Title
Management science implications of adaptive control methods

Authors
Rausser, Gordon C.
Pekelman, Dov

Publication Date
1980-04-01

117

Division of Agricultural Sciences
UNIVERSITY OF CALIFORNIA

Working Paper No. 117

MANAGEMENT SCIENCE IMPLICATIONS OF ADAPTIVE CONTROL METHODS

Gordon C. Rausser and Dov Pekelman

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
April 1980

MANAGEMENT SCIENCE IMPLICATIONS OF ADAPTIVE CONTROL METHODS

Gordon C. Rausser and Dov Pekelman

1. Introduction

The term 'adaptive' is often used to indicate the utilization of recent information in the decision making process. There are, however, various levels of information which can be employed for different purposes in the decision making process. In particular, there are two general categories of adaptivity which can be identified. To illustrate these two categories, consider some measured information which describes the state of the system, be it the monthly sales of a product or the output of a set of machines. This information can be used for making subsequent advertising, pricing or maintenance decisions, without any modification of the model describing the system. Alternatively the information can be utilized first to modify the model and its parameters and then as a basis for computing the desired decision. The common use of the term adaptive in managerial applications refers to the first approach. In this paper, the adaptive process will be viewed in the context of the second, more comprehensive approach. Not only is available information utilized for adapting the model, its parameters and the control decisions, but the quality of the information itself is controlled as well.

From the pragmatic standpoint, the adaptive control framework[1] recognizes that as a system progresses through time more data become available with which to update or revise estimates of the influence of alternative control variable settings on various performance measures. In general, revisions of system representations should not be regarded as separate from the derivation of an optimal policy. Different decisions may reveal more or less information about the actual system via different sets of the resulting data obtained. On the benefits side, such information can result in an "improved representation" of system structure resulting in superior future control. The incurred cost of such information emanates, in part, from choosing a current decision which is less than optimal from the pure control point of view.[2] In a concrete context, as Dreze [1972, p. 15] has suggested, "a monopolist may wish to depart from the price which maximizes expected profit, simply to learn more about his demand function."

Thus, optimal adaptive controls require a simultaneous solution to a combined control and sequential design of experiments problem and thus are dual in nature. This dual feature of the adaptive control framework was first recognized by Feldbaum [1965]. Early examinations of the dual control technique advanced by Feldbaum are available in Sworder [1966], Aoki [1967], Meier [1965] and Meier, Peschon, and Dressler [ ]. From this work, the anatomy of the dual control approach may be characterized by three principle elements, viz., direct control, learning, and design of experiments.

Neglecting the relationship between current information and future measurements, the control element takes into account the direct effect of decisions on the criterion function. The learning element is motivated by the existence of a set of sufficient statistics which are conditional on information related to the current state of the system and on the more recent estimate of the probability distribution function for the unknowns of the decision problem. With respect to the latter, as we proceed into the future, additional sample information becomes available. The use of these sample additions allows learning to take place regarding unknown parameters, unknown states, and the like.[3] Finally, since the generation of such sample information has a direct bearing on future estimates of the probability distribution functions, which in turn influence the performance of alternative decisions, an experimental dimension is involved in current adaptive control actions.

Unfortunately, it has not as yet been possible to express the adaptive or dual controls solutions in analytical form. Although numerous authors present the relevant recursive equations (see for example Feldbaum [1965] and Aoki [1967]), these equations do not yield explicit solutions. For this reason, from the methods standpoint, much in the way of research effort has been directed toward the development of analytical approximations and numerical feasible adaptive control rules.

A significant issue in the above applications along with other potential applications, is whether the systems encountered by the management scientist can indeed benefit from an adaptive approach which induces learning about its parameters. There seem to be three general cases where the importance of adaptive control is apparent. The first is when a new system is introduced and its parameters are generally unknown. Examples will be the demand equation for a new product; employee reactions to a new incentive scheme; the output-maintenance relations in new machines; or the production function of a new technology. The second case is where the system is established, but the parameters tend to vary over time. In order to ensure an acceptable control performance these parameters should be monitored and actively learned.

The third category, technically viewed as a subset of the first two, is where improvement of the model representing the system is sought. In particular, if two candidate models are competing, and the control is determined on the basis of a composite model formed by a convex combination of the two, the weighting scheme can be modified continuously as a result of learning. If the two models are nested, the discrimination between the two may be improved by adaptive control and hence the performance of the system may be enhanced.

In this paper we structure our examination by first formulating the general adaptive stochastic control problem and defining the associated information elements. Stochastic control policies are then classified with regard to their information usage. This classification, although not exhaustive, provides a framework which allows various adaptive control approximations to be generated. In order to motivate the need for these approximations we formulate a simple two period problem to illustrate the inherent analytical difficulties in its solution. This problem is also utilized to demonstrate the distinction between passive learning and active learning. Schemes based on a passive framework admit only accidental learning and thus do not involve an experimental dimension. All the active schemes conform to what Raiffa and Schlaifer [1961] have referred to as "preposterior analysis." Such schemes involve an experimental dimension which probes the system in anticipation of the value of information to be derived from future observations. Following descriptions on both the passive and active learning schemes, available results on the performance of these schemes are reported.

Since the first computationally feasible algorithms for general adaptive control problems are of very recent vintage, the empirical applications of these schemes to real world problems are rather meager. The few applications which are available are reported in Section 7. Finnaly, a concluding section offers a number of problems which would benefit substantially from the perspective provided by the adaptive control approach and presents an overview of issues faced in the further development of adaptive control methods.

2. General Model Specification

The general adaptive stochastic control problem may be defined as follows. The state of the system at time k, $\underline{x}(k)$ evolves according to:

(1) $\underline{x}(k+1) = f[k, \underline{x}(k), \underline{p}(k), \underline{u}(k), \underline{v}(k)], \; k=o,\ldots,N-1$

where $\underline{p}(k)$ is a vector of parameters, $\underline{u}(k)$ is the vector controls applied at time k and $\underline{v}(k)$ is the process noise.

At time k, prior to applying the control, the state of the system $\underline{x}(k)$ may be observed. However, as a result of the measurement noise $\underline{w}(k)$, only partial observation y(k) can be achieved, where

(2) $\underline{y}(k) = \underline{h}[k, \underline{x}(k), \underline{w}(k)], \; k=1,2,\ldots,N.$

The variables $\underline{x}(k)$, $\underline{p}(k)$, $\underline{u}(k)$, $\underline{v}(k)$, $\underline{y}(k)$, w(k) are vectors with the following dimensions; (n x 1), (r x 1), (m x 1), (n x 1), (s x 1) respectively. The statistics of the random elements

(3) $\underline{x}(0), \; \{\underline{v}(k)\}_{k=0}^{N-1}, \; \{\underline{w}(k)\}_{k=0}^{N-1}$

and the functional form f(.), h(.) are assumed known.

The objective function is represented by the minimization of the expected cost

(4) $J = E\,\{C(N, X^N, U^{N-1})\}$

where the expectation is taken over all the underlying random variables, C is a real-valued function, and U^k and X^k, k=0,...,N are defined as:

(5) $U^k \equiv \{\underline{u}(i)\}_{i=0}^{k}$

$X^k \equiv \{\underline{x}(i)\}_{i=0}^{k}$

Accordingly, the minimization of the expected cost is performed with respect to the sequence of controls U^{N-1} applied during the N-stage control process.

The set of observations from time k=1 to time j, after the sequence U^{j-1} has been applied is denoted by

(6) $Y^j \equiv \{\underline{y}(k)\}_{k=1}^{j} \qquad 1 \leq k \leq j \leq N$.

The adaptive control problem, as a special case of the general stochastic control problem, treats the case where the vector of parameters $\underline{p}(k)$ evolves according to

(7) $\underline{p}(k+1) = \underline{g}(k,\underline{p}(k), \underline{\theta}(k))$

where $\underline{g}(.)$ and the statistics of $\underline{p}(0)$ are known. This formulation admits the special case of constant, but unknown parameters. Clearly, $\underline{p}(k)$ can be regarded as a state and augmented to $\underline{x}(k)$ so that the new state vector is

(8) $\underline{x}'(k) = [\underline{x}(k), \underline{p}(k)]$

The various informational elements contained in the above system may be captured by a few additional definitions. In particular, the information about the dynamics of the measurement system between time 0 and time j may be denoted as:

(9) $m^j \equiv \{h(k,.,.)\}_{k=0}^{j}$,

i.e. m^j denotes the knowledge on ths structure of the observation program up to time j. Along with this definition, the following joint density distributions will prove useful in characterizing alternative adaptive control schemes

(10) $L^k \equiv dP[\underline{x}(0),\underline{p}(0), \{v(i)\}_{i=0}^{N-1}, \{\theta(i)\}_{i=0}^{N-1}, \{w(i)\}_{i=0}^{k}], k=1,\ldots,N-1.$

$L^o \equiv dP[\underline{x}(0),\underline{p}(0), \{v(i)\}_{i=0}^{N-1}, \{\theta(i)\}_{i=0}^{N-1}],$

Note that this representation implies that L^o does not contain any information on the measurement statistics.

To illustrate the adaptive control framework consider a machine maintenance problem where both the rate of deterioration and the effectiveness of investment in maintenance of the machine output are unknown and may possibly vary over time. Let x(k) be the machine output in period k which is generated in accordance with

(11) $$x(k+1) = \alpha(k)x(k) + \beta(k)u(k) + \varepsilon(k), \qquad k=0,\ldots,N$$

and where the control u(k) is the dollar amount invested in maintenance, $\alpha(k)$ and $\beta(k)$ are unknown coefficients expressing the deterioration rate in output and effectiveness of investment in maintenance respectively, and ε_k is a random variable with known statistics. The objective is to find the sequence of investment (u(0),u(1),...,u(N-1)) which will maximize the expected discounted profit over a given horizon, i.e.,

(12) $$\max \; E\,\{C(N)x(N)\delta(N) + \sum_{k=0}^{N-1} [C(k)x(k)-u(k)]\delta(k)\}$$

where C(k) expressed the profit per unit output and $\delta(k)$ is the discounting factor.

In order to have a well formulated problem the evolution of the parameters $\alpha(k)$ and $\beta(k)$ should be specified. The assumption of constant but unknown parameters can be expressed as:

(13) $$\alpha(k+1) = \alpha(k) \quad k = 1, \ldots, N;$$
$$\beta(k+1) = \beta(k) \quad k = 1, \ldots, N;$$

where initial values $\alpha(0)$ and $\beta(o)$ have known prior distributions. Similarly, a representation of random parameters may be modeled as:

(14) $$\alpha(k+1) = \alpha(k) + \theta_1(k)$$
$$\beta(k+1) = \beta(k) + \theta_2(k)$$

where $\theta_1(k)$, $\theta_2(k)$ are random variables whose statistics may be known or unknown, but typically whose prior distributions are known. Of course, by augmenting the coefficients $\alpha(k)$, $\beta(k)$ to the state of the system, x(k), the problem assumes the form of a general stochastic control formulation.

3. Classes of Stochastic Control Problems

The various classes of stochastic control problems can be defined by the amount of information utilized to determine the control for each period. These classes are not exhaustive and are advanced only to provide some overall structure. It is particularly important to note the difference between the feedback policy and the closed loop policy. For both policies available information is utilized to determine current decisions, however, only in the case of the latter policy does the controller anticipate future information and take it into account in his current control calculations.

3.1 The open loop policy

During the horizon [0,...,N-1] no information is available to the controller, i.e. all decisions are made on the basis of the information available at time 0. Therefore the control rule has the form

(15) $\underline{u}(k) = \underline{u}[k, U^{k-1}, L^{o}], k=0,\ldots,N-1.$

3.2 The Open Loop Feedback Policy

At each period k, as the data becomes available, the controller observes this information but in his determination of the control he assumes that no future observations will be available to him. In other words, the controller for each period k is presumes to have no knowledge about future observations. Hence, the control rule in this case has the form:

(16) $\underline{u}(k) = \underline{u}(k, U^{k-1}, Y^{k}, m^{k}, L^{k}], \quad k=0,\ldots,N-1.$

3.3 The M-measurement feedback policy

For this policy, in addition to the current data, the subsequent M-measurements with their statistics are available to the controller. i.e., the control decision is based on the current measurement and M additional anticipated measurements. Here we have:

(17) $\underline{u}(k) = \underline{u}[k, U^{k-1}, Y^{k}, m^{k+M}, L^{k+M}], \quad k=0,\ldots,N-1.$

3.4 The closed loop policy

In this policy the current information as well as all future anticipated information and their corresponding statistics are taken into account, i.e. it is known that the loop will be closed through all the planning horizons. For this policy the rule will have the form:

(18) $\underline{u}(k) = \underline{u}[k, U^{k-1}, Y^{k}, M^{N-1}, L^{N-1}], \quad k=0,\ldots,N-1.$

Note that for a deterministic system all four classes are identical.

3.5 Illustration of alternative policies

Further insight into the above policies may be gained by examining a stochastic problem which is not adaptive in its parameters, viz. an inventory system where the control decision is the amount to order, and the observation is the demand realization (or inventory on hand). For an N-period problem the open loop policy will consist of N decisions, one for each of the N periods. These ordering decisions will be made at time zero, on the basis of the inventory level at that time and the demand distributions for all future periods. No further observations on future demand realization will be made. In the feedback policy the ordering decision will be made in each period k after the observation of the realized demand of the period. However, it will be assumed that no future demand observations will be available, i.e. corrective orders in future periods will not be possible. Note that what constitutes a feedback policy for the horizon [0, N] is actually the set of a first-period solutions of an open loop policy with horizons [0,N], [1,N],...[N,N]. This policy will differ from what is commonly coined as "rolling horizon" problem, i.e. when the length of the horizon remains constant and the set of first-period solutions is generated for the horizons [0,N],[1, N+1],...,[N,2N].

To illustrate the M-measurement policy consider the case where during the N-period problem observations of demand realizations will be available only for the first M periods. Then the ordering decision at time k=0 will be based on the knowledge that corrective ordering can be made during the coming M periods. In period k=1 there will be only M-1 corrective orders in the future and therefore the policy will be an M-1 measurement policy and in general for period k, $k \leq M$, we have an M-k measurement policy until k=M where the policy will become a feedback or zero-measurement policy. The closed loop policy is the more standard inventory policy where observations of demand realization are available for all subsequent periods. For an N-period planning horizon the closed-loop policy is an N-measurement policy.

