Stochastic Optimal Control of Linear Dynamic Systems

JASON L. SPEYER AND DONALD E. GUSTAFSON
Charles Stark Draper Laboratory, Inc., Cambridge, Mass.

1. Introduction

THE most widely used engineering result in stochastic control theory is that of the solution to the Linear-Quadratic-Gaussian (LQG) problem. The problem statement of the LQG problem is to find a nonanticipative control law based on the measurement history which minimizes the expected value of a *quadratic* function of the state and control variables, subject to *linear* time-varying dynamics and measurement function, with additive *gaussian* white noise. The resulting feedback law is a set of gains, calculated from the deterministic equivalent problem,[1] which operate on the minimum variance estimate of the present state, calculated from a Kalman filter.[2,3] This celebrated result, first presented to the control literature in Refs. 4, 5 for the discrete case and in Refs. 6, 7 for the continuous case, has matured into a useful engineering tool as discussed in Ref. 8. This issue hosts an exceptional bibliography.[9]

The class of stochastic control problems considered here will be limited to certain extensions of the formulation of the Linear-Quadratic-Gaussian problem, for which usable solutions have been obtained. In particular, the dynamics are linear, forced by not only an additive noise process but by multiplicative noise, in which either white noise processes or randomly jumping parameters multiply the state and control variables. Non-quadratic performance criteria, such as an absolute value criterion and an exponential cost function of a quadratic form of the state and control variables, are considered. The exponential criterion represents a multiplicative performance criterion, in contrast to the usual time additive criteria. Dynamic programing solutions are obtained in discrete time in order to simplify the theoretical arguments, although the continuous time case is used for some illustrations. Besides obtaining usable results from an engineering viewpoint, this class of problems illustrates such concepts as the Separation Theorem and the Certainty Equivalence Principle. Because of the limited scope of this paper, areas of interest such as optimal measurement selection strategies, and nonclassical information pattern are not considered.

To a large degree, because of the parallel development of minicomputers, the linear feedback law of the LQG problem becomes a practical control synthesis technique for time varying problems and multiple input-output systems, and an alternative to classical frequency design techniques for time-invariant systems. For example, perturbation guidance schemes lead to linearized systems about a prechosen nominal path. If the nominal path is a deterministic optimal path, the neighboring optimum perturbation guidance scheme[10-14] is the result of solving a problem with linearized dynamics and a quadratic performance index formed naturally by expanding the performance criterion to second order. This case is exceptional since, in general, the weightings in the quadratic criterion are chosen by engineering intuition. The preceding problem reduces to the LQG problem when linearized measurements with additive noise are included as well as a noise term added to the linearized dynamics.

The design philosophy for the LQG problem is expounded by Athans in Ref. 15. He suggests that the choice of the noise not only reflects the designer's appraisal of the uncertainty associated with the additive noise but also models the inaccuracy due to linearization and other parameter inaccuracies. However, as shown by Speyer,[16] small errors in the system coefficients multiplying the state and control variables must be taken into account directly; otherwise, they may lead to instabilities no matter what additive noise variance is chosen. For this reason, our survey begins with the extension of the LQG problem to include a class of random coefficients in the linear system, giving a more direct approach to handling parameter variations than "jacking up" the noise variance or arbitrarily readjusting the weightings in the quadratic performance criterion. It is shown that the random coefficients introduce equivalent state and control weightings in a very natural matter.

For the perturbation guidance problem, Denham,[17] Fitzgerald[18] and Vander Stoep[19] noted that improvements in the expected value of the performance could be obtained if the trajectory is shaped to enhance the assumed linear estimation process. This occurs because the covariance of the error in the estimate is affected by the nominal trajectory through the coefficients in its propagation equation (i.e., the coefficients are partial derivatives of the original nonlinear functions, evaluated

Dr. Jason L. Speyer received his B.S. from MIT in the Department of Aeronautics and Astronautics and his Ph.D. from Harvard in 1968 in Applied Mathematics. He is presently on the research staff at the Charles Stark Draper Laboratory, Inc., and serves as a lecturer in the Department of Aeronautics and Astronautics, MIT. During the academic year 1972–1973, he was on leave as Senior Weizmann Fellow at the Weizmann Institute of Science, Israel. His research interests are in optimization theory, numerical optimization methods, stochastic optimal control, and estimation theory. Dr. Speyer is a member of Sigma Xi, Institute of Electrical and Electronic Engineers, and the AIAA.

