadvantageous state is expected to be short. Since the performance index attaches importance to only the mean of J, there is no explicit control over how big J may be.

There is another more subtle difficulty with the mean optimal policy. The first term in (2) weights state deviation and the matrix $\tilde{Q}$ can often be justified by the importance accorded to deviations in different state variables. The positive matrix $\tilde{R}$, on the other hand, weights the size of the endogenous variable. One common choice is

$$\tilde{R}(t) = \text{diag}(V_{1,m}^{-2}, \ldots, V_{k,m}^{-2})$$

where $V_{i,m}$ is the maximum permissible magnitude of the i'th endogenous variable. Such a choice is very subjective. It encourages the over use of control through the endogenous variable in those dynamic modes not requiring such external control to provide satisfactory performance. Political considerations often make it advantageous to use as small a value for v as dynamic conditions permit. Generation of such a "low profile" policy is the aim of this paper.

It will be assumed that there is a worst dynamic state of operation of (1). To be precise, it will be assumed one of the dynamic states $w \in [1, \ldots, N]$ is such that

(4) $$(A_w - A_i)'K + K(A_w - A_i) \quad + K[B_i\tilde{R}^{-1}B_i' - B_w\tilde{R}^{-1}B_w']K \geq 0$$
$$i = 1, \ldots, N$$

for all nonnegative symmetric K. As will become apparent later, the matrix $[A_w, B_w]$ is least favorable state of system operation.

Equation (4) appears to be rather technical, but it has a natural interpretation in some special cases.

(4a) (a) If $A_w = A_i$, condition (4) $\Longrightarrow B_i\tilde{R}^{-1}B_i' \geq B_w\tilde{R}^{-1}B_w'$

(4b) (b) If $B_w = B_i$, condition (4) $\Longrightarrow A_w \geq A_i$

If the two systems have identical A matrices, the less favorable system has a smaller gain matrix (see 4a). If the gain matrices are the same, the less favorable system has poles uniformly to the right of those of the more favorable system (4b). It will be assumed that Q is such that all states of r lead to w.

The basic problem can now be delineated. A policy v^* is sought which minimizes the maximum cost; i.e.

(5) $$J(v^*, \omega^*) = \min_{v} \max_{\omega} J(v, \omega)$$

The minimum is taken over all decision rules, and the maximum is over all sample functions or r.

3. Solution Algorithm

The method of solution to this problem uses a well known result of decision theory which states that if v^* is a decision rule such that

(6) $$J(v^*, \omega) = C \text{ for all } \omega \in \Omega$$

and v^* is a Bayes rule with respect to some probability measure on Ω, then v^* is minimax [1]. The measure with respect to which v^* is Bayes is termed least favorable, and a decision rule satisfying (6) is called an equalizer.

An equalizer decision policy will first be constructed. Let K_w be the solution of the matrix differential equation

(7) $$\dot{K}_w = -A_w'K_w - K_wA_w - Q + K_wB_w\tilde{R}^{-1}B_w'K_w$$
$$K_w(T) = 0$$

It is well known that K_w is the cost matrix associated with the optimally controlled deterministic system satisfying $r \equiv w$.

Suppose a decision rule of the form

$$v(t) = -\tilde{R}^{-1}B_i'F_i x \quad \text{if } r(t) = i \tag{8}$$

is used in this problem. Direct calculation leads to the conclusion that

$$E\left\{\int_t^T (x'\tilde{Q}x + v'\tilde{R}v)\,dt \mid x(t) = x, r(t) = r\right\} = x'\overline{K}_r(t)x \tag{9}$$

where

$$\begin{aligned} \dot{\overline{K}}_i(t) &= -(A_i - B_i\tilde{R}^{-1}B_i'F_i)'\overline{K}_i - \overline{K}_i(A_i - B_i\tilde{R}B_i'F_i) - Q \\ &\quad -F_i'B_i\tilde{R}^{-1}B_i'F_i + \sum_j q_{ij}K_j; \end{aligned} \tag{10}$$