In general, the optimal policy for a stochastic control problem will be closed loop. However, due to analytical difficulties in deriving the closed loop control rule the other policy classes are commonly used as an approximation. The order of the analytical difficulty usually decreases as we reduce the amount of information available to the controller.

The common formulation of the adaptive control problem has some inherent nonlinearities since the unknown parameters are usually the coefficients of the state and control variables. Therefore, the adaptive control problem typically constitutes a nonlinear stochastic control problem and requires an approximate solution. In the following section we first demonstrate by a particular example the nature of the closed loop solution and its inherent analytical difficulties. We then introduce the notion of actively adaptive and passively adaptive solutions and subsequently relate them to various approximation methods.

4. Approximate Solutions

To properly classify various schemes, it is important to first distinguish between approximations performed on the original system and those performed in the process of deriving the optimal rule. Most approximations resulting from the first approach will fall in the category which we define as passively adaptive schemes while the latter type approximations will generally be actively adaptive. In order to motivate the need for approximation we solve a simple two-period problem and show the analytical difficulty in deriving the optimal solution. We utilize this example to illustrate as well the nature of passively adaptive versus actively adaptive solutions.

4.1 Example

Consider the two-period problem:

$$(19) \quad J = \min E \sum_{k=1}^{2} (x_k + 1/2 \cdot u_k^2)$$

subject to

$$(20) \qquad x_k = \alpha u_k + \varepsilon_k , \quad k=1,2$$

where $\varepsilon_k \sim N(0,q)$ for k=1,2, cov $\varepsilon_1 \varepsilon_2 = 0$; and α is the unknown parameter with the prior distribution at time k=0 given as

$$(21) \quad \alpha \sim N(\alpha_o, \sigma_o).$$

The first decision will be u_1 which when imposed will result in x_1. After the realization of the random variable x_1, the distribution of the parameter

α will be reestimated. Using a Bayesian procedure, the aposteriori mean and variance of α , denoted as α_1, and σ_1, are:

$$(22) \qquad \alpha_1 = \frac{\alpha_o q + \sigma_o u_1 X_1}{u_1^2 \sigma_o + q}$$

$$(23) \qquad \sigma_1 = \frac{\sigma_o q}{\sigma_o u_1^2 + q}$$

Clearly the selection of u_1 and the resulting x_1 will effect these estimates. In particular, if $|u_1|$ is large, the updated variance of α will be small. Hence a large control value may assist in learning about α_o. In order to determine the optimum value of u_1, the value of usefulness of additional information on α must be captured. This value, in the context of a given control problem, allows the trade off between current control and learning for future control to be estimated. If the estimate of α_1 can be improved, a better decision u_2 can obviously be made. We call this argument the active learning part of the decision since deliberate learning in terms of improved estimation is exercised.

The other dominant consideration in determining u_1 is control, i.e. the impact on the objective function for the current period. In this particular example, the only interaction between the two periods is through the learning argument. The underlying system is inherently static since the state x_{k+1} is independent of x_k. Therefore, we will be able to isolate the static terms corresponding to current <u>control</u> and the dynamic one which corresponds to <u>learning</u>.

Let us turn now to the solution of the problem. Using dynamic programming we have:

$$(24.1) \qquad J_3^* \equiv 0$$

$$(24.2) \qquad J_2^* = \min_{u_2} E\{(x_2 + 1/2\, u_2^2 | x_1, u_1) + J_3^*\}$$

$$(24.3) \qquad J_1^* = \min_{u_1} E\{(x_1 + 1/2\, u_1^2 | \alpha_o, \sigma_o) + J_2^*(u_1)\}$$

Substituting for x_2 and applying the expectation operator we obtain:

$$(25) \qquad J_2^* = \min_{u_2}\{\alpha_1 u_2 + 1/2\ u_2^2\}$$

which yields

$$(26) \qquad u^*_2 = -\alpha_1$$

and

$$(27) \qquad J^*_2 = -\ 1/2\ \alpha_1^2$$

where $\alpha_1 = E\{\alpha|X_1,\ u_1\}$, i.e., α_1 is the aposteriori mean of the parameter after the decision u_1 is taken and the realization x_1 is obtained. Substituting (22) into (27) for J_2^* we have

$$(28) \qquad J_2^* = -\ 1/2\left(\frac{\alpha_o q + \sigma_o u_1 x_1}{u_1^2 \sigma_o + q}\right)^2$$

which is a random variable due to x_1. The cost of learning is negative and is represented by the expected minimum cost

$$(29) \qquad E(J_2^*) = -1/2(\alpha^2 + \frac{\sigma_o^2\, u_1^2}{u_1^2 \sigma_o + q})$$

This result indicates that learning is beneficial (negative cost) to future control.

Shifting one stage backward and substituting (28) for $J_2^*(.)$ into (24.3) we have

$$(30) \qquad J_1^* = \min_{u_1} E\ \{(\alpha u_1 + \epsilon_1 + 1/2\ u_1^2) - 1/2\ \left(\frac{\alpha_o q + \sigma_o u_1 x_1}{u_1^2 \sigma_o + q}\right)^2 \Big|\ \alpha_o, \sigma_o\ \}$$

and after taking expectations

$$(31) \qquad J_1^* = \min_{u_1}\ \{(\alpha_o u_1 + 1/2\ u_1^2) - 1/2 \quad (\alpha_o^2 + \frac{\sigma_o^2\ \ u_1^2}{u_1^2\ \ \sigma_o + q})\}$$

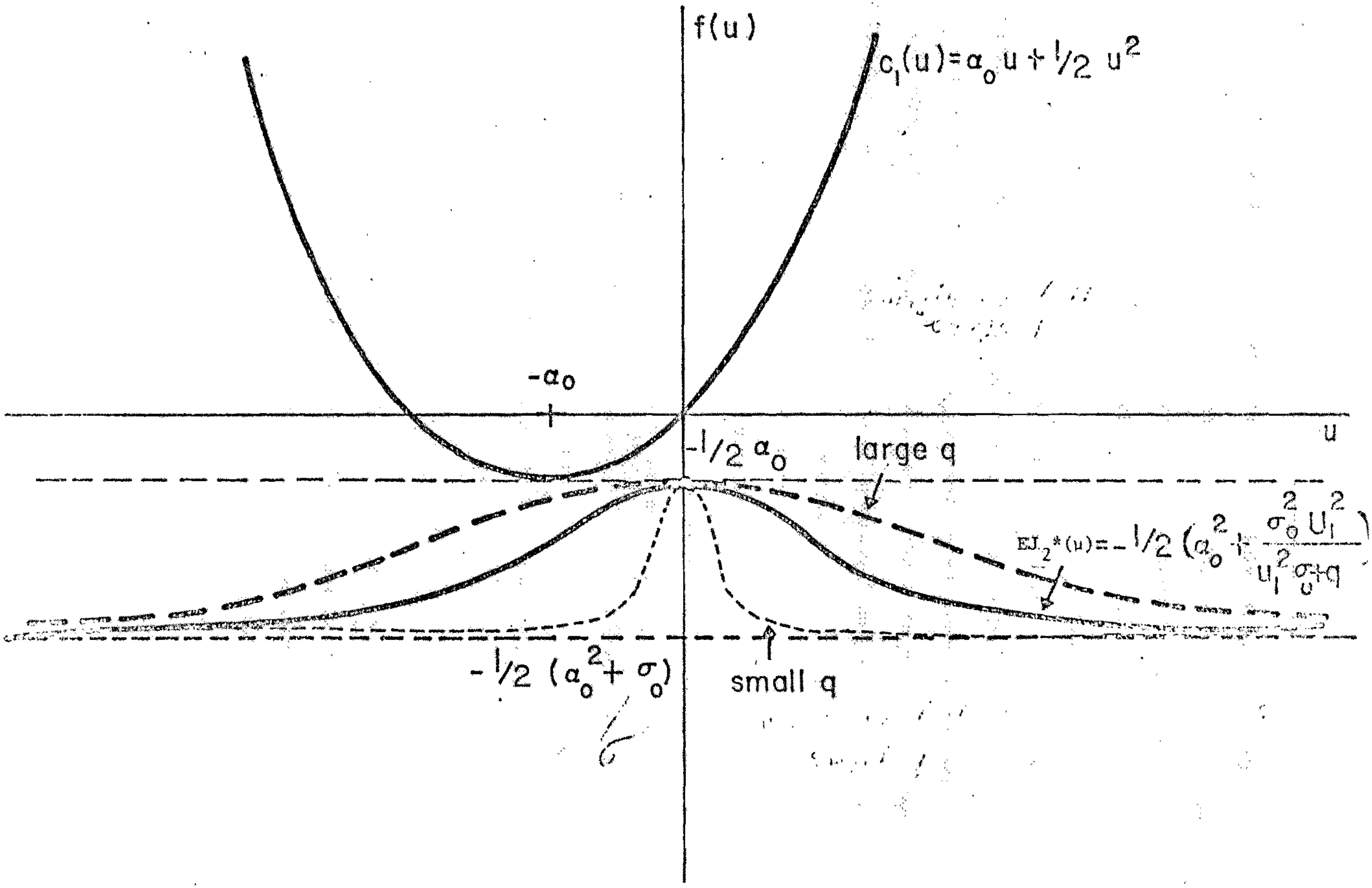
where the first term of (30) refers directly to control while the second term refers directly to learning. The optimum value of the u_1 must achieve some balance between present control (the first term) and learning for future control (the second term).

Unfortunately the minimization of the two "dual" cost terms in (31) is not a simple analytical task. As shown in Figure 1, the current control $C_1(u)$, and future learning, $EJ_2^*(u)$, components of the dual cost suggests that (where α_o is assumed to be positive): (i) if q is small, the optimal u_1 is approximately $-\alpha_o$; (ii) if q is moderate, the optimal u_1 lies to the left of $-\alpha_o$; and (iii) if q is very large, the u_1 is approximately $-\alpha_o$. For case (i), the small value of q implies unplanned (or accidental) learning is sufficient to reduce the updated variance (23) while for case (iii) a large value of q, suggesting high noise intensity, implies planned learning does not pay off and thus there is no need to look beyond a one-stage optimization problem. Only if q is of moderate size does planned learning prove beneficial. Note that a similar conclusion can be drawn for different values of σ_o.[4]

Although the system (20) is static, dynamic interactions are introduced by the reestimation of α after the realization u_1, x_1, i.e. the effect of u_1 and x_1 on α_1 and therefore on the decision in the second stage. The decision u_1 in the first stage and hence the realization x_1 may improve the estimation of α and consequently the second period decision u_2 along with the resulting payoff J_2^*. Hence, when the decision u_1 is made the value J_2^* is taken into account. This is represented by the second term in expression (31).

In this particular example, the effects of learning and control can be clearly separated. This is possible because of the static nature of system i.e. current control effect directly current state, but not future states. However, even with such a simple formulation, the three period proolem cannot be solved analytically. Solving the three period problem is equivalent to proceed one additional step backward in the above analysis. This is analytically infeasible since we cannot find a closed form solution for u_1 to substitute in the value function. Therefore some approximations are required to obtain a solution.

Figure 1 Current Control Versus Future Learning



When the decision is concerned with parameter system learning as well as with the control we will call it active learning. If the active learning of α is ignored, i.e., if we eliminate the second term in the J_1^* expression, the problem will be completely static, and the optimal strategy is simply:

$$(32.1) \qquad u_1^* = -\alpha_o$$

$$(32.2) \qquad u_2^* = -\alpha_1$$

The learning element will still exist, since α will be reestimated, however it will be completely passive, i.e., the amount of learning about α will not be deliberately directed by any previous decision. Various approximation approaches involve the partial or complete neglect of active learning. In this example, an obvious approximation is generated by adding a simplifying assumption to the original system. Specifically, the mean estimate of α_1 is presumed to be the true value of α_1. This assumption causes the removal of the learning term in (31). As noted later, this simplifying assumption leads to the heuristic certainty equivalent specification.

An example of an approximation performed during the derivation of the rule, rather than on the basic system, would be an expansion up to a second order of the active learning term in J^* around some nominal values. If we select the nominal value in this case to be $u_1 = -\alpha_o$, we will obtain after the expansion to the second order the following:

$$(33) \qquad J_1^o = \min_{u_1} \{\alpha_o u_1 + 1/2\, u_1^2 + Q_o + \frac{\partial Q_o}{\partial u_1}(u_1+\alpha_o) + \frac{\partial^2 Q_o}{\partial u_1^2}\, 1/2\, (u_1+\alpha_o)^2\}$$

where Q_o is the active learning term in J_1^* evaluated at the nominal trajectory $u_1^o = -\alpha_o$. Differentiating and substituting we have:

$$(34) \qquad J_1^o = \min_{u_1} \{\alpha_o u_1 + 1/2\, u_1^2) - 1/2\, (\alpha_o^2 + \frac{\sigma_o^2 \alpha_o^2}{\sigma_o \alpha_o^2 + q})$$
$$- \frac{\alpha_o \sigma_o^2 q}{(\alpha_o^2 \sigma + q)^2}(u_1+\alpha_o) + 1/2\, [\frac{3\alpha_o^2 \sigma_o^3 q - \sigma_o^2 q^2}{(\alpha_o^2 \sigma_o + q)^3}]\, (u_1+\alpha_o)^2\}$$

and hence a closed form expression for u_1 exists, and multiple-period problems can then be solved recursively. Note that the active learning part is approximated but will still appear in the control rule.

To recapitulate, two approximate forms were introduced, one on the original system by ignoring the uncertainty on the parameter and the other on the optimal rule after the expectations were taken. We noted that the first approximation generates a control rule with passive learning while the second form preserves the element of active learning. For many of the common adaptive control approximations this general observation holds. That is, when the approximation is on the original system, the control rule is often passive while if the approximation is performed in the derivation of the rule some of the active learning elements are preserved.

5. Passive Adaptive Schemes

We describe here four basic schemes, the certainty equivalence, the open loop feedback, the sequential decision and the linear quadratic gaussian scheme. All these schemes involve some modification of the original system.