Donald E. Gustafson received his Ph.D. in Instrumentation from MIT in September 1972. His principal interests are the theory and applications of estimation and stochastic control to a variety of problems, including guidance and navigation, communications, and biomedical systems. He has worked on guidance and control for Polaris, the LM landing guidance and navigation system, and automatic guidance and navigation systems for station-keeping. Recently, he has been investigating the application of modern filtering theory to acquisition and synchronization in communication systems. Presently, he is Chief of the Biomedical Analysis Section, investigating the application of signal analysis techniques to automatic classification of electrocardiograms. Dr. Gustafson is a member of Sigma Xi, Institute of Electrical and Electronics Engineers, and an Associate Fellow of AIAA.

Received March 5, 1973; revision received December 7, 1973. This work was done while the first author was at the Weizmann Institute, Rehovot, Israel, September 1972–September 1973.
Index categories: Navigation, Control and Guidance Theory; Spacecraft Navigation, Guidance, and Flight-Path Control Systems.

along the nominal trajectory). This dual role of the control variable, in which the estimation process and thereby the performance criterion are improved, is characteristic of the general stochastic control problem and was conceptually introduced by Fel'dbaum[20] under the term "dual control." In this Survey, we avoid this difficulty by restricting ourselves to systems linear in the state variables. The result of this restriction is that the probability density function for the state variables, conditioned on the measurement sequence and the past control sequence, is independent of the controller design. This is essentially the definition of separation of estimation and control due to Witsenhausen,[21] in which the dual control concept no longer applies. Although this assumption greatly simplifies the problem, the resulting formulation is still not usually tractable. For a general survey on stochastic optimal control, see Fleming[22] or Wonham.[23] The object here is to survey results obtained from problems retaining some of the aspects of the LQG problem, but relaxing others which give further insight into the optimal stochastic control problem and, under *certain restrictions*, give usable results for engineering purposes. In Sec. 2, the LQG problem is extended to include multiplicative state and control noise, and the effect of nongaussian noise, random jump parameters, and stochastic differential games are discussed. In Secs. 3 and 4, the performance criterion is generalized. First, the expected value of an absolute value function is shown to give radically different control strategies than obtainable from the LQG problem. Then, the expected value of an exponential function of the quadratic form used in the LQG problem is considered. This performance criterion allows not only the first moments of the quadratic form to be minimized (as in the LQG problem) but some linear combination of all the moments. The LQG problem is seen to be a special case of this problem.

These extensions are *all* characterized by controllers which do *not* obey the "Certainty Equivalence Principle"[14] but for which there is a separation between estimation and control.[21] The Certainty Equivalence Principle states that the optimal control design calculated from the noise-free problem is the same as that for the stochastic problem. This principle holds for the LQG problem. Although separation holds for the extensions considered here, the resulting solutions may be so complex that only approximate solutions are mechanizable. The approximate solutions given here illustrate certain fundamental ideas and there is no attempt to survey ad hoc approximations which lead to suboptimal or adaptive control.

The theoretical results are obtained by using dynamic programing.[24,25] In certain cases, this formulation allows useful simplification in that the solutions can be mechanized. In other instances, approximations will be made so that the problem can be formulated by a finite number of statistics. Within this new framework, the problem is solved by deterministic optimization methods.

In this survey, the authors are undoubtedly biased by their opinions, which are shaped to some extent by their associates and personal prejudices. They beg the reader's forgiveness in omitting important and useful work of others.

2. Stochastic Optimal Control with Multiplicative State and Control Noise

The structure of the linear dynamic system of the LQG problem is here extended to include white multiplicative noise on the state and control variables (these will be denoted state- and control-dependent noise, respectively). This class of problems is particularly interesting because the system parameters are not assumed known. Their randomness is modeled, if only in an approximate way, by a white noise process. The critical assumption is that the multiplicative noise bandwidths are much larger than the frequencies of the signals of interest. As shown in Ref. 16, small errors in some of the parameters of the LQG problem may lead to instabilities. Furthermore, noise of this class does occur naturally (see McLane[26] for some aerospace examples and Gustafson and Speyer[27] for an example in communication theory).