$$\overline{K}_i(T) = 0 \quad i = 1, \ldots, N$$

Let the $\{F_i\}$ be selected in such a way that

$$\overline{K}_i(t) \equiv K_w(t) \quad i = 1, \ldots, N \tag{11}$$

From (3)

$$\sum_j q_{ij} = 0$$

To satisfy (11) then

$$\begin{aligned} -(A_i - B_i\tilde{R}_i^{-1}\tilde{B}_iF_i)'K_w - K_w(A_i - B_i\tilde{R}_i^{-1}B_i'F_i) - F_i'B_i\tilde{R}^{-1}B_i'F_i \\ = -(A_w - B_w\tilde{R}^{-1}B_w'K_w)'K_w \\ - K_w(A_w - B_w\tilde{R}^{-1}B_w'K_w) - K_wB_w\tilde{R}^{-1}B_w'K_w \end{aligned}$$

or equivalently

$$\begin{aligned} F_i'B_i\tilde{R}^{-1}B_iF_i - F_i'B_i\tilde{R}^{-1}B_i'K_w - K_wB_i\tilde{R}^{-1}B_i'F_i \\ = -(A_i - A_w)'K_w - K_w(A_i - A_w) - K_wB_w\tilde{R}^{-1}B_w'K_w \end{aligned} \tag{12}$$

Completing the square on the left hand side of (12)

$$\begin{aligned} (F_i - K_w)'B_i\tilde{R}^{-1}B_i'(F_i - K_w) &= (A_w - A_i)'K_w + K_w(A_w - A_i) \\ &\quad - K_wB_w\tilde{R}^{-1}B_w'K_w + K_wB_i\tilde{R}^{-1}B_i'K_w. \end{aligned} \tag{13}$$

But $K_w \geq 0$ from (7), and consequently (4) implies that the right side of (13) is nonequative. Taking the nonnegative square roots in (13),

$$\begin{aligned} F_i = K_w \pm (B_i'\tilde{R}^{-1}B_i)^{-1/2}[(A_w - A_i)'K_w - K_w(A_w - A_i) \\ + K_w(B_i\tilde{R}^{-1}B_i' - B_w\tilde{R}^{-1}B_w)K_w]^{1/2}. \end{aligned} \tag{14}$$

If either of the values for F_i given in (14) are substituted into (8), a decision rule is produced which has constant mean cost independent of r.

The matrix F_i can be viewed as the gain of the policy, and (9) indicates that $E\{J \mid x, r\}$ can be made insensitive to r with a high gain (the positive sign in (14)) or a low gain (the negative sign in (14)). This latter policy will be referred to as a "low-profile" policy and its properties investigated.

The statement that (8) yields constant mean performance is not sufficiently strong for the purposes of this paper. Let $t_{i,i-1}$ be the (random) time at which r makes the transition from i to $i-1$. It was shown in [2] that

$$\operatorname{Var} J = E\left\{\sum_{r_0} x(t_{i,i-1})' \Delta\bar{K}(t_{i,i-1})\, x(t_{i,i-1})\right\}$$

where

$$\Delta\bar{K}(t_{i,i-1}) = \bar{K}_{i-1}(t_{i,i-1}) - \bar{K}_i(t_{i,i-1})$$

By selecting F_i according to (14)

$$\Delta\bar{K}(t) \equiv 0$$

and as a consequence

$$\operatorname{Var} J = 0$$

Thus (8) yields not only a cost that has constant mean value independent of r, but the realized value of the cost is independent of r with probability one.

Since (8) has constant cost independent of w, it is an equalizer. To prove it is minimax, it need only be shown that (8) is Bayes with respect to a probability measure on Ω. Let $\omega_0 \in \Omega$ be characterized by

$$r(t, \omega_0) \equiv w \tag{15}$$

Since (8) with $F_i \equiv K_w$ is the optimal solution to the nonstochastic problem satisfying (15), (8) is Bayes with respect to the measure assigning probability one to $\omega \in \Omega$.[1] Thus (8) is minimax.