5.1. Certainty Equivalence (C.E.). In this scheme it is assumed that the certainty equivalence property holds i.e. that the mean value of all random variables in the system is their true value. For every period k, the current estimate of the mean value of the random variables is used to replace the random variables and the problem is solved as a deterministic problem. If the control rule for a completely deterministic problem is:

$$\underline{u}^d(k) = \underline{u}^d\ (k,\ \underline{p}(k),\ \underline{x}(k)) \tag{35}$$

then the C.E. control rule is:

$$\underline{u}^{ce}(k) = \underline{u}^d(k,\hat{\underline{p}}(k),\hat{\underline{x}}(k)) \tag{36}$$

where $\hat{\underline{p}}(k)$, $\hat{\underline{x}}(k)$ are the recent estimates of the parameter and the state means.

In terms of the information used, the C.E. control rule has the form:

$$\underline{u}^{ce}(k) = \underline{u}[\ k,\ U^{k-1},\ Y^k,\ M^k]\ . \tag{37}$$

The rule is a feedback rule, since the parameters and the states are reestimated every period. However, no account is taken for future measurements and thus future uncertainty.

Under certain conditions, (see Simon [1956] and Theil [1957]) the C.E. property holds.[5] In this case the C.E. control rule is the optimal one, and there is nothing to be gained by utilizing information on future measurements and uncertainty. A closely related property is the separation property. This property holds if the information about the parameters and the states utilized in the optimal control rule can be captured by their mean value, i.e.,

$$(38) \quad \underline{u}(k) = u^{s}(k, \hat{\underline{P}}(k), \hat{\underline{x}}(k)).$$

Here the functional form of $u^{s}(.)$ may differ from the deterministic rule $u^{d}(.)$. The C.E. property has two basic characteristics. First the functional form of the control rule is equivalent to the deterministic one, and second the random variables can be replaced by their mean values. The separation property has only the second characteristic and therefore the C.E. property implies the separation property but not vice versa. Discussion of the general conditions under which these properties hold can be found in Wonham [1968] and Bar-Shalom and Tse [ ]. The word separation is used in indicate the complete independence between the procedures used for estimation of the mean value and those used for the control calculation.

The C.E. property which was first formulated by Simon [1956] and Theil [1957] was shown to hold for linear systems with quadratic objective functions. In most adaptive control systems the C.E. property does not hold as a result of the nonlinearity introduced by the product of the parameters and states. The use of the C.E. assumption as an approximation has also been called forced separation (Ku and Athans [1973]) and heuristic certainty equivalence (Norman[197

5.2 Open Loop Feedback. This approach was suggested by Dreyfus [1965] and studied by Aoki [1967] , Bar-Shalom and Sivan [1969], Curry [1969], Ku and Adams [1973], and Tse and Athans [1972]. The approach is also called open loop optimal feedback. At time k when the control u_k is calculated, it is assumed that no measurement will be obtained in the future. The stochastic elements of the parameters as well as those of the states are taken into account. The use of information in this scheme is identical to our previous definition of the feedback class.

In the process of the derivation of the optimal rule for these schemes, we face the difficult task of determining the expectation of the product of the parameters and the states along with the parameters and the controls. Tse and Athans [1972], for example, avoid the problem by assuming that the parameter multiplying the state is nonrandom, while Ku and Athans [1973] assume that the expected product of the parameter and the state is equal to the product of the expectation. Both of these assumptions obviously represent misspecifications.

For a time-varying parameter problem with quadratic objective function, a linear system and Gaussian noise (we denote it linear quadratic Gaussian -- LQG) this approach will yield a linear rule of the following form:

$$(39) \quad \underline{u}(k) = - G(k)\hat{\underline{x}}(k) + \underline{\gamma}(k)$$

where $G(k)$ is the adaptive gain matrix and $\underline{\gamma}(k)$ is the adaptive correction term. These two quantities are a function of the estimated and predicted mean and covariances of the parameters. It is interesting to note that for a linear-quadratic system where the certainty equivalence assumption holds the optimal rule has the same form, but $G(k)$, $\underline{\gamma}(k)$ are functions only of the mean values of the parameters.

5.3 Sequential Stochastic Control. Rausser and Freebairn [1974a], Yoshida and Nakamure [1973], Zellner [1971] Chow [1975] and Prescott [1971] have investigated this approach. The derivation of the control rule is based on the assumption that observations will indeed be available in the future. However, they will not be used to adapt the distribution of the parameters. Chow [1975] has referred to this approach as control without learning. In the case of a linear system and quadratic objective function the optimal rule here is linear in the mean value of the state as in the open loop feedback. In essence, all of this work treats the unknown paramers as independent, identiacally distributed random variables in each period.

5.4 Linear Quadratic Gaussian Approximations. This scheme follows the same arguments as presented in Athans (1971). If the system is nonlinear and/or the objective function nonquadratic, an expansion around some nominal value may be performed as follows. Let $\{\underline{u}^o(k)\}$ be the nominal control sequence, $\{\underline{v}^o(k)\}$ the nominal noise input, $\underline{p}^o$ the nominal parameter values and $\{\underline{x}^o(k)\}$ the resulting nominal trajectory. The nominal values can be generated by imposing on the original problem the certainty equivalence assumption.

Performing a first order perturbation on the non-linear system

(40) $\underline{x}(k+1) = f(k, \underline{x}(k), \underline{p}(k), \underline{u}(k), v(k))$

(41) $\underline{y}(k+1) = h(k, \underline{x}(k), \underline{p}(k), \underline{u}(k), w(k))$

around the nominal values we obtain the following:

(42) $\delta\underline{x}(k+1) = \left.\frac{\partial f(.)}{\partial \underline{x}(k)}\right|_o \delta\underline{x}(k) + \left.\frac{\partial f(.)}{\partial \underline{p}(k)}\right|_o \delta\underline{f} + \left.\frac{\partial f(.)}{\partial \underline{u}(k)}\right|_o \delta\underline{u}(k) + \left.\frac{\partial f(.)}{\partial \underline{v}(k)}\right|_o \delta\underline{v}(k)$

(43) $\delta\underline{y}(k) = \left.\frac{\partial h(.)}{\partial \underline{x}(k)}\right|_o \delta\underline{x}(k) + \left.\frac{\partial h(.)}{\partial \underline{p}(k)}\right|_o \delta\underline{p}(k) + \left.\frac{h(.)}{\underline{u}(k)}\right|_o \delta\underline{u}(k) + \left.\frac{\partial h(.)}{\partial \underline{w}(k)}\right|_o \delta\underline{w}(k)$

where $\Big|_o$ indicates that the derivatives are evaluated at the nominal values, and at the mean value of $\underline{w}(k)$, and

(44) $\delta x(k) = \underline{x}(k) - \underline{x}^o(k)$

(44.2) $\delta\underline{p} = \underline{p}(k) - \underline{p}^o(k)$

(44.3) $\delta\underline{u}(k) = \underline{u}(k) - \underline{u}^o(k)$

If the nominal values are regarded as desired or target trajectories, the objective function may be formulated such that any deviation from these trajectories will be penalized. In particular we minimize the following quadratic objective function:

(45) $E\{\delta\underline{x}'(N)F\delta\underline{x}(N) + \sum_{k=0}^{N-1} \delta\underline{x}'(k)Q(k)\delta\underline{x}(k) + \delta\underline{u}'(k)R(k)\underline{u}(k)\}$

The selection of the matrices F, Q, and R are discussed in Athans

The solution to the above system assuming the input noise is normal is

(46) $\delta\underline{u}^*(k) = -\underline{G}_1(k)\,\delta\hat{\underline{x}}(k) - \underline{G}_2(k)\,\delta\hat{\underline{p}}(k)$

where $\underline{G}_1(k)$, $\underline{G}_2(k)$ are gain matrices, and $\delta\hat{\underline{x}}(k)$, $\delta\hat{\underline{p}}(k)$ are conditional estimates for $\delta\underline{x}(k)$ and $\delta\underline{p}$ based on the realizations,

(47) $\{\delta\underline{y}(1),\ldots,\delta\underline{y}(k)\}$ and $\{\delta\underline{u}^o(0),\ldots,\delta\underline{u}^o(k-1)\}$.

Although this scheme is relatively easy to implement, its performance will deteriorate if the uncertainty in the parameters or in the input noise is high.

5.5 Averaging Across Deterministic Control Rules. This scheme was developed by Deshpande, Mpadhyay and Laniotis (1973) and extended in Upadhyay and Laniotis (1974). It considers a linear system with quadratic performance and Gaussian noise, where the parameters of the system may not be known. Let $\underline{\theta}$ be the vector of unknown parameters, and assume for expositional purposes, that $\underline{\theta}$ can take discrete values $\underline{\theta}=\underline{\theta}_i$ $i = 1\ldots, n$ with probabilities $P_{\underline{\theta}}(\underline{\theta}_i)$ $i=1,\ldots n$. For a particular value $\underline{\theta}_i$ the control problem becomes the standard linear quadratic gaussian problem with known parameters. It, therefore, has the linear solution:

$$u(\underline{\theta}_i,k) = K(\underline{\theta}_i, k)\,\hat{x}(k/k)$$

where $K(\underline{\theta}_i,k)$ is the gain matrix for period k based on the particular values of the parameters $\underline{\theta}_i$, and $\hat{x}(k,/k)$ is the state vector estimate based on all observations up to period k.

The optimal control law for the adaptive control problem is obtained by weighting each particular control law $u(\underline{\theta}_i,k)$ by the appropriate probability $P_{\underline{\theta}}(\underline{\theta}_i)$, i.e.

$$u^*(k) = \sum_i u(\underline{\theta}_i,k)\,P_{\underline{\theta}}(\underline{\theta}_i) = \sum_i K(\underline{\theta}_i,k)\hat{x}(k/k)P_{\underline{\theta}}(\underline{\theta}_i)$$

The learning element is introduced by adjusting the distribution $P_{\underline{\theta}}(.)$ after each observation by a Bayesian rule. The matrix $K(\underline{\theta}_i,k)$ for all i,k can be computed off-line, and $\hat{x}(k/k)$ be estimated by kalman filter. In addition to the advantage of simplicity, this approach provides an interesting interpretation. In particular, this framework expresses the system by competing models each represented by the particular value of $\underline{\theta}_i$ and its probability $P_{\underline{\theta}}(\theta_i)$ of being the 'true' model.

6. Actively Adaptive Schemes

For situations in which the need for learning is explicitly recognized, knowledge can be accumulated in an active or dual control fashion. A number of schemes have been advanced in the literature for approximating actively adaptive control problems. Most of these schemes provide a consistent approach to the entire planning horizon and thus are approximations to the optimal closed loop solution. When these schemes are applied only to a subset of periods within the planning horizon, they approximate the m-measurement feedback solution.[6] Each of these schemes view the future by utilizing in various fashions what is currently known about the information to be obtained later.

The key distinguishing feature of the alternative schemes is how they deal with the dependence of future information on present controls. For the general problem, (1)-(4), the information state along with the optimal-cost-go associated with each information state must be characterized. The information state denoted by Ψ_k, is directly influenced by the conditional density, $p(x(k)|Y^k, U^{k-1})$. The general expression for how future observations will be made and utilized may be represented by the optimal-cost-go, $\Lambda^*_{k+1}\{\Psi_{k+1}, k+1\}$. Given these definitions along with a criterion function (4) which is separable across k, a stochastic dynamic programming solution expresses how $\Lambda^*_k\{\Psi_k, k\}$ can be computed (in principle) recursively by

(48) $\Lambda^*_k\{\Psi_k, k\} = \min E\{C[x(k), u(k), k] + \Lambda^*_{k+1}\{\Psi_{k+1}[\Psi_k, u(k)], k+1\}|Y^k, U^{k+1}\}$

where $\Psi_{k+1}[\Psi_k, u(k)]$ monitors the evolution of the information state. Three of the principle difficulties which arise in attempting to deal with (8.35) are: (i) the information in either infinite dimensional or finite but grows with time; (ii) the optimal-cost-go associated with the information state is generally not an explicit function; and (iii) storage of the control value associated with each information state at time k, k = 0,..., N - 1 is indeed difficult due to the large dimensionality. Approximations are offered by the following schemes to deal with these difficulties and thus all the schemes involve some simplification of the experimental or active learning dimension of dual control.

6.1 Adaptive covariance control. An actively adaptive scheme which admits perhaps the simplest computational framework has been developed by MacRae [1972, 1975]. This procedure has been referred to as linear experimentation by MacRae and as the adaptive covariance solution by Rausser and Freebairn [1973]. Essentially, it alters the original dual control structure of the problem so that equations defining the approximate solution may be derived while maintaining the interaction between estimation and control. Moreover, the equations defining the approximate solution readily admit terms associated with such economic concepts as the price of information and the value of estimation.

In the development of this scheme, (i) the system (1) assumes the linear form,

(49) $$x_k = Ax_{k-1} + Bu_k + Cw_k + \varepsilon_k$$

$$= Dz_k + \varepsilon_k$$

i.e., $D = [A, B, C]$, $z_k' = [x_{k-1}', u_k', w_k']$; (ii) no measurement errors (the state vectors x_k are completely accessible for all k periods), i.e., $y_k = x_k$; the criterion function has the quadratic form

$$(50) \qquad J = E\left\{ \sum_{k=1}^{N} (1+\delta)^{-k} \left[2g_k' x_k + 2h_k' u_k - x_k' G_k x_k - u_k' H_k u_k\right]\right\}$$

where N is the terminal horizon and δ is the discount rate. Note that the parameters contained in the matrix D are presumed temporally invariant, normally distributed, and uncorrelated with ε. The prior mean of D is assumed to be available, D_o, as is the variance-covariance matrix of D, Γ_o.[7] The disturbances, ε_k, are assumed intertemporally independent and normally distributed with zero expectation and (known) stationary variance-covariance matrix Ω.

For the above specification, the posterior or conditional distribution of D, given observations through period k, will be multivariate normal (Zellner [1971]). Thus, revisions in conditional estimates of Γ and D can be determined recursively by

$$(51) \qquad \Gamma_k^{-1} = \Gamma_{k-1}^{-1} + \Omega^{-1} \otimes z_k z_k'$$

and

$$(52) \qquad L(\bar{D}_k') = \Gamma_k[\Gamma_{k-1}^{-1} L(\bar{D}_{k-1}') + (\Omega^{-1} \otimes z_k)\, x_k]$$

where L denotes the stacking operator and $\otimes$ is the Kronecker product For this formulation, the relevant information state is completely characterized by (51) and (52).