2.1 Problem Formulation

In discrete time, the dynamic programing approach of Aoki[24] and Tou[25] is now considered. The problem is to find a control sequence $\{u_i; i = 1, \ldots, N\}$ which minimizes the quadratic criterion*

$$J = E\left[\sum_{i=1}^{N} (x_{i+1}{}^T Q_{i+1} x_{i+1} + u_i{}^T R_i u_i)\right] \tag{1}$$

where $E[\cdot]$ denotes the expectation operation, x_i is an n-dimensional vector of state variables at stage i, u_i is a p-dimensional control vector at stage i, Q_i is a known $n \times n$ non-negative definite matrix, R_i is a known $p \times p$ non-negative definite matrix, and superscript T indicates matrix or vector transpose. The performance criterion of Eq. (1) is to be minimized, subject to the discrete dynamical system equations

$$x_{i+1} = A_i x_i + B_i u_i + \Gamma_i w_i \tag{2}$$

where A_i is an $n \times n$ transition matrix which is represented as a white noise sequence with known mean $\bar{A}_i$ and variance, B_i is an $n \times p$ white noise control matrix sequence with known mean $\bar{B}_i$ and variance, w_i is an m-dimensional zero mean white noise sequence with known variance, and the initial condition x_1 is a random variable with known mean and variance; and the noisy linear measurement vector

$$z_i = H_i x_i + v_i \tag{3}$$

where z_i is a q-dimensional vector of measurements available to the controller at each stage, H_i is a $q \times n$ output matrix white noise sequence with known mean $\bar{H}_i$ and variance, and v_i is a q-dimensional zero mean white noise sequence with known variance. The random variables A_i, B_i, H_i, w_i, v_i for $i \in [1, N]$‡ are all mutually independent. The restriction that w_i and v_i be white can be relaxed by augmenting additional state variables.[14] The restriction to mutually independent processes can also be relaxed, at the expense of more complex algebraic manipulations. In the following, the general problem is first solved and then specialized to the LQG problem.

2.2 Dynamic Programing Solution

By following the dynamic programing formulation of Aoki[24] and Tou,[23] a feedback control law is obtained. A *nonanticipative* feedback law is to be found which is a function of the past measurement history defined as the composite vector

$$Z_i{}^T = [z_1{}^T, z_2{}^T, \ldots, z_i{}^T] \tag{4}$$

which is an $i \cdot q$-dimensional vector. The control is to be a function† of Z_i as

$$u_i = u_i(Z_i) \tag{5}$$

Consider writing J as a sequence of nested expectations as

$$J = \underset{\{u_i; i\in[1,N]\}}{\text{Min}} E\big[E\{x_2{}^T Q_2 x_2 + u_1{}^T R_1 u_1 + E[x_3{}^T Q_3 x_3 + u_2{}^T R_2 u_2 + \cdots + E(x_{N+1}{}^T Q_{N+1} x_{N+1} + u_N{}^T R_N u_N / Z_N) \ldots / Z_2] / Z_1\}\big] \tag{6}$$

where $E[\,.\,/\,.\,]$ denotes conditional expectation and where all the information in Z_i is contained in Z_{i+1}. Applying the fundamental lemma of Ref. 29, the operations of expectation and minimization are interchanged in Eq. (6) to give

$$J = E\big[\underset{u_1}{\text{Min}}\, E\{x_2{}^T Q_2 x_2 + u_1{}^T R_1 u_1 + \underset{u_2}{\text{Min}}\, E[x_3{}^T Q_3 x_3 + + u_2{}^T R_2 u_2 + \cdots + \underset{u_N}{\text{Min}}\, E(x_{N+1}{}^T Q_{N+1} x_{N+1} + u_N{}^T R_N u_N / Z_N) \ldots / Z_2] / Z_1\}\big] \tag{7}$$

---

* Small letters are used to indicate scalars and vectors; capital letters are used for matrices.

† These admissible controls are required to be measurable functions[28] of the past measurement history, to assure that they are random variables.

‡ The notation $i \in [1, N]$ means that i belongs to the set of integers from 1 to N.

From an inspection of Eq. (7) it is convenient to define the optimal return function $J(Z_i)$ as the cost to go from stage i to $N+1$, given the measurement sequence Z_i and the optimal control from stage i to stage N, i.e.,

$$J(Z_i) = \operatorname*{Min}_{u_i} E[x_{i+1}{}^T Q_{i+1} x_{i+1} + u_i{}^T R_i u_i + \cdots + \operatorname*{Min}_{u_N} E(x_{N+1}{}^T Q_{N+1} x_{N+1} + u_N{}^T R_N u_N / Z_N) \ldots / Z_i] \quad (8)$$

A *recursion relation* can now be written between the optimal return functions at Z_{i+1} and Z_i as

$$J(Z_i) = \operatorname*{Min}_{u_i} E\{[x_{i+1}{}^T Q_{i+1} x_{i+1} + u_i{}^T R_i u_i + J(Z_{i+1})]/Z_i\} \quad (9)$$

By repeated use of Eq. (9), from the final stage to the initial stage, the cost in Eq. (7) is evaluated.