It is interesting to note that the policy given by (8) is not necessarily admissible in a decision theoretic sense. The classical theorem on the admissibility of minimax rules (see [1], Theorem 2.3) requires that the support set of the least favorable probability measure be broad. The singular measure found to be least favorable here certainly violates the broadness hypothesis.

4. An Example

To illustrate the mechanics involved in finding the low gain, minimax controller for a stochastic system, consider the following problem. Sup-

[1] If $r(t_0) \neq w$, (8) is actually extended Bayes. The conclusion of this section is still true in this case (see [1], theorem 3.9).

pose the system to be controlled is described by the scalar differential equation

(16)
$$\dot{x} = ax + v$$
$$x(0) = x_0$$

The coefficient process $a(t)$ is a Markov jump process with two possible states

$$a(t) = \begin{cases} 0; & r(t) = 1 \\ 4; & r(t) = 2 \end{cases}$$

The state $r(t) = 1$ is absorbing. ($q_{11} = 0$) and the residence time of r in state 2 is exponentially distributed with mean value 0.125.

The cost functional J is given by

$$J = \int_0^\infty (x^2 + v^2)\, dt$$

The mean-optimal control for this system was derived in [3] and is given by

(17)
$$v(t) = \begin{cases} -x; & r = 1 \\ -3x; & r = 2 \end{cases}$$

Further

(18)
$$E\{J \mid x_0, r(0) = 1\} = x_0^2$$
$$E\{J \mid x_0, r(0) = 2\} = 3x_0^2$$

Substituting (17) into (16) it follows that the closed loop, mean optimal system satisfies the equation

(19)
$$\dot{x} = \begin{cases} -x & r = 1 \\ x & r = 2 \end{cases}$$

Although mean optimal, it is still true that

(20)
$$\max_\omega J(\omega) = \infty \text{ if } r(0) = 2$$

From (4), the state $r = 2$ satisfies the criterion for being the least favorable state. From (7) (using the positive stationary solution)

(21)
$$K_w \cong 8.12 = F_2$$

From (14)

$$F_1 = 8.12 - \sqrt{8(8.12)}$$
$$= 0.10$$

Hence, the "low profile" control is

$$v = \begin{cases} -0.10x & \text{if } r = 1 \\ -8.12x & \text{if } r = 2 \end{cases} \tag{22}$$

The policy given by (22) is minimax;

$$\max_{\omega} J = 8.12x^2 \tag{23}$$

The gain used in (22) is far higher in the "worse" mode than is that used in (17). This stabilizes the system on all sample functions. The gain in the "better" mode is decreased in (23). Since operation is permitted to take place at increased incremental cost, a very low gain yields acceptable performance. The guaranteed performance (23) is inferior to the mean performance given in (18), but (23) is an assured level of performance rather than an average over the set of all possible system realizations.

4. Conclusion

This paper presents a policy which employs a minimal expenditure of energy to achieve acceptable performance. This policy makes performance insensitive to the realization of the exogenous disturbance which acts upon the system. The resulting system is inferior in average utility to that derived using the policy derived in [3], but the uncertainty surrounding the value of J is eliminated. Although the performance index is path invariant, the penalty associated with control utilization is a random variable.

References

[1] D. D. Sworder, *Optimal Adaptive Control Systems*, Academic Press, New York, 1966.
[2] D. D. Sworder, "A Simplified Algorithm for Computing Stationary Cost Variances for Optimally Controlled Stochastic Systems," (to appear).
[3] D. D. Sworder, "Feedback Control of a Class of Linear Systems with Jump Parameters, *IEEE Trans. on A.C.*, Vol. AC-14, No. 1, February 1969, pp. 9–14.

Department of Electrical Engineering
University of Southern California