The simplification introduced by MacRae [1972, 1975] involves replacing unknown future observations by expected observations in the above conditional mean (51) and covariance update (52) equations. In other words, (51) and (52) are approximated by:

$$\Gamma_k^{-1} = \Gamma_{k-1}^{-1} + \Omega^{-1} \otimes E\{z_k z_k'\} \tag{53}$$

and

$$L(\bar{D}_k') = \Gamma_k[\Gamma_{k-1}^{-1} L(\bar{D}_{k-1}') + E[(\Omega^{-1} \otimes z_k)x_k] \tag{54}$$

These modified updating equations eliminate much of the randomness resulting from the dependence of the covariance matrice and the conditional means on uncertain states and uncontrollable exogenous variables. Moreover, the modified equations imply that the approximate updated means are all equal to each other and to the prior mean, i.e., $L(\bar{D}_k') = L(\bar{D}_{k-1}')$ for all k. To obtain this equivalence, substitute

$$\begin{aligned} E(x_k|x_{k-1}) = \bar{D}_{k-1}z_k &= L(\bar{D}_{k-1}z_k) \\ &= L(z_k' \bar{D}_{k-1}') \\ &= (I \otimes z_k')L(\bar{D}_{k-1}') \end{aligned} \tag{55}$$

into, i.e.,

$$\begin{aligned} L(\bar{D}_k') &= \Gamma_k[\Gamma_{k-1}^{-1}L(\bar{D}_{k-1}') + E\{(\Omega^{-1} \otimes z_k z_k')L(\bar{D}_{k-1})\}] \\ &= \Gamma_k[\Gamma_{k-1}^{-1} + \Omega^{-1} \otimes E\{z_k z_k'\}]L(\bar{D}_{k-1}) \\ &= \Gamma_k\Gamma_k^{-1} L(\bar{D}_{k-1}) = L(\bar{D}_{k-1}). \end{aligned} \tag{56}$$

Hence, for the approximate updating equations (53) and (54) lead to replacing the conditional distributions of D by unconditional distributions of indepenpent random variables, D_k with means all equal to the prior mean and with variances and covariances given by (53). Hence, the modification introduced by MacRae leads to an approximate information state which depends upon the evolution of the covariance matrix (53) and the constant prior mean vector $L(\bar{D}_o')$.

Given actual observations on the state vector of the previous period, the adaptive covariance modification leads to an essentially open-loop feedback rule which is derived from the maximization of (50) subject to the stochastic constraints in (49) and the deterministic updating constraints (53). The inclusion of the deterministic constraints also requires the introduction of an associated matrix of Lagrangean multipliers, V_k, which define the prices of additional information. The resulting augmented criterion function is:

$$(57) \qquad \Lambda = E(J) + \sum_{k=1}^{N} (1+\delta)^{-k} \; V_k \circledast [\Gamma_k^{-1} - \Gamma_{k-1}^{-1} - \Omega^{-1} \otimes E\{z_k z_k'\}]$$

$$= E\{J + \sum_{k=1}^{N} (1+\delta)^{-k} \; V_k \circledast (I \otimes z_k)\, \Omega^{-1}\, (I \otimes z_k')\}$$

$$+ \sum_{k=1}^{N} (1+\delta)^{-k} \; V_k \circledast (\Gamma_k^{-1} - \Gamma_{k-1}^{-1}),$$

where the star product, $\circledast$, is a special form of partitioned matrix multiplication.[8] The maximization of (57) may be carried out in two steps. First, for given values of $V_1, \ldots, V_N$, dynamic programming is used to derive conditions that solutions for the vectors $u_1, \ldots, u_N$ and the matrices $\Gamma_1, \ldots, \Gamma_N$ must satisfy subject to the system of stochastic constraints (49). Since the values of Γ_k obtained in this manner (for given V_k) will not, in general, satisfy the updating equation (53), the second step is to select values for matrices V_k such that the updating covariance constraint is satisfied. At the conclusion of both steps, the maximization of Λ will be equivalent to the constrained minimization of $E(J)$.

For period k, the above solution procedure results in the following expression:

$$\Lambda_k(u_k, x_k, \Gamma_{k-1}) = \underset{u_k}{\mathrm{Max}}\ [E\{2\, q_k'\, x_k + 2\, h_k'\, u_k - x_k'\, G_k\, x_k - u_k'\, H_k\, u_k \tag{58}$$

$$+ V_k \circledast (I \otimes z_k)\, \Omega^{-1}\, (I \otimes z_k') - (V_{k-1} - V_k) \circledast \Gamma_{k-1}^{-1} +$$

$$(1 + \delta)^{-1}\, \Lambda_{k+1}\, (\cdots)]$$

Substituting (49) for x_k; expanding in terms of A, B, and C; and partially differentiating (58) with respect to u_k and setting the result equal to zero we obtain the following stragegy rule for u_k in terms of x_{k-1}, Γ_{k-1}, and V_k:

$$u_k = - N_k^{-1}\, (F_k\, x_{k-1} + f_k) \tag{59}$$

where[9]

$$N_k = \bar{B}' S_k\, \bar{B} + S_k \circledast \Gamma_{k-1}^{BB} - \Omega^{-1} \circledast V_k^{BB} + H_k, \tag{60}$$

$$F_k = \bar{B}' S_k\, \bar{A} + S_k \circledast \Gamma_{k-1}^{BA} - \Omega^{-1} \circledast V_k^{BA}, \tag{61}$$

$$f_k = [\bar{B}' S_k\, \bar{C} + S_k \circledast \Gamma_{k-1}^{BC} - \Omega^{-1} \circledast V_k^{BC}]\, \bar{w}_k \tag{62}$$

$$- \bar{B}'\, r_k - h_k,$$

and K_k and R_k are defined as

(63) $$r'_k = g'_k + (1+\delta)^{-1} k'_{k+1}$$

(64) $$S_k = G_k + (1+\delta)^{-1} K_{k+1},$$

with $r'_N = g'_N$ and $S_N = G_N$. The matrices Γ^{BB}, Γ^{BA}, Γ^{BC}, etc., are partitioned submatrices of Γ and the vector k'_k and the matrix K_k enter the "maximum" expected gain Λ_k $(\cdots)$ obtained from employing the u_k controls.

Along with (60) through (62), the optimal value for Γ_{k-1} in terms of u_k, x_{k-1}, V_k, and V_{k-1} is obtained by substituting (49) into (58), applying the expectation operator, simplifying and partially differentiating with respect to Γ_{k-1}. The result is[10]:

(65) $$V_{k-1} = (1+\delta)^{-1}\{V_k + \Gamma_{k-1}[E(I \otimes z_k)S_k(I \otimes z'_k)]\Gamma_{k-1}\},$$

with $V_N = 0$, i.e., V_N must be zero for a finite minimum of Λ to exist. Substituting this result along with (59) into (57) and simplifying, we have[11]

(66) $$\Lambda_k(x_{k-1}, \Gamma_{k-1}) = 2\, k_k\, x_{k-1} - K_k\, x_{k-1} + Q_k$$

where

(67) $$k'_k = r'_k\bar{A} + f'_k N_k^{-1} F_k - \bar{w}'_k(\bar{C}' S_k C + S_k \circledast \Gamma^{CA}_{k-1} - \Omega^{-1} \circledast V^{CA}_k)$$

(68) $$K_k = \bar{A}' S_k \bar{A} + S_k \circledast \Gamma^{AA}_{k-1} - \Omega^{-1} \circledast V^{AA}_k - F'_k N_k^{-1} F_k$$

$$Q_k = f_k' N_k f_k + 2 r_k' \bar{C}_k \bar{w}_k - S_k \circledast \Omega - \bar{w}_k'(\bar{C}S_k\bar{C} + S_k \circledast \Gamma_{k-1}^{CC} - \Omega^{-1} \circledast V_k^{CC}) \bar{w}_k - (\bar{C}'S_k\bar{C} + S_k \circledast \Gamma_{k-1}^{CC} - \Omega^{-1} * V_k^{CC}) \circledast \Gamma_k^W - (V_{k-1} - V_k) \circledast \Gamma_{k-1}^{-1} + \sum_{i=1}^{N-k} (1 + \delta)^{-i} Q_{k+i}, \tag{69}$$

where Γ_k^w denotes the covariance matrix.

To complete the solution to (57), the second step involves selecting matrices V_k so that values for the Γ_k and strategy rules for u_k obtained in the first step will satisfy the covariance update equations. Since V_N must be zero, Γ_N will be determined by the covariance update constraint for the last period. The values of $V_1, \ldots, V_{N-1}$ will be determined implicitly by the covariance constraints for periods 1 through N-1; hence, both $V_1, \ldots, V_{N-1}$ and Γ_N are determined by equation (53). This result along with (59), (49), (53), and (65) requires the solution of a two-point boundary value problem involving u, x, Γ, and V. Equation (59) characterizes the strategy for the policy variable u_k as a function of the state variable, x_{k-1}, given Γ_{k-1} and V_k. The path of x is given by (49) and the matrices Γ and V are functions of the paths of both x and u. The matrix Γ is determined by (53) and V is given by the backward difference equation (65). The net result of the two-point boundary value problem is the determination of matrices Γ_k and V_k and the strategy rules for u_k subject to (49), (53), and (65) given initial conditions x_o and Γ_o and the terminal condition $V_N = 0$.

The solution to the above two-point boundary value problem can be obtained using standard numerical methods. In addition, the solution readily admits some useful economic interpretations and comparisons to other decision strategies. First, the updated certainty equivalence rules are obtained when the covariance Γ_k is zero for every period. For these rules, the price of additional information is zero ($V_k = 0$) since nothing further can be learned. For the stochastic updating rules, Γ_k is not zero and thus neither is V_k. Nevertheless, even though there is some value to increasing the information level for these rules, the settings on the stochastic control variables do not influence the levels of Γ_k and hence the price of information V_k plays no role in the determination of these decision rules.

In the adaptive covariance approximation the relevant stock of information is Γ_k^{-1}. Hence, additions to this stock refer to more efficient estimates of the coefficient matrices A, B, and C. To determine the value of such additional information, we differentiate $\Lambda_k(\cdot)$ with respect to the covariance matrix or stock of information, Γ_k. This is simply the price of information V_k represented by equation (65). Solving this equation recursively for $V_N = 0$ (since no actions are anticipated beyond N) we obtain

$$(70) \qquad V_k = \sum_{j=k}^{N} (1 + \delta)^{-(j-k+1)} \Gamma_j [E(I \otimes z_{j+1} S_{j+1}) (I \otimes z'_{j+1})] \Gamma_j .$$

Thus, as MacRae [1975, p. 901] notes, V_k "is the present discounted value of the stream of future rental on information."[12]

Finally, the term $\Omega^{-1} \circledast V_k$ may be interpreted as the value of estimating, i.e., the benefits derived from changing variables of the model for purposes of improving the estimates (MacRae [1972, p. 443]). As indicated by the first set of equations, the value matrix V never appears alone; it is always multiplied by Ω^{-1}, a matrix which may be interpreted as the amount of information ultimately available for the system (49). The latter matrix, if very "small" may indicate that (49) is too noisy for operational use in either predicting the path of future x or improving estimates of the unknown coefficients. If the value of estimating, $\Omega^{-1} \circledast V$, is "high" then the positive quadratic term $z' (\Omega^{-1} \circledast V) z$ in the augmented expected value function (58) may offset the effects of choosing current policies which are non-optimal from a direct control point of view. On the other hand, if this value is "low", either due to "small" V or a noisy dynamic system (49), less benefits will accrue in deviating from a pure control solution for the purpose of improving the estimated system (49).

6.2 Approximate numerical differentiation controls. This actively adaptive scheme has been developed by Chow [1975] for the case of a quadratic but non-additive criteria function, a linear system (49) and active learning with regard to the unknown parameters. For this specification, a complex nonlinear stochastic control problem is obtained which Chow's method approximates by applying a second order Taylor series expansion in perfectly measured states of the relevant value function. This expansion is taken about some tentative path for the states over the complete planning horizon using numerical differentiation. Chow suggests selecting the initial tentative nominal path on the basis of either the certainty equivalent feedback controls or the sequential stochastic feedback controls. Of course, after solving for the optimal control equations by the approximate second order scheme, a new nominal path of the states can be generated. This new path can be utilized as a second tentative path for computing a new solution. We can continue to iterate in this fashion until convergence of the optimal controls for the first period are obtained.

The specific nonadditive criterion function employed by Chow [1975] is of the form:

$$(71)\quad J = 1/2 \sum_{k=1}^{N} x_k' G_{kk} x_k + \sum_{k=1}^{N} \sum_{s<k} x_k' G_{k,s} x_s + \sum_{k=1}^{N} g_k' x_k + d$$

where $G_{k,s} = G^t_{s,k}$, g_k, and d are known constants. The admission of the cross products of x_k and x_s for $k \neq s$ introduces the nonadditivity. From this expression, the conditional expectation of J, given all the data up to N - 1 can be characterized. The result, $\mathrm{Min}_{u_N} E_{N-1} J$, is approximated by quadratic function in $x_{N-1}, x_{N-2}, \ldots, x_1$. Then $E_{N-2}(\mathrm{Min}_{u_N} E_{N-1} J)$ is minimized with respect to x_{N-1}, and so on until we reach:

$$(72)\quad \min_{u_1} E_o (\cdots\{\min_{u_{N-2}} E_{N-3}[\min_{u_{N-1}} E_{N-2} (\min_{u_N} E_{N-1} J)]\}\cdots)$$

At each step, a second-order Taylor expansion about a nominal path $x_{01}, x_{02}, \ldots, x_{0N-1}$ is taken and the relevant first and second derivatives with respect to $x_1, x_2, \ldots, x_{N-1}$ are obtained numerically.[13] Due to these approximations, the information state for this scheme is approximate by only the first two moments of the posterior probability distributions on the states. However, this approach treats the unknown parameters of linear state transformation equations more directly than the adaptive covariance scheme. Hence, it comes closer to dealing with the actual joint control, estimation problem and thus the experimental component of adaptive control than the MacRae Approach.

A principal limitation of the Chow procedure is the exogenous nature of the nominal path. There is clearly no assurance that the selected control settings will conform to globally optimal controls. However, this scheme can be easily implemented particularly if targets appearing in a quadratic criterion function are prespecified, i.e., the decision problem in essence a tracking problem.

More recently, Chow [1976] has advanced another approximate solution procedure for the case in which the system representation is nonlinear. That is, the criterion function remains quadratic, the parameters are unknown, but the system representation is nonlinear. In the development of this procedure, Chow [1976] assumes a solution exists and advances a simple first-order Taylor series expansion of the system representation to obtain a linear approximation.