2.3 Evaluation of Optimal Return Function at Stage N

The optimal return function at the last stage of control, given Z_N, is

$$J(Z_N) = \operatorname*{Min}_{u_N} \int (x_{N+1}{}^T Q_{N+1} x_{N+1} + u_N{}^T R_N u_N) p(x_{N+1}/Z_N) dx_{N+1} \quad (10)$$

where $p(x_{N+1}/Z_N)$ is the conditional probability density function of x_{N+1}, given Z_N. Although this density function is, in general, quite difficult to obtain,[30] it is assumed that it is available. The state x_{N+1} can be eliminated in terms of x_N, u_N and the white noise processes at stage N by using the dynamic constraint Eq. (2) and the Chapman-Kolmogorov equation[30] where

$$p(x_{N+1}/Z_N) = \int p(x_{N+1}/x_N, Z_N, A_N, B_N, w_N) \times p(x_N, A_N B_N, w_N/Z_N) dx_N\, dA_N\, dB_N\, dw_N \quad (11)$$

Because of the assumed independence of (A_N, B_N) and w_N

$$p(x_N, A_N, B_N, w_N/Z_N) = p(A_N, B_N) p(w_N) p(x_N/Z_N) \quad (12)$$

Using Eqs. (2, 11, and 12), Eq. (10) becomes

$$J(Z_N) = \operatorname*{Min}_{u_N} \int [(A_N x_N + B_N u_N)^T Q_{N+1} (A_N x_N + B_N u_N) + u_N{}^T R_N u_N] p(A_N, B_N) p(x_N/Z_N) dA_N\, dB_N, dx_N + E[w_N{}^T Q_{N+1} w_N] \quad (13)$$

Rewriting Eq. (13) and taking the variance of w_i as W_i

$$J(Z_N) = \operatorname*{Min}_{u_N} \int [x_N{}^T E[A_N{}^T Q_{N+1} A_N] x_N + 2u_N{}^T E[B_N{}^T Q_{N+1} A_N] x_N + u_N{}^T \{R_N + E[B_N{}^T Q_N B_N]\} u_N] \times p(x_N/Z_N) dx_N + tr(Q_{N+1} W_N) \quad (14)$$

where tr denotes the trace operation.§

Since u is a deterministic function of Z_N,¶ then the simple parameter minimization indicated in Eq. (14) results in an optimal control law as

$$u_N{}^* = -\Lambda_N \hat{x}_N \quad (15)$$

where

$$\Lambda_N = \{R_N + E[B_N{}^T Q_{N+1} B_N]\}^{-1} E[B_N{}^T Q_{N+1} A_N] \quad (16)$$

and

$$\hat{x}_N = \int x_N\, p(x_N/Z_N) dx_N \quad (17)$$

where $\hat{x}_N$, the conditional mean of x_N, is, in general, quite difficult to evaluate.[30] If the matrix to be inverted in Eq. (16) is singular, the pseudo-inverse should be used.[31]

If Eqs. (15) and (16) are introduced into Eq. (14) and the term $E[x_N{}^T \pi_N x_N/Z_N]$ is added and subtracted, then Eq. (14), after some manipulation, becomes

$$J(Z_N) = E[x_N{}^T L_N x_N + v_N)/Z_N] \quad (18)$$

where

$$\pi_N = E[A_N{}^T Q_{N+1} B_N](R_N + E[B_N{}^T Q_{N+1} B_N])^{-1} E[B_N{}^T Q_{N+1} A_N]$$
$$L_N = E[A_N{}^T Q_{N+1} A_N] - \pi_N \quad (19)$$
$$v_N = tr(Q_{N+1} W_N) + E[(x_N - \hat{x}_N)^T \pi_N (x_N - \hat{x}_N)/Z_N] \quad (20)$$

§ The trace has the cyclic property; $tr[ABC] = tr[CAB] = tr[BCA]$. Furthermore, $\partial tr[ABC]/\partial A = (BC)^T$.