For the multiperiod solution, six basic steps are involved in the application of this approximate scheme. First, on the basis of some initial control settings, solve for the initial set of state values for the complete planning horizon by the Gauss-Seidel method. Second, for each period of planning horizon linearize the nonlinear model about the values generated in the first step. Third, compute the relevant expectations by using numerical integration, Monte Carlo methods (by averaging over repeated random drawings from the distributions underlying the unknown parameters), or some other approximate procedure. Fourth, employ the expectations generated from the third step to generate the relevant recursive matrices appearing in the criteria function after substituting the constraints. Fifth, utilize the results of steps one through five to generate the optimal approximate controls for all periods of the planning horizon. Finally, the above sequence of steps is based upon initial states which are employed in the generation of the linear approximation. These initial estimates may deviate substantially from the true ones and thus some improvement may be achieved through recomputing these values by applying the controls obtained in step five. Utilizing these solutions to generate new estimates of the initial states, steps one through five can be repeated once again. We can continue in this fashion, of course, until some measure of convergence is achieved.

6.3 Adaptive "Closed-Loop" Controls (TBM). The most widely publicized actively adaptive scheme is based on the work of Bar-Shalom, Tse, and Meier [1973], Bar-Shalom and Tse [1975], Bar-Shalom and Tse [1976], and Tse and Bar-Shalom [1973]. In essense, the wide-sense dual control procedure advanced by these authors decomposes the complete adaptive control problem into three components (a) current control, (b) future deterministic controls, and (c) a future perturbation control. The perturbation or experimental control component is partioned into a caution and probing term. The caution term reflects the effects of uncertainty at time k and subsequent process noise on the criterion function. The probing term summarizes the effect of uncertainties when subsequent decisions are made.

If current decisions through experimentation can reduce future updated parameter's covariances, they also reduce the cost appearing in the objective function. A weighting matrix is included in this term which yields an approximate estimate of the value of future information for the problem under consideration. The sum of the caution and probing terms provide a basis for evaluating the future benefit of learning and when compared with the other components a mechanism is established for allocating between this benefit and the cost of current control settings.

To provide some flavor for this scheme, consider the general model specification advanced in Section 2. Problems arising from information states which are either infinite dimensional or finite but grow with time are dealt with by utilizing only the first two moments of the state estimates. In other words, the actual information state is approximated by

(73) $$\Psi_k = \{\hat{x}(k|k), \Sigma(k|k)\} ,$$

i.e., the updated conditional mean and covariance of x_k.[14] These updated estimates can be computed by any one of a number of approximate methods including the extended Kalman filter (Sorenson [1966], Jazwinski [1970]), the adaptive filter (Jazwinski [1970], Tse, et al. [1973]), second order filters (Tse et al [1973]), or the optimum filter (Bucy and Senne [1970], Alspach and Sorenson [1970]).

The optimal future benefits or costs-to-go are stated in terms of the above approximate information states. Specifically, invoking Bellman's [1961] principle of optimality, we obtain the following characterization for the closed-loop-optimal expected cost-to-go at time k

$$(74) \qquad \Lambda_k^*(\Psi_k) = \min_{u_k} E\{C[x(k), u_k, k] + \Lambda_{k+1}^*(\Psi_{k+1}|\Psi_k)\}.$$

In general, it is not possible to obtain an explicit expression for $\Lambda_{k+1}^*(\cdot)$ which preserves the desired closed-loop feature and thus adequately reflects the "value" of future observations. Hence, the TMB approach transfers what is in essence a nonlinear joint control estimation problem for $\Lambda_{k+1}^*(\cdot)$ to an approximate linear-quadratic problem. This transference is made by noting that associated with each control setting u_k there is a future (fictitious) nominal control sequence u_{oj}, $j = k + 1, \ldots, N - 1$. The choice of the nominal is quite flexible and dependent upon the class of problem under consideration. There is, of course, a different nominal associated with each u_k and the sole purpose of the nominal is to obtain an approximate value of $\Lambda_{k+1}^*(\cdot)$ associated with the information state at the next stage. These nominal controls are then employed to generate a nominal state sequence $[x_{oj}, j = k + 1, \ldots, N - 1]$ about which a second order perturbation analysis is carried out. In this fashion an approximate otpimal cost-to-go is obtained that explicitly reflects the future learning and control performance that can result from alternative settings of u_k.

The approximation to $\Lambda^*_{k+1}(\cdot)$ follows from

(75) $$C(k+1) = C_o(k+1) + \Delta C_o\ (k+1)\ ,$$

where

(76) $$C_o(k+1) = \sum_{j=k+1}^{N-1} C[x_{oj}, u_{oj}, j]\ ,$$

i.e., the cost along the nominal; and $\Delta C_o(k+1)$ is derived as the variation of the cost about the nominal and thus is stated in terms of the deviation of actual states and controls from their respective nominals. The corresponding result $\Lambda^*_{k+1}(\cdot)$ is: [15]

(77) $$\Lambda^*_{k+1}(\Psi_{k+1}) = \min_{u_{k+1}} E\{\ldots \min_{u_{N-1}} E[C(k+1)|\Psi_{N-1}]\ldots|\Psi_{k+1}\}$$
$$= \Lambda_o(k+1) + \Delta\Lambda^*_o(k+1),$$

where

(78) $$\Lambda_o(k+1) = C_o(k+1)\ ;$$

(79) $$\Delta\Lambda^*_o(k+1) = \min_{\delta u_{k+1}} E\{\ldots\min_{\delta u_{(N-1)}} E[\Delta C_0(k+1)|\Psi_{N-1}]\ldots|\Psi_{k+1}\}\ ;$$

and

(80) $$\delta u_j = u_j - u_{oj}\ .$$

Manipulating $\Delta\Lambda^*_o(k+1)$ and operating with terms up to second order, TBM are able to illustrate the closed-loop property of $\Delta\Lambda^*_o$, i.e., its relationships with future uncertainties, along with its existence and uniqueness.

Substituting this result along with $\Lambda_o(k+1)$ into the expression for $\Lambda_k^*(\Psi_k)$ and dropping all terms which do not depend on u_k, they obtain:

$$\Lambda^{CL}(k) = \Lambda_D(k) + \Lambda_C(k) + \Lambda_P(k) \quad . \tag{81}$$

On the basis of this representation their approximate "closed-loop" adaptive control solution is obtained. The first term, $\Lambda_D(k)$ incorporated deterministic current effects along with future (deterministic) costs along the nominal, i.e., $C_o(k+1)$. The next two terms incorporate various elements of $\Lambda\Delta_o^*(k+1)$. The term $\Lambda_C(k)$ reflects the effect of uncertainty at time k and subsequent process noises which should be minimized. Thus, this term may be referred to as the caution component. The term $\Lambda_P(k)$ summarizes future uncertainty effects associated with decisions over the period k+1, ..., N-1. If control settings can reduce future updated covariances which compose this term, then experimentation or probing must be given serious consideration. This term is referred to as the probing component and it must be balanced with the deterministic and caution components of the criterion function. This balance will be achieved in the minimization of $\Lambda^{CL}(k)$ by properly assessing the tradeoff between present and future actions in accordance with the information available at the time the corresponding decisons are made.

To find the TBM closed-loop control u_k, the minimization of (74) is performed using a search procedure.[16] The basic steps of this search first select a value or values of u_k, then evaluate explicit functional relationships between u_k and the future covariance matrices, and finally use the resulting functional values to select the next values of u_k to examine. This scheme proceeds until satisfactory convergence is obtained.

In addition to the search method, the approximate cost-to-go must be computed and this requires the derivation of (1) the nominal path; (2) a second order Taylor series expansion about this nominal path; (3) a solution of the perturbation problem for the approximate otpimal cost-to-go; (4) a partial solution for a set of difference equations to provide a clear separation of the stochastic terms from the non-stochastic terms appearing in the approximate cost-to-go; (5) a deterministic approximation for the cost-to-go; (6) projections of the covariance matrices; and (7) an updating process for the covariance matrix.[17] The complete algorithm begins at period k with estimates of the relevant updated mean vector and covariances. A search is then begun to find the optimal control u_k to be imposed in period k. For each choice of u_k the optimal cost-to-go function is evaluated and this process is continued until satisfactory conversion is obtained. Clearly this adaptive scheme is computationally expensive, and for some systems this computational expense is prohibitive.

Finally, although both the TBM approach and the Chow or numerical differentiation control schemes are closed loop approximations, they differ in a number of important respects. First, the TBM procedure treats the unknown parameters by augmenting them to the state vector while the Chow approximation does not. Second, the TBM approach makes the computation of the normal path endogenously, i.e., the nominal trajectory for the entire planning horizon is a function of the current controls. In contrast, the Chow procedure treats the compuation of the nominal path exogenously.

Third, the Chow approach approximates the entire otpimal expected cost-to-go by a second order expansion about a tentative path after taking expectations using numerical derivatives, while the procedure advanced by TBM applies the second order expansion prior to taking the expectations. The latter approach allows the derivation of an explicit representation for the perturbation costs.

6.4. M-Measurement Feedback Controls. A number of variations of the above two schemes can and have been developed. One of the most important is concerned with the notion of M-measurement feedback control. This approach, developed by Curry [1969], has been analyzed by Rausser and Freebairn [1973], Early and Early [1972], and Tse and Bar-Shalom [1973]. As previously noted, it represents an intermediate approach to the above adaptive schemes by assuming that each period, new information about the systems will become available only at some M future stages of the controlling horizon.

Assume that the relevant stages are the M successive periods in the immediate future, i.e., $k + 1, \ldots, k + M$. The M-measurement feedback controls (u_k^M, $k = 1, \ldots N$) are then those policy decisions which utilize all past and present information as well as the knowledge that over the next M periods new information (observations) will become available. More formally, these controls require the specification of an integer (I_k) which is the smaller of M and the remaining periods in the planning horizon, i.e.,

$$I_k = \min \ \{M, N - k - 1\}, \quad k = 1, \ldots, N. \tag{82}$$

In other words, this integer is equal to M if the *assumed* number of future observations taken into account is less than the *actual* number of future observations which will become available during the controlling process.

Given I_k, it is possible to specify the M-measurement feedback controls as generalization of any other passive or active scheme. For example, u_k^M reduces the sequential stochastic controls (Section 5.3) for $M = 0$ if u_k^M along with $u_{k+1}^M, \ldots u_{k+I_t}^M$ are determined by

$$(83) \qquad \Lambda_{Mk}(x_{k-1}, \Psi_{k-1}) = \underset{u_k \ldots u_{k+I_k}}{\text{Max}} \; [E\{ \sum_{i=k}^{k+I_k} C_i\ (x_i, u_i, i)$$

$$\Lambda_{ss},\ _{k+I_k+1}\ (x_{k+I_k}, \Psi_{k+I_k} \mid \Psi_{k+I_k+1}, u_k^{k+I_k}, w_k^{k+I_k})\} \mid x_{k-1}, \Psi_{k-1}].$$

where $u_k^{k+I_k} = (u_k, \ldots, u_{k+I_k})$, and $w_k^{k+I_k} = (x_k, \ldots, x_{k+I_k})$. For the u_k^M controls resulting from (83), two limiting forms are immediate. First, if $M = 0$, the availability of future information is neglected and thus u_k^M is equivalent to the sequential stochastic controls. Second, if $M = N - k - 1$, the availability of future information over all remaining periods of the planning horizon is taken into account and thus u_k^M is equivalent to the adaptive or dual control solution for period k. In addition, note that if the second term on the right hand side of (83) is appropriately modified, the u_k^M controls for $M = 0$ could reduce to any of the previous control schemes.

As with most other adaptive schemes, the dependence of future covariances on current control is nonlinear and quite complicated for the M-measurement feedback control procedure. Even for the case of $M = 1$ this procedure involves a search in an h-dimensional space where h is a dimension of the control vector. Unless the control space is bounded, this search will result in a local minimum with respect to u_k. Thus, as with the other adaptive schemes discussed so far, numerical procedures are required.

6.5 Two-Step Adaptive Control. An approach motivated by the notion of M-measurement feedback control has been advanced by Sarris and Athans [1973] which they characterize as the two-step optimal adaptive control. The optimization for this scheme pertains to only two more periods and imposes the assumption that the unknown parameters are constant and equal to the predicted magnitudes of these parameters. On the basis of this assumption a control law is derived which is a highly nonlinear function of the state variables. The interesting feature of this scheme is that it takes into account future adaptation of the conditional means but not the variances and covariances associated with estimates of the unknown parameters. This dual control scheme has the advantage of not involving the solution to any iterative system of equations that are encountered with most of the other methods.

6.6 First Order Schemes. Due to the complexity of the above schemes, especially the TBM scheme, some interest has emerged in further simplifications of these schemes for large models. In particular, Norman [1976] has introduced a first-order Scheme which involves certain simplifications of the TBM scheme. Like the TBM approach the first order scheme decomposes the problem into a current control, future deterministic control, and future perturbation control components. However, in contrast to the TBM approach the first-order scheme neglects all second-order terms involving the interaction between random parameters and states in the derivation of the future perturbation component. In essence, this simply means that the basic difference between the first-order scheme and the TBM procedure is that the former contains no covariances of the states and unknown parameters.

Two versions of the first-order scheme have been investigated by Norman [1976]. The first augments the original state vector with the unknown parameters while the second does not. The second version has some rather obvious advantages, particularly in the computation of the relevant Ricatti matrices which are cubic in the number of states. Hence, the nonaugmented first-order scheme incurs smaller computational which indeed may be significant for larger models.

6.7 Other Actively Adaptive Schemes. Other actively adaptive schemes have also been examined in the literature. For example, Norman [1976] has examined another modified version of the TMB approach. This second order schemes simply involves a non-augmentation of the state vector by the unknown parameters. The reason for this simplification is that for many problems augmentation of the state vectors by the unknown parameters can prove to be a costly procedure and fail to result in improved performance.

Most of the remaining schemes available in the literature generally apply to specialized structures. For example, Aoki [1967], Saridis and Lobbia [1972], and Wieslander and Wittenmark [1971] have advanced an adaptive control law based upon a step-by-step minimization of

$$J(\underline{u}_k) = E\{C^1_{k+1}(\underline{x}_{k+1}) + C^2_{k+1}(\underline{u}_k)/Y^k, U^{k-1}\} \tag{84}$$

where the cost functional $C_k(\underline{x}_k, \underline{u}_k)$ may be decomposed into $C^1_{k+1}(\underline{x}_{k+1})$ and $C^2_{k+1}(\underline{u}_k)$. In other words, their results are based on a criterion function which is separable across the states and controls. For this one step problem the availability of future information is irrelevant in the computation of u_k. In this sense the method is similar to the open loop feedback method. This approach has been augmented by Alster and Belanger [1974] to admit dual control. In particular they guarantee some level of learning by minimizing (84) subject to the constraint

$$E\{\mathrm{tr}\, \underline{P}^{-1}_{k+1/k+1} / Y^k, U^{k-1}\} \geq \gamma_{k+1} \tag{85}$$

where $\underline{P}_{k+1/k+1}$ is the updated covariance of the states based on information up to time k+1. Of course, a difficult problem is faced in appropriately specifying γ_{k+1}.

Murphy [1968] and Gorman and Zaborsky [1968] have also investigated a specialized structure of the adaptive control problem, viz. where only the parameters associated with the controls are unknown. Moreover, a number of other methods, often appearing under the heading of self-organizing control, are available. For these methods, the control scheme is treated as a fixed structural form with some adjustable parameters. In general, the parameters are adjusted adaptively by one of two alternative procedures. The first adjustment procedure is based on improved learning of the system parameters and is often referred to as parameter-adaptive schemes (Saridis and Stein [1968], Saridis and Dao [1972], and Tsypkin [1973]). The second procedure adjusts the parameters such that the control settings will improve some "subgoal" generated from the criterion function and has been referred to as performance-adaptive schemes (Saridis and Gilbert[1970], Saridis [1971], and Tsypkin [1973]). For the first procedure an attempt is made to estimate the true parameters of the system while for the second procedure it is the control law's feedback parameters that must be estimated.

Other adaptive schemes utilize second-order perturbation analysis and have been referred to in the literature as trajectory-shaping (Denham [1964], Meier [1965], and Van der Stoep[1968]). Much of this literature is a simple modification of the linear quadratic Gaussian approach. However, instead of employing the deterministic trajectory, another trajectory which minimizes the cost criterion obtained by utilizing second-order analysis is selected as the reference trajectory. This approach shares similar advantages and disadvantages with the linear quadratic Gaussian method discussed in section 5.4.

Still another approximate suboptimal adaptive control schemes have been advanced in the literature. For example, Deshpande, Upadhyay and Lainiotis [1972] have extended the work of Stein and Saridis [1969] and Saridis and Dao [1972] by providing a computationally simpler scheme. Specifically, for a quadratic criteria function and a linear representation of the system, these authors specify a method for computing the *a posteriori* probabilities which incorporate some elements of learning. It is assumed that the unknown parameters have a discrete probability distribution which is adapted with each new observation. For each particular selection from this discrete probability distribution there corresponds a unique optimal control. These authors suggest weighting the controls by the appropriate probability values and computing a weighted measure of the approximate optimal control.[18]

Finally, another approach which is particularly useful when the underlying probability distributions are non-Gaussian or nonunimodal has been developed by Alspach [1972]. Following some earlier work, he utilizes the Gaussian sum approximation to the required a posteriori densities generated by adaptive control approach. Based upon these approximations suboptimal adaptive controls can be computed. Unfortunately, the procedure is not feasible in real time or even off-line if the number of terms in the a posteriori densities by the Gaussian sum method and the curse of dimensionality. For these reasons, Alspach suggests employing a version of the M-measurement feedback control approach. In particular, Alspach's approach would be utilized for stages one through M to approximate the a posteriori densities and for the remaining stages of the planning horizon (namely, M+1 through N), some other simple scheme such as certainty equivalence can be utilized. One of the advantages of the procedure is that instead of operating with estimates of means or covariances, an approximation is made on the entire a posteriori density function.[19]

6.8 Comparative Performance of Adaptive Control Schemes. A large number of comparisons of some of the alternate suboptimal schemes are available but almost all of these comparisons are based on rather questionable Monte Carlo simulation designs. This includes the work of Bar-Shalom and Tse [1976], Chow [1975], MacRae [1975], Norman [1976], Prescott [1971], Sarris and Athans [1973], and Zellner [1971]. Generally this work examines only the case of a scalar-state and control variable, a quadratic criterion function, a linear plant representation, and few if any of these studies employ any of the useful results from the statistical literature on sequential experimental designs. Moreover, due to the computational complexity and resulting costs of the various methods, the number of Monte Carlo simulations carried out are generally meager. Furthermore, the data base used for comparisons among various authors generally differ and thus realistic comparisons of all the alternative suboptimal adaptive schemes are generally non-existent. Nevertheless, some comparative performance results are available which are worth briefly summarizing here.

Most of the results obtained thus far conform to a priori expectations. For example, Prescott [1971] compares three alternative schemes: two passive adaptive schemes and the dual control scheme. The first passive scheme is a certainty equivalence approach while the second is a myopic decision rule which minimizes the criterion function only for the current period and thus completely neglects the future. He finds that the certainty equivalence approach is a reasonable procedure when uncertainty in the unknown parameter is small; specifically when the ratio of the prior's mean to its standard deviation is at least 4 in absolute value. When the ratio of the prior's mean to its standard deviation is smaller than 2, however, experimentation becomes a relevant consideration; i.e., it pays to select a decision larger in absolute value than the one which minimizes current expected loss in order to obtain additional information about the unknown parameter. Another result is that the more periods remaining in the planning horizon, the more important is experimentation.

MacRae [1972,1975] compares three alternative schemes, viz. certainty equivalence, the sequential stochastic, and the adaptive covariance controls. For a dynamic model (i.e., current states depend upon lag states) in which the parameters of the system representation associated with the lag states are known, it is possible for the adaptive covariance scheme to geneate a more conservative current policy than the sequential stochastic rule. This result for the adaptive covariance scheme simply reflects the possibility that the best policy may well be to do a very little at first to avoid the relatively large cost of uncertainty and subsequently compensate later when the effect of the policy is known with more precision. As MacRae [1972, p. 903] points out, "this paradoxical response to the introduction of learning possibilities can occur only in dynamic models and appears more likely with longer planning horizons." This result is modified, however, if the parameter summarizing the relationship between lag and current states in the system representation is unknown. In this event, parameters associated with current controls and with lag states can be correlated and thus no general implications can be drawn about the relative magnitudes of the policy variables for the three types of schemes examined by MacRae. For this specification, although larger variances imply larger uncertainty involving the model parameters, larger covariances imply more information. In addition, there is no particular reason to presume that either the variances or covariances (both of which appear in the sequential stochastic and adaptive covariance rules) will dominate.

Several Monte Carlo simulations carried out by Tse and Bar-Shalom and their colleagues at Systems Control have evaluated the comparative performance of the TBM method with the certainty equivalence rule. For most cases examined by these authors, the TBM method gave substantial improvement over the certainty equivalence rule. In contrast, Chow [1975, 1976, pp. 267-76] reports a comparison among certainty equivalence, sequential stochastic and the Chow active adaptive scheme for both a single equation and a two-equation system representation. He finds that the various rules do not lead to numerical values of the controls which differ to any substantial degree. An explanation of these apparent conflicting results is that the model representation employed in the Chow comparisons are more precisely estimated than those examined by Tse, Bar-Shalom, and others. In particular, for the former model representations, (i) the standard errors of the estimated coefficients are small relative to the prior estimates of the unknown parameters and (ii) the sample period utilized to obtain prior estimates of the unknown parameters is long compared with the planning horizon.

More recent results involving a comparison of the TBM method, the Chow actively adaptive scheme, the adaptive covariance scheme, and such passively adaptive schemes as open loop feedback and certainty equivalence have been reported by Bar-Shalom and Tse [1976]. For the same dynamic model examined by MacRae [1972] they found that the more sophisticated actively adaptive schemes did not always perform better than the open loop feedback or even the certainty equivalence rule. Similar small differences among the performance of the various algorithms were found by Sarris and Athans [1973]; in particular they did not obtain a distinct ordering among the certainty equivalent and sequential stochastic rules.

Two other comparative performance studies by Norman (1976] have addressed the trade-off between computational complexity and performance. Utilizing the same model as MacRae [1975] and Bar-Shalom and Tse [1976], he compares augmented and non-augmented variations of the first and second order perturbation control schemes along with heuristic certainty equivalence and open loop feedback. His results clearly indicate that the performance of the alternative schemes is problem-specific and for some cases parameter estimation error can lead to poorer performance for the first-order scheme than the open loop feedback rule. Finally, when computational cost is explicitly considered, Norman finds some specifications for which the open loop feedback as well as the certainty equivalence appear to be the most desirable schemes.

7. Applications

There have been a large number of passively adaptive applications in management science, especially in the context of the open-loop feedback control approach (e.g. Ying [1967]). One of the principal areas of these applications is operations management. There are literally myriads of passively adaptive schemes employed in automated planning and control systems. Most forecasting-inventory systems utilize an open loop feedback framework to update cost and demand parameters in computing lot-size formula and buffer stock levels. These and other passively adaptive applications are far too numerous to survey here; instead we shall concentrate on application of actively adaptive control schemes.

Since the empirical applications of active computationally feasible algorithms for the general adaptive control system were constructed only in the last few years, few schemes have appeared in the literature. They include Pekelman and Tse [1975], Little [1966, 1977], Chong and Cheng [1975], Bar-Shalom and Stengel [1976], Rausser and Freebairn [1974], Freebairn and Rausser [1974], Rausser and Hochman [1977], Kendrick and Majors [1974], Able [1975], Chow [1975], and Upadhyay [1975]. In what follows, the basic structure of each of these applications will be briefly outlined.

7.1 Advertising - Pekelman and Tse [1975], Little [1966, 1977]. In Pekelman and Tse [1975] the problem is to determine the optimal advertising expenditures which will achieve the appropriate amount of learning (experimentation) about the parameters as well as the control in terms of discounted profit. A sales-advertising relationship represents the plant equation. The selected specification, based upon regional data, states current sales in terms of lagged sales and the ratio of firm advertising to total industry advertising. Given some additional relationships for industry sales and competitive advertising along with the temporal evolution of the sales-advertising parameter, the adaptive control rule is derived using a slight modification of the TBM method.

Various simulations demonstrate the superiority of the adaptive scheme when the sales-advertising parameter is changing over time. Analysis of the sales-advertising system offers the rationale for the specific form of the learning policy; in particular, the significance of improving the estimation of different parameters over time.

Scenarios when the adaptive scheme is not materially better are also investigated. For example, when the noise is high either in the advertising-sales equation or in the competitive advertising equation, all the schemes, including the open-loop scheme, perform equally bad. In scenarios where the prior estimation is poor, but the errors in the mean value of the various parameters may compensate for each other, the dual control scheme is not necessarily superior to the certainty equivalence.

Little's [1966, 1977] formulation appeared first in a scalar form in 1966, and was later extended to the multivariate case. His control problem, ignoring estimation, is static, and its unique characteristic is the form of the experimental design.[20] The utilization of information is equivalent to m=1 measurement control. For a linear criterion function, and a quadratic state equation relating sales to promotion (with random effects), simulation experiments are advanced to compare various control rules; viz., truly optimal, constant promotion, optimal adaptive, adaptive with half of the optimal size experiment, and a simplified adaptive rule. The results show that the performances of all adaptive schemes are approximately the same while they are considerably superior to the constant promotion rules. The relative performance of the adaptive rule is shown to be stable when k is changing.

In Little's multivariate extension there are r controls, and the experimental design is a 2^r factorial problem. The profit function, after substitution of the state equation, becomes the sum of a linear and quadratic forms in the control variables, and again only the coefficient of the linear term is assumed unknown.[21]

7.2 Pricing. Chong and Cheng [1975], Bar-Shalom and Stengel [1976]. In this work a linear demand equation is postulated with unknown parameters. Expected profit is maximized subject to a linear demand equation. For Chong and Cheng [1975], the problem is essentially static since current demand does not depend upon lagged demand. The only connection between different periods is due to the effect of present decision variables on the accuracy of subsequent estimates. Moreover, the monopolist envisaged by this framework is assumed to have linear risk, constant marginal cost, finite planning horizon and instantaneous production capabilities.

Chong and Cheng show that if the only unknown parameter is the slope of the demand curve, the optimal solution is the one period solution. Except for this special case, the multiperiod nature of the problem has important effects on the pricing policy of the monopolist. In particular, additional data collected by the monopolist can always be employed to update his information on the demand relationship. To approximate the resulting active learning framework, these authors use the scheme developed by TBM. They subsequently compare the results of this approximate control rule with a certainty equivalence scheme and find via a number of simulation experiments that active learning clearly improves the performance of the monopolist.

Bar-Shalom and Stengel approach the problem from a slightly different direction. They devise a scheme which expands the value function, after expectation and minimization, around the expected values of the unknown parameters. They demonstrate that perfect knowledge of the parameters is not required to obtain the optimal decision. This result is due to an "indifferent region" in which the parameter estimates can lie without affecting the performance index. This result is demonstrated both within the context of a Bayesian and a non-Bayesian framework.

7.3 Trade, Pricing and Inventory Reserves (Rausser and Freebairn [1974a, 1974b], Freebairn and Rausser [1974], Rausser and Hochman [1977]. In Rausser and Freebairn [1974a], governmental control of U.S. beef imports are examined in the context of consumer welfare, beef breeding producer welfare (cow-calf operators), and cattle feeder welfare. The weights associated with these welfare measures in (criterion function) were quantitatively estimated using a revealed perference approach and the notion of policy leverage points (Rausser and Freebairn [1974b]). Since the welfare measures depend upon meat product prices, demand, supply, production, and cattle inventory levels which in turn depend upon beef import levels, the control variable, an econometric model of the U.S. livestock sector was constructed (Freebairn and Rausser [1974]). The specific components of this model include (i) consumer meat demand, (ii) margin and producer meat prices, (iii) cattle producer supply and inventory behavior, (iv) pork producer supply and inventory behavior, and (v) poultry producer supply and inventory behavior. Prior estimates of the unknown parameters appearing in the resulting twenty-nine equation model were updated by applying the extended Kalman filter to obtain the relevant probability distributions at the beginning of each period of the planning horizon.