¶ The derivation in Aoki[24] demonstrates this.

The error in the estimate, $x_N - \hat{x}_N$, can be shown to be independent of the past measurements or functions of these measurements.[32] Therefore, the estimation errors are independent of the past controls [c.f. Eq. (5)]. Only the first term in the expectation of Eq. (18) is influenced by previous controls.

2.4 Optimal Feedback Control Law at Stage i

To compute the control at stage $N-1$, introduce Eq. (18) into the recursion formula (9)

$$J(Z_{N-1}) = \operatorname{Min} E[\{x_N{}^T (Q_N + L_N) x_N + u_{N-1}{}^T R_{N-1} u_{N-1}\}/Z_{N-1}] + E[v_N/Z_{N-1}] \quad (21)$$

The last term does not contain the control. The first term which is a function of u_{N-1} is now in a form identical to Eq. (10). The procedure for determining u_{N-1} is now identical with that given for u_N in Sec. 2.3. By induction, the optimal feedback control at stage i is

$$u_i{}^* = -\Lambda_i \hat{x}_i \quad (22)$$

where

$$\Lambda_i = \{R_i + E[B_i{}^T T_{i+1} B_i]\}^{-1} E[B_i{}^T T_{i+1} A_i]; \; T_i \equiv Q_i + L_i \quad (23)$$

and the conditional mean of x_i is $\hat{x}_i = [x_i/Z_i]$. Similar to Eqs. (19) and (20)

$$\pi_i = E[A_i{}^T T_{i+1} B_i]\{R_i + E[B_i{}^T T_{i+1} B_i]\}^{-1} E[B_i{}^T T_{i+1} A_i]$$
$$T_i = E[A_i{}^T T_{i+1} A_i] + Q_i - \pi_i \quad (24)$$
$$v_i = v_{i+1} + tr[T_{i+1} W_i] + E[(x_i - \hat{x}_i)^T \pi_i (x_i - \hat{x}_i)/Z_i] \quad (25)$$

2.5 The Certainty Equivalence Principle and the Separation of Estimation and Control

Here, as in the LQG problem, the feedback control is a set of gains operating on the present estimate of the state. These gains may be precomputed in computationally limited applications. However, these gains are functions of the variance of the multiplicative noise and obviously not obtainable from the deterministic equivalent problem found by setting the noise to zero. The Certainty Equivalence Principle[14] does not hold. Intuitively, this result is appealing in that the controller now "hedges" or acts differently depending upon the amount and type of uncertainty.

The separation of estimation and control sometimes referred to as the "separation theorem" finds its most general definition in the work of Witsenhausen.[21] Suppose there exists a conditional probability density function of the state x_i, given the variables Z_i and $U_{i-1} = \{u_1, \ldots, u_{i-1}\}$, as $p(x_i/Z_i, U_{i-1})$. The separation result states that there is no loss in obtaining the minimum cost when the set of designs is restricted to a function of $[p(x_i/Z_i, U_{i-1})]$. That is, there exists a function $\mathscr{L}_i$ such that a minimizing control** satisfies $u_i{}^* = \mathscr{L}_i[p(x_i/Z_i, U_{i-1})]$. The filtering problem is entirely independent of the criterion and of the control laws. In the cases surveyed here, this is equivalent to the Separation Theorem defined by Wonham.[7] Separation exists when the optimal control law exists in a subclass of controls which depends only on the expected value of the current state, given the past measurement sequence (the state estimate), i.e., $u_i{}^* = \mathscr{L}_i(\hat{x}_i)$.

The separation theorem is clearly satisfied for the controller Eq. (22). The feedback gains can be calculated a priori and independent of the filter calculations. Note that only the moments of (A_i, B_i) are needed. This stochastic control system is a controller in cascade with the estimator.

Although this solution is elegant, it is unfortunately not usable, since the conditional mean $\hat{x}_i$ (or minimum variance estimate[30]) is not easily attainable because of the presence of state- and control-dependent noise in the dynamics and observations. A nonlinear filter is required, which is usually never realizable in practice. If the state variables are known perfectly, then the controller is easily mechanized, since $\hat{x}_i$ is simply x_i in Eq. (24). Note that the convexity on the control $R_i + E[B_i{}^T T_{i+1} B_i]$ is increased over the no-noise case where $B_i = \bar{B}_i$. This is particularly significant in the continuous case[26,33–36] where this

** The minimizing control may not be unique.

convexity is retained because of the characteristics of the white noise. Thus, the restriction that R in the continuous case be positive definite can be relaxed. This can be of importance in practical problems without natural control penalties. In such cases, LQG theory requires the artificial introduction of control penalties to bound the control. The use of control-dependent noise to account for control influence uncertainties produces convexities which are more realistic.

If there is no multiplicative noise, then the gain, Eq. (23), and the accompanying recursion relationships reduce to those of the LQG problem. The expectations involving (A_i, B_i) are replaced by their averages $(\bar{A}_i, \bar{B}_i)$, whereas additive process noise has *no* effect on the gain program. However, unless the measurement noise, process noise, and initial condition uncertainty are gaussian, the conditional mean is still generated by a nonlinear filter.[37] Curry[38] showed that the result of Eq. (23) still holds even if the measurements are nonlinear.

2.6 Solution to the LQG Problem

In the gaussian noise case with (A_i, B_i) deterministic, the conditional probability density function of the state variables is also gaussian and, therefore, depends only on the conditional mean and the covariance of the estimation errors. This result, due first to Kalman,[2] is given as

$$\hat{x}_i = \bar{x}_i + K_i s_i \tag{26}$$

where $\bar{x}_i$ and s_i, the measurement residual, are

$$\bar{x}_i = E[x_i/Z_{i-1}] = A_{i-1}\hat{x}_{i-1} + B_{i-1}u_{i-1}, \qquad s_i = z_i - H_i\bar{x}_i \tag{27}$$

Given that the white gaussian zero mean noise sequences v_i and w_i and the zero mean normally distributed initial state, x_i, have variances

$$E[v_i v_j^T] = V_i\delta_{ij}, \quad E[w_i w_j^T] = W_i\delta_{ij}, \quad E[x_1 x_1^T] = X_1 \tag{28}$$

where δ_{ij} is the Kronecker function, the weighting K_i, the Kalman filter gain, is

$$K_i = P_i H_i^T V_i^{-1} \tag{29}$$

where the aposteriori error variance P_i is

$$P_i \equiv E[(x_i - \hat{x}_i)(x_i - \hat{x}_i)^T] = M_i - M_i H_i^T (H_i M_i H_i^T + V_i)^{-1} H_i M_i \tag{30}$$

obtained using Eqs. (26) and (27), and where M_i is the extrapolated covariance of the error in the estimate obtained using Eq. (2)

$$M_i = E[(x_i - \bar{x}_i)(x_i - \bar{x}_i)^T] = A_{i-1}P_{i-1}A_{i-1}^T + \Gamma_{i-1}W_{i-1}\Gamma_{i-1}^T; \; M_1 = X_1 \tag{31}$$

The error in the estimate, e_i, is independent of the measurement history Z_i and also independent of $\hat{x}_i$. The measurement residual, s_i, is a white noise sequence[32] independent of the measurements Z_{i-1}. Since only $\hat{x}_i$ contains Z_i and not the error variance, $\hat{x}_i$ is called a sufficient statistic[39] since it summarizes completely the measurement sequence Z_i. Note that this filter in cascade with the controller, Eq. (22), forms the solution to the LQG problem when there is no multiplicative noise and only gaussian additive noise.

For the discrete singular problem ($R_i = 0, i \in [1, N]$), it is still possible, in principle, to determine the optimum gains from Eq. (23). However, it follows from Eqs. (23) and (24) that L_i is singular since $(A_i^{-1}B_i)^T L_i (A_i^{-1}B_i) = 0$ and A_i^{-1} is of full rank by definition of a transition matrix. This can lead to numerical problems of ill-conditioning and indefiniteness. An alternate approach has been used by Ho,[40] in which the optimal control is generated using a dynamic system of order $n-p$ and maintains the state on a singular arc.†† An initial impulse is required in the continuous case to meet initial and terminal boundary conditions. An especially interesting result occurs if $n = p$ and rank $(B_i) = n \forall i$; the optimal control is

$$u_i^* = -B_i^{-1}A_i K_i z_i \tag{32}$$

†† In Sec. 3.1 singular arcs are defined for the continuous case. However, in the discrete case, dynamic programing eliminates many of the technical difficulties.

and $\hat{x}_i = 0 \forall i$ except possibly at the boundaries. The control is a memoryless function of the present measurement and filter gain matrix.

2.7 Approximate Solution by the Calculus of Variation

An approximate solution to the stochastic control problem with multiplicative noise in the dynamics and measurements is obtained by fixing the structure of the controller and filter to be linear.[36] Since the performance index is quadratic and the controller and filter are linear, the problem is reformulated in terms of the covariances of the state and the estimation error of the state. In this statistical space, the problem is now deterministic and deterministic optimization techniques are applicable.[14] In particular, since the dynamic constraints are now matrix difference equations (or, in continuous time, matrix differential equations) the matrix calculus of variations approach[41,42] supplies the necessary conditions and numerical procedure. A two-point boundary value problem results because the control now affects both the mean and variance of the estimation error. This is an example of dual control (first introduced by Fel'dbaum[20]) where both the filter gains and the control are used to improve the filter performance. This should be contrasted with the dynamic programing solution where the control does not affect the variance of the minimum variance estimator, as it does the variance of the *linear* minimum variance estimator.

In Ref. 43, this scheme was applied to the space shuttle re-entry guidance and navigation problem. Control dependent noise was used to account for aerodynamic coefficient uncertainties. For this perturbation guidance scheme, the linearized equations are time-varying. The difficulty is in choosing the proper time-varying weightings (Q_i, R_i) in the performance criterion. The significant result is that even though R_i is chosen as a constant, the additional convexity caused by control dependent noise is time-varying; automatically, an effective time-varying control convexity is obtained. In the re-entry problem, it is seen that the effective convexity on the control reached peaks in regions of high dynamic pressure. Thus, the control gains decrease when the control effectiveness increases.

2.8 LQG Differential Games

A zero sum deterministic differential two-player game involving a quadratic cost subject to the linear dynamics of both players was solved in Ref. 44 with the resulting strategy that each player uses a controller composed of linear gains operating on the state variables. In Ref. 45, the measurement available to the evader was linear in the state with additive white noise, while the pursuer had perfect state information available. Although, in general, the controllers for the players in the continuous problem are infinite-dimensional, for an important subclass of problems, which includes the intercept problem, a finite controller results for both players. For the pursuer, the control gains are the same as the deterministic controller; the certainty equivalence principle holds. However, the evader's controller contains an additional state which is the pursuer's estimation error. If both players make noisy linear measurements, then the controllers depend upon the past measurement history. The controllers then depend upon a growing memory, which in the continuous case, becomes infinite-dimensional.[46]

2.9 LQG Problem with Randomly Jumping Parameters

An interesting class of LQG problems discussed in Refs. 33 and 47 considers linear stochastic dynamics whose coefficients are functions of Markov step processes. The coefficients are assumed to be a continuous time Markov chain with an exponential transition probability to a finite number of states of dimension q. The state of the dynamic system and the states of the Markov step process form a Markov process. For the LQG problem where this enlarged state space is perfectly known, the solution is a linear gain operating on the dynamic state, although these gains are functions of the state of the Markov step process. In order to determine these gains, a coupled set of q, $n \times n$ matrix Riccati equations must be computed.

3. Nonquadratic Performance Criterion

Consider problems with linear dynamics and measurements with additive noise only, but with nonquadratic performance criterion. As in the multiplicative noise case, for a large class of performance criteria the Separation Theorem holds. The problem is to find the control sequence which minimizes the nonquadratic functional

$$J = E\left\{\sum_{i=1}^{N} [L_i(x_i, u_i)]\right\} \tag{33}$$

subject to the dynamic constraints, Eq. (2), and the observations, Eq. (3), where the nonrandom means $(\bar{A}_i, \bar{B}_i, \bar{H}_i)$ are assumed, w_i and v_i are white gaussian noise processes with zero mean and covariance defined in Eq. (28), and the initial state is normally distributed with zero mean and variance given in Eq. (28).

It has been shown by Striebel[39] and Deyst,[48] for the discrete time case, and Wonham,[7] for the continuous time case, that the separation theorem holds and that the optimal controller has the form

$$u_i^* = \Lambda_i(\hat{x}_i) \tag{34}$$

where $\hat{x}_i$ is the present estimate of the state (the conditional mean) and is a sufficient statistic since it combines all the information in Z_i needed to define the optimal control. This is a specialization of the definition of separation by Witsenhausen.[21] Furthermore, given the linearity and gaussian assumptions, the conditional mean $\hat{x}_i$ is calculated by the Kalman filter given by Eqs. (26–31).‡‡ The estimator is independent of this particular class of performance criteria, Eq. (33). One distinguishing feature of this class of performance criteria is that it is an additive function of the present state. This should be contrasted with the multiplicative performance criteria given in Sec. 4. The difficulty with this class of problems is in obtaining the nonlinear feedback law, Eq. (34) which, in general, is found by dynamic programing.

3.1 Midcourse Guidance Illustration

An interesting example of the preceding class of problems is the planetary midcourse guidance problem, where the expected value of the fuel is minimized subject to a constraint on the covariance of the terminal miss. The one-dimensional problem was first solved approximately by Striebel and Breakwell[49] and in three dimensions by Breakwell and Tung.[50] Work in this area is summarized in greater detail than here by Breakwell.[51]

Since small perturbations in the state about a nominal path are assumed, the equations of motion are linear in the perturbed state. With this assumption, the Separation Theorem holds and the control law will be of the form of Eq. (34). An approximate solution is obtained by restricting the controller, Eq. (34), to be linear in the state estimate. The problem is reduced to finding the best set of gains which makes this *linear* controller optimum.

The perturbation equations are given in continuous time because the solution becomes more transparent. Denote $x(t)$ as the terminal miss that would result if no corrective action were taken after time t. $x(t)$ is a scalar position perpendicular to the nominal direction of travel. The time history of $x(t)$ is obtained by solving

$$dx(t)/dt = b(t)u(t) \tag{35}$$

where the "velocity effectiveness" $b(t)$ is $\partial x(T)/\partial c(t)$, $c(t)$ being the instantaneous velocity perpendicular to the nominal path. This partial derivative is evaluated along a nominal trajectory which requires no corrections.

‡‡ Even if the control is nonlinear in the dynamics, as long as the dynamics are linear in the state, the estimate of the state is still found from a Kalman filter.[7]

The problem is to find a corrective acceleration command policy $u(t)$ so as to achieve a satisfactory mean square terminal miss $E[x^2(T)]$ (T is the nominal time-of-arrival) and minimize the expected value of fuel

$$J = E\left[\int_o^T |u(t)|\, dt\right] \tag{36}$$

subject to the dynamic constraint, Eq. (35), and continuous linear observations with gaussian additive noise

$$z(t) = x(t) + v(t) \tag{37}$$

where $v(t)$ is a white gaussian noise process with zero mean and covariance

$$E[v(t)c(\tau)] = V(t)\,\delta(t-\tau) \tag{38}$$

where $\delta(t-\tau)$ is the dirac delta function. In the rest of this section, it will be assumed all functions are implicit functions of time and the explicit notation will be dropped.

3.2 Solution by Approximation

A dynamic programing solution to this problem is avoided if the controller is assumed linear in the state estimate as

$$u = -\Lambda\hat{x} \tag{39}$$

where the feedback gain, Λ is determined so as to minimize Eq. (36) and may contain delta functions. Define the covariances of the state, the estimate, and the error in the estimate as

$$p(t) = E[\hat{x}^2], \quad s(t) = E[x^2], \quad q(t) = E[(\hat{x}-x)^2] \tag{40}$$

The estimate of the state and the error ($e \equiv \hat{x} - x$) are propagated as

$$\dot{\hat{x}} = b\Lambda\hat{x} + k(z-\hat{x}), \quad \dot{e} = -ke + kv \tag{41}$$

where $k = qV^{-1}$ and the covariances of Eq. (40) are propagated simply as ($s = p + q$)

$$\dot{q} = -q^2V^{-1}, \quad \dot{p} = -2b\Lambda p + q^2V^{-1}, \quad q(0) = s(0), \quad p(0) = 0, \quad p(T) = s(T) - q(T) \tag{42}$$

Note that the error is uncorrelated with the estimate, i.e., $E[e(t)\hat{x}(t)] = 0$. The quantity to be minimized, Eq. (36), can be written in terms of the covariance as

$$J = E\left[\int_o^T |u|\, dt\right] = [2/\pi]^{1/2}\int_o^T \Lambda(t)[p(t)]^{1/2}\, dt \tag{43}$$

The preceding equation holds exactly in the one-dimensional problem and is an upper bound on the expected value in the multidimensional problem, c.f. Ref. 50.

Reformulated so that the statistics p, q, s form the state space, the deterministic problem is to find $\Lambda(t)$ for $t \in [0, T]$ which minimizes Eq. (43