In the above setting, a number of schemes were compared, including certainty equivalence, sequential stochastic, adaptive covariance, and m-measurement feedback (m=2) control schemes. In general, the latter two schemes led to more (less) extreme settings of the control levels in the first (last) few periods of a given planning horizon that the sequential stochastic controls. Furthermore, the adaptive covariance and m-measurement feedback approximations generally resulted in (i) control settings which exceeded in some cases by substantial amounts solutions obtained for the certainty equivalent approximations and (ii) expected gains which exceeded all other control strategies. These results suggest that the beef trade public decision makers should find it beneficial to incur the cost of learning by substituting knowledge accumulation in current periods for expected gains at some later date. To assist public decision makers in this regard, the study found that among the control schemes examined, the m-measurement feedback control scheme led to the best performance.

A similar approach may be found in Freebairn and Rausser [1974], where the decision maker is the Australian government. The government's objective is to obtain a desired stock of wheat which will provide a reserve to satisfy export demand (trade) and domestic demand. The framework involves the maximization of consumer and producer welfare subject to state equations for the level of inventory and the production of wheat, where the price paid to producers is a control variable set by the government. Since data allowed rather precise estimation of the plant equations (production response and demand) appearing in this model, only certainty equivalence and sequential stochastic control strategies were investigated. The performance of the latter strategy far exceeded the performance resulting from applying the certainty equivalence controls.

In the context of commodity marketing boards, Rausser and Hochman [1977] extend the earlier work on pricing discussed in section 7.2. In these studies the control variables include output prices, production quotas distributed among various producers, and for some applications the amounts produced and sold. In contrast to the earlier work on pricing, nonlinear risk is allowed and the possibility of inventory accumulation is explicitly admitted. Certainty equivalence, sequential stochastic, adaptive covariance, and m-measurement feedback controls are compared in these empirical applications with the m-measurement feedback control scheme generally providing the best results; in some instances significantly improved performance over the remaining schemes.

7.4 Macroeconomics. (Kendrick and Majors [1974], Able [1975], Upadhyay [1975], Chow [1975]). In Kendrick and Majors, the basic macroeconomic model developed by Pindyck [1973] (linear system with a quadratic objective) is examined. It is assumed that one of the parameters is unknown. Three schemes are compared; viz. an open loop, closed loop assuming the parameter is known, and an approximated close loop corresponding to the linear quadratic Gaussian scheme presented in section 5.4. The structural equations of the linear econometric model used by Pindyck [1973] have also been examined by Upadhyay [1975] employing the scheme developed by Deshpande, Upadhyay and Laniotis [1972].

Able [1975] examines the performance of three control rules which differ in their treatment of learning. The system consists of two equations; the controls are fiscal and monetary policy and the objective function is a quadratic welfare function. The three investigated policies are certainty equivalence, the sequential stochastic rule which recognizes the uncertainty in the parameters but ignores the possibility of learning (section 5.3), and the active scheme of Chow (section 6.2). His results imply that for the linear system the optimal first-period policy is insensitive to the uncertainty in the parameters and the specific control rule.

Chow [1976] has presented an illustrative application of his nonlinear, active scheme to the Klein-Goldberger macroeconometric model. A principal interest in this illustrative application was the computational efficiency of the scheme for a large, (23 structural equations) nonlinear model. For a ten period planning horizon, five target variables, and four control variables, only 32.7 seconds was required on an IBM 360-91 computer. This result suggests that is not computationally prohibitive to incorporate uncertainty by essentially evaluating numerical derivatives.

8. Concluding Remarks and Research Issues

The various adaptive control schemes presented above have been compared in terms of their utilization of information and the nature of the approximations they impose. As we have seen existing simulation experiments do not offer any general conclusions about the relative performance of these schemes. Moreover, the recently developed approximation schemes are not based on arguments which imply improved performance, but instead they are motivated by numerical considerations. Unfortunately, no conceptual framework has been advanced on which to base an evaluation and comparison of the various schemes. Currently, the selection of a particular adaptive control method for a particular application can only be determined on a trial and error basis.

A conceptual framework is required which provides a clear exposition of the interaction between learning and control in a stochastic, dynamic setting. Our principal interest in the development of such a framework should be to obtain (i) measures of the learning capability for various adaptive schemes and their relationship with performance; (ii) relationships of the structural properties of the optimal control scheme with the system structure, observation program, and performance criterion; (iii) a clear indication of how information flows through the system are influenced by the control scheme and how both are related to overall system performance; (iv) formal measures of the trade-off between computational complexity of the various schemes and their performance; and (v) tight bounds on the performance of alternative schemes. An initial attempt with regard to (i), (ii), and (iii) has been advanced by Tse [1976]. In this work, the learning capabilities of various control laws are formalized in terms of Shannon's information measure. Furthermore, Faden and Rausser [1976] advance a framework which can be employed to formally investigate (iv) and (v).

To be sure, for the number of applications of adaptive control to increase, the development of numerically simpler schemes with high capability of learning is critical. Most of the existing schemes which are operational involve some form of numerical search. The search at times may be lengthy, due mainly to the non-convexity of the value function and the existence of multiple local extrema, e.g., Pekelman and Tse [1975]. Hence, the introduction of multiple controls would require for most active schemes a multidimensional search which may be numerically prohibitive.

In addition to the numerical issue, the development of new approaches should also consider aspects common to both management science and economic problems. The first and foremost element missing in the existing approaches is the capability of handling inequality constraints on the controls as well as the states. Although constraints on state variables, in general, can be assured to hold only in probability their incorporation would enable us to solve the common resource allocation problem in an adaptive form where the coefficients of the return function as well as those specifying the technology could be continuously learned. Another element is concerned with whether or not the system or plant representation is in structural form (current states depending upon other current state variables) or reduced form. All the adaptive control frameworks available in literature assume the latter form but unfortunately many models in management science and economics are more naturally stated in their structural form. Furthermore, when these structural forms are highly nonlinear it is not possible to derive a corresponding, unique reduced form.

Another issue of importance is model discrimination. A common criticism of adaptive control is that although significant parameter uncertainty may indicate the potential value of an adaptive control scheme, it may also indicate that the model is misspecified. Clearly, identification of the model structure is essential for any quantitative application, but in the adaptive control framework, the experimentation may be directed toward the discrimination between competing models as well as learning of their parameters. Some work has been accomplished in this area (Saridis and Stein [1968], Saridis and Dao [1972], Athans and Wilner [1973]) but much remains to be done. In particular, when different model representations appear equally possible the appropriate adaptive control framework should be couched in terms of a simultaneous hypothesis testing and control problem. This is an especially difficult problem when a single model must be selected from a class of non-nested models and where a convex combination of models is not allowed.

A related issue is the impact of model misspecification, e.g., omitted variable, functional form, on the control and prediction performance of the various schemes. In the econometrics literature the reestimation of parameters is usually justified (even for constant parameters) by its ability to capture some of the model misspecification, and hence reduce the prediction error. However, it is not clear that control performance is also enhanced by this procedure. Here again, much in way of additional conceptual work on adaptive control schemes would appear to be necessary.

In the applications sphere, a number of management science problems readily lend themselves to the adaptive control framework. These include, as previously noted, new product development; effectiveness of new employee incentive schemes; output-maintenance relations for new machines; the production of new technology; and systems where the parameters tend to vary over time. Other potentially rewarding applications of adaptive control include <u>inter alia</u> (i) the selection of private or public projects where the rate of return is unknown, but the selection process repeats itself; (ii) promotion policy where a periodic decision is made on offering discounts, free samples, and the like; (iii) periodic market sample surveys to determine likely consumer responses to product specifications; (iv) public regulatory management where periodic monitoring decisions must be made to determine compliance and private sector response to regulatory controls (Rausser and Howitt [1975]); and (v) periodic agricultural and natural resource management decisions concerned with the effect of alternative input combinations and land allocations on uncertain plant growth. The adaptive control framework indeed appears promising but much additional evidence should be collected before its true value as a research tool for management science is ascertained; clearly the jury is still out.

ACKNOWLEDGMENT

The authors wish to express their appreciation to Paul Kleindorfer for his penetrating comments on an earlier draft of this paper. Any errors or misinterpretations appearing in this paper may be regarded as his responsibility.

FOOTNOTES

1. The first appearance of a formal definition of 'adaptive control system' may be traced to Tsein [1954] in 1954. But as late as 1965 the term 'adaptive control system' in the engineering literature was used to describe systems with controllable parameters (Donalson and Kisch [1965]). The output of the controlled system in the framework advanced by these authors is compared to some desired output and the parameters of the system are modified such that the system output approaches the desired output (see Landau [1974]). As we shall see, this framework misses some important features of the adaptive control framework currently accepted in the literature.

2. The adaptive control approach corresponds to Bellman's [1961], pp. 198-290] as well as Feldbaum's [1965], ppl 24-31] third class of control systems. This class is characterized by some unknown quantities about which uncertainty changes by learning as the process evolves. It includes the active accumulation of information.

3. Formal procedures for utilizing sample additions to update probability distributions specifications include Baysian methods (Zellner [1971]), least squares revision methods (Albert and Sittler [1966], Kalman type filters e.g. Kalman [1960]) as well as other methods.

4. In fact, the optimum u_1 is determined by the ration q/σ_o and σ_o. Furthermore, the result that both large and small q imply that u_1^* is approximately $-\alpha_o$ is easily seen by examining the optimal decision rule associated with (28). It is:

$$\alpha_o + u + \frac{\sigma_o^3 u^3 - \sigma_o^2 u (\sigma_o u^2 + q)}{(u^2 \sigma_o + q)^2} = 0$$

For both $q \to \infty$ and $q \to o$, the third term approaches zero, and hence $u_1^* = -\alpha_o$.

5. Under certain well-known conditions, stated first by Simon [1956] and Theil [1957], the C.E. solution is optimal. Under these conditions, nothing can be gained by utilizing information on future measurements and uncertainty.

6. For the remaining periods of the planning horizon $[N-(k + m)]$, one of the passive adaptive schemes can be employed to approximate the dual control solution. Nevertheless, as with active schemes that are applied to the entire planning horizon, the m-measurement active scheme approximation is quite complicated due to nonlinear dependencies of future variances and covariances on current controls. For illustrations of these nonlinearities and various numerical solution procedures, see Curry [1969], Rausser and Freebairn [1974a], Early and Early [1972], Tse, et al. [1973] and Tse and Bar-Shalom [1973].

7. This covariance matrix is specified for the elements of D arranged by rows.

8. This multiplication is defined by $C = A * B \equiv \sum_{i,j} a_{ij} B_{ij}$ where a_{ij} is the ijth element of an m by n matrix A, and B_{ij} is the ijth submatrix of an mp by np matrix B. Note that C has dimension p by q.

9. It is assumed that the given matrices V_k are such that N_k is negative definite so that the first order condition (59) defines a maximum.

10. The positive definiteness and symmetry of matrices Γ_k follows from the covariance update equation (53). This property becomes important in the second step of the maximization process.

11. By proceeding to period k-1 it is a simple matter to demonstrate that Λ_{k-1} has the same form as Λ_k and thus, by induction, the strategy implied by (59) and (65) is optimal for the modified structure.

12. This interpretation follows immediately since (65) satisfies the result from capital theory that the discount rate equals the own rate of return plus price appreciation. That is,

$$\delta I = \Gamma_k[E(I \otimes z_{k+1}) S_{k+1} (I \otimes z_{k+1}] \Gamma_k V_k^{-1} + (V_{k+1} - V_k) V_k^{-1}.$$

The first term is the own rate of return. The rental on information is the marginal reduction in short-run loss that would take place if extra information were made available for one period; this rental is $\Gamma_k[E(I \otimes z_{k+1})S_{k+1}(I \otimes z_{k+1})]\Gamma_k$, the negative derivative with respect to Γ_k of the uncertainty term $E\{z'_{k+1} (S_{k+1} \circledast \Gamma_k) z_{k+1}\}$.

13. For the details of this scheme, see Chow [1975, pp. 254-261].

14. Note that for this approximation, the dimension of the information state remains constant in time.

15. Note that $\Lambda^*_{k+1}(\cdot)$ is obtained by the closed-loop minimization of C(k+1).

16. For a thorough exposition of this approach, see Kendrick and Kang [1976].

17. Projection of the covariance matrices must be done as many times as there are search steps in each time period while updating the covariances is done only once each time period.

18. For further details on this approach see the applications section 7.4.

19. Prescott [ ] has also examined the control approach in the context of a very simple model involving one unknown parameter, a single control and state. His adaptive controls were derived numerically with the aid of dynamic programming. In other words, he provides no new method of solution; his results were computed by enumeration of all possibilities. This approach is of course computationally infeasible for the multivariate case.

20. The experimentation is executed as follows. In $\frac{n}{2}$ markets the control is $u(k) = u^0(k) + \frac{\Delta}{2}$ where $u^0(k)$ is the certainty equivalence control and in another $\frac{n}{2}$ markets, $u(k) = u^0(k) - \frac{\Delta}{2}$.

21. An implementation of a slightly modified version of Little's model is reported by Rich [1971] and an extension of the univariate case where error of measurement is introduced is presented by Fitzroy [1967].

REFERENCES

1. ABLE, A.B., "A Comparison of Three Control Algorithms as Applied to the Monetarist-Fiscalist Debate," Annals of Economic and Social Measurement, Vol 4, No. 2 (Spring 1975), pp. 23[illegible]-252.

2. ALBERT, A., AND SITTLER, R.W., "A Method of Computing Least Squares Estimators that Keep Up with Data," SIAM J. Contr., Ser. A, Vol. 3 (1966), pp. 384-417.

3. ALSPACH, D., "Dual Control Based on Approximate A Posteriori Density Functions," IEEE Trans. Aut. Contr., Vol. AC-17 (October 1972) pp. 689-693.

4. ALSPACH, D.L., AND SORENSON, H.W., "Approximation of Density Function by a Sum of Gaussian for Nonlinear Bayesian Estimation," Proc. Symp. Nonlinear Estimation Theory and Its Application, San Diego, Calif., 1970.

5. ALSTER, J., AND BELANGER, P.R., "A Technique for Dual Adaptive Control," Automatica, Vol. 10 (1974), pp. 627-634.

6. AOKI, M., Optimization of Stochastic Systems, Academic Press, 1967.

7. ASTROM, K.J., Introduction to Stochastic Control Theory, Academic Press, 1970.

8. ASTROM, K.J., "On Self Tuning Regulators," Automatica, Vol. 9 (1973), pp. 185-199.

9. ATHANS, M., "The Discrete Time Linear-Quadratic - Gaussian Stochastic Control Problem," Annals of Economic and Social Measurement, Vol. 1, No. 4 (1972), pp. 449-491.

10. ATHANS, M., AND WILLNER, D., "A Practical Scheme for Adaptive Aircraft Flight Control Systems," Proc. Symposium on Parameter Estimation Techniques and Applications in Aircraft Flight Testing, NASA Flight Research Center (April 1973).

11. ATHANS, M., WISHNER, R.P., AND BERTOLINI, A., "Suboptimal State Estimation for Continuous-Time Nonlinear Systems from Discrete Noisy Measurements," IEEE Trans. Aut. Cont., Vol. AC-13 (October 1968), pp. 504-514.

12. BAR-SHALOM, Y., AND SIVAN, R., "On the Optimal Control of Discrete-Time Linear Systems with Random," IEEE Trans. Automatic Control, Vol. AC-14 (1969) pp.3-8.

13. BAR-SHALOM AND STENGEL

14. BAR-SHALOM, Y., AND TSE, E., "Caution Probing, and the Value of Information in the Control of Uncertain Systems," Annals of Economic and Social Measurement Vol. 5, No. 3 (Summer 1976), pp. 323-338.

15. BAR-SHALOM, Y., AND TSE, E., "Concepts and Methods in Stochastic Control," in C.T. Leondes, (ed), Control and Dynamic Systems: Advances in Theory and Applications, Academic Press, 1975.

16. BAR-SHALOM, Y., AND TSE, E., "Dual Effect, Certainty Equivalence and Separation in Stochastic Control," IEEE Trans. Automatic Control, Vol. Ac-19 (October 1974), pp. 494-500.

17. BAR-SHALOM, Y., TSE, E., AND LARSON, R.E., "Some Recent Advances in the Development of Closed-Loop Stochastic Control and Resource Allocation Algorithms," Proc. IFAC Symposium on Stochastic Control, Budapest, Hungary, 1974.

18. BELLMAN, R., Adaptive Control Process: A Guided Tour, Princeton, N.J.: Princeton Univ. Press, 1961.

19. BUCY, R.S., AND SENNES, K.D., "Realisation of Optimum Discrete-Time Nonlinear Estimators," Proc. Symp. Nonlinear Estimation Theory and Its Application, San Diego, Calif., 1970.

20. CHONG, C.Y., AND CHENG, D.C., "Multistage Pricing Under Uncertain Demand," Annals of Economic and Social Measurement, Vol. 4, No. 2 (Spring 1975), pp. 311-324.

21. CHOW, G.C., Analysis and Control of Dynamic Economic Systems, New York: Wiley, 1975.

22. CHOW, G.C., "An Approach to the Feedback Control of Nonlinear Econometric Systems," Annals of Economic and Social Measurement, Vol. 5, No. 3 (Summer 1976), pp. 297-310.

23. CURRY, R.E., "A New Algorithm for Suboptimal Stochastic Control," IEEE Trans. Auto. Control, Vol. AC-14 (October 1969), pp. 533-536.

24. DENHAM, W., "Choosing the Nominal Path for a Dynamic System with Random Forcing Function to Optimize Statistical Performance," Div. Eng. Appl. Phys. Harvard Univ., Cambridge, Mass., Tech. Report 449, 1964.

25. DESHPANDE, J.G., AND UPADHYAY, T.N. AND LAINIOTIS, D.G. "Adaptive Control of Linear Stochastic Systems," Automatica, Vol. 9 (1972), pp. 107-115.

26. DONALDSON, D.D., AND KISHI, F.M., "Review of Adaptive Control System Theories and Techniques," in C. Leondes (ed.), Modern Control Systems Theory, Vol.II, McGraw-Hill, 1965.

27. DREYFUS, S.E, Dynamic Programming and The Calculus of Variation (Mathematics in Science and Engineering, (Vol. 21). New York: Academic Press 1965.

28. DREZE, J., "Econometrics and Decision Theory," Econometrica, Vol. 40 (1972), pp.1-18.

29. EARLY, R.H., AND EARLY, B.N., "On the Relative Performance of the Optimal Control System with M Measurements," IEEE Trans. an Automatic Control, Vol. AC-17, 1972.

30. FADEN, M.A., AND RAUSSER, G.C., "Econometric Policy Model Construction: The Past Bayesian Approach," Annals of Economic and Social Measurement, Vol. 5, No. 3 (Summer 1976), pp. 349-362.

31. FELDBAUM, A.A., Optimal Control System, Academic Press, 1965.

32. FITZROY, P.T., "An Adaptive Model for Promotional Decision Making," Special Report, Marketing Science Institute, Philadelphia (April, 1967)

33. FREEBAIRN, J.W. AND RAUSSER, G.C., "An Adaptive Control Approach to Agricultural Policy," Australian Journal of Agricultural Economics, Vol. 18, No. 3 (December 1974), pp. 208-220.

34. GORMAN, D., AND ZABORSZKY, J.,"Stochastic Optimal Control of Continuous Time Systems with Unknown Gain," IEEE Trans. Automatic Control, Vol. AC-13 (1968), pp. 630-638.

35. JAZWINSKI, A., Stochastic Processes and Filtering Theory, New York: Academic, 1970.

36. KALMAN, R.E., "A New Approach to Linear Filtering and Prediction Problems,"ASME J. Basic Eng., Vol. 82 (1960), pp. 35-45.

37. KENDRICK, U., AND KANG, BO HYUN,"An Economists' Guide to Wide Sense Dual Control," Working Paper 75-1, Department of Economics, The University of Texas at Austin.

38. KENDRICK, D.A., AND MAJORS, J., "Stochastic Control with Uncertain Macroeconomic Parameters," Automatica, Vol.10, No. 2 (1974),pp. 587-594.

39. KU, R., AND ATHANS, M., "On the Adaptive Control of Linear Systems Using the Open Loop Feedback Optimal Approach," IEEE Trans. Automatic Contr., Vol. AC-18 (1973),pp. 489-493.

40. LANDAU , I.D., "A Survey of Model Reference Adaptive Techniques - Theory and Applications,"Automatica,Vol. 10 (July 1974), pp. 353-380.

41. LARSON, R.E., AND TSE,E., "Modal Trajectory Estimation and Parallel Computers," Proc. 2nd Symp. Nonlinear Estimation Theory, 1971.

42. LITTLE, J.D.C., "A Model of Adaptive Control of Promotional Spending," Operations Research, Vol. 14, No. 6 (1966), p. 1075.

43. LITTLE, J.D.C., "A Multivariate Adaptive Control Model," IEEE Trans. Automatic Control, Forthcoming.

44. MACRAE, E.C., "An Adaptive Learning Rule for Multiperiod Decision Problems," Econometrica, Vol. 5-6 (1975),pp. 893-906.

45. MACRAE, E.C., "Linear Decision with Experimentation," Annals of Economic and Social Measurement, Vol. 1 (1972), pp. 437-447.

46. MEHRA, R.A., "Approaches to Adaptive Filtering" IEEE Trans. Aut. Contr., Vol. AC-17 (October 1972),pp. 693-698.

47. MEIER L.,"Adaptive Control and the Combined Optimization Problem,"Stanford Res. Inst., Menlo Park,Calif.,Contract NAS 12-59, Memo 6. SRI Project 5578, February 10, 1966.

48. MEIER, L., "Combined Optimal Control and Estimation," in Proceedings, 3rd Annual Allerton Conference, Circuit and System Theory, 1965.

49. MURPHY, W.J., "Optimal Stochastic Control of Discrete Linear Systems with Unknown Gain," IEEE Trans. Automatic Control, Vol. AC-13 (1968), pp. 338-344.

50. NORMAN, A.L., "First Order Dual Control," Annals of Economic and Social Measurement, Vol. 5, No. 3 (Summer 1976), pp. 311-322.

51. PEKELMAN, D., AND TSE, E., "Experimentation and Control in Advertisement: An Adaptive Control Approach," Working Paper 76-04-01, Decision Science Department, University of Pennsylvania, Philadelphia, Pa.

52. PINDYCK, ROBERT S., Optimal Planning for Economic Stabilization, North Holland Publishing Co., 1973.

53. PRESCOTT, E.C., "Adaptive Decision Rules for Macroeconomic Planning," Western Economic Journal, Vol. 9 (1971), pp. 369-37.

54. PRESCOTT, E.C., "The Multi-Period Control Problem under Uncertainty," Econometrica, Vol. 40 (1972), pp. 1043-1058.

55. RAIFFA H., AND SCHLAIFER R., Applied Statistical Decision Theory, Boston: Harvard Business School Press, 1961.

56. RAUSSER, G.C., AND FREEBAIRN, J.W., "Approximate Adaptive Control Solutions to U.S. Beef Trade Policy," Annals of Economics and Measurement, Vol. 3 (January 1974), pp. 177-203.

57. RAUSSER, G.C., AND FREEBAIRN, J.W., "Estimation of Policy Preference Functions: An Application to U.S. Beef Import Policy," Review of Economics and Statistics, Vol. 56 (November 1974), pp. 437-449.

58. RAUSSER, G.C., AND HOCHMAN, E., Dynamic Agricultural Systems: Economic Prediction and Control, Elsevier-North Holland (Dynamic Economic Series), 1977.

59. RAUSSER, G.C., AND HOWITT, R., "Stochastic Control of Environmental Externalities," Annals of Economic and Social Measurement, Vol. 4, No. 2 (Spring 1975), pp. 271-292.

60. RICH, K.L., "An Adaptive Optimization Model for Setting Advertising Budgets," Working Paper P-41, Marketing Science Institute, Philadelphia, 1971.

61. SARIDIS, G.N., "On a Class of Performance Adaptive Self-Organizing Control Systems," in K.S. Fu, (ed.), Pattern Recognition and Machine Learning, Plenum Press, 1971.

62. SARIDIS, G.N., AND DAO, T.K., "A Learning Approach to the Parameter-Adaptive Self-Organizing Control Problem," Automatica, Vol. 8 (September 1972), pp. 589-5[illegible].

63. SARIDIS, G.N., AND GILBERT, H.D., "Self-Organizing Approach to the Stochastic Fuel Regulator Problem," IEEE Trans. On Systems Sciences and Cybernetics, (July 1970), pp. 186-191.

64. SARIDIS, G.N., AND LOBBIA, R.N., "Parameter Identification and Control of Linear Discrete-Time Systems," IEEE Trans. Automatic Control, Vol. AC-17 (February 1972), pp. 52-60.

65. SARIDIS, G.N., AND STEIN, G., "An Approach to the Parameter Adaptive Control Problem," TR-EE68-35, Purdue University, School of Electrical Engineering, December, 1968 .

66. SARRIS, A.H., AND ATHANS, M., "Optimal Adaptive Control Methods for Structurally Varying Systems," Working Paper #24, National Bureau of Economic Research, Cambridge, Mass., December 1973 .

67. SIMON, H.H., "Dynamic Programming under Uncertainty with a Quadratic Function," Econometrica, Vol. 24 (1956), pp. 74-81

68. SORENSON, H.W., "Kalman Filtering Techniques," in C.T. Leondes, (ed.), Advances in Control Systems, Vol. 3, Academic Press, 1966.

69. STEIN, G., AND SARIDIS, G.N., "A Parameter-Adaptive Control Technique," Automatica, Vol. 5 (November 1969), pp. 731-739.

70. SWORDER, D., Optimal Adaptive Control Systems, New York: Academic Press, 1966.

71. TAYLOR, J.B., "Methods of Efficient Parameter Estimation in Control Problems," Annals of Economic and Social Measurement, Vol. 5, No. 3 (1976), pp. 339-348.

72. THEIL, H., "A Note on Certainty Equivalence in Dynamic Planning," Econometrica, (April 1957), pp. 346-349.

73. THEIL, H., Economic Forecasts and Policy, (2nd ed.) Amsterdam: North-Holland Publishing Co., 1970.

74. TSE, E., "Sequential Decision and Stochastic Control," Working Paper, 1975, Engineering Economics Dept., Stanford University, Stanford, California.

75. TSE, E., AND ATHANS, M., "Adaptive Stochastic Control for a Class of Linear Systems," IEEE Trans. Automatic Contr., Vol. AC-17 (February 1972), pp. 38-52.

76. TSE, E., AND BAR-SHALOM, Y., "An Actively Adaptive Control for Discrete-Time Systems with Random Parameters," IEEE Trans. Automatic Control, Vol. AC-18 (April 1973), pp. 109-117.

77. TSE, E., AND BAR-SHALOM, Y., "Adaptive Dual Control for Stochastic Nonlinear Systems with Free End-Time," Proc. 1974 IEEE Conference on Decision and Control, Phoenix, Arizona (November 1974).

78. TSE, E., BAR-SHALOM, AND MEIER, L., "Wide-Sense Adaptive Dual Control of Stochastic Nonlinear System," IEEE Trans. Automatic Control, Vol. AC-18 (April 1973), pp. 98-108.

79. TSIEN, H.S., Engineering Cybernetics, New York: McGraw-Hill Book Co., 1954.

80. TSYPKIN, Y.Z., Foundations of the Theory of Learning Systems, Academic Press, 1973.

81. UPADHYAY, T., "Application of Adaptive Control to Economic Stability Policy," International Journal of System Science, Vol. 6, No. 10 (1975).

82. VAN DER STOEP, D.R., "Trajectory Shaping for the Minimization of State Variable Estimation Errors," IEEE Trans. Automatic Contr. (June 1963), pp. 284-286.

83. WIESLANDER, J., AND WITTENMARK, "An Approach to Adaptive Control Using Real Time Identification," Automatica, Vol. 7 (May 1971).

84. WITSENHAUSEN, H.S., "Separation of Estimation and Control for Discrete-Time Systems," Proceedings of the IEEE, Vol. 59, No. 11, Nov. 1971

85. WONHAM, W.M., "On the Separation Theorem of Stochastic Control," SIAM J. Control, Vol. 6, No. 2 (1968), pp. 312-326.

86. YING, C.C., "Learning By Doing An Adaptive Approach to Multiperiod Decisions," Operations Research, Vol. 15, 1967.

87. ZELLNER, A., An Introduction to Bayesian Inference in Econometrics, New York: John Wiley and Sons, Inc., 1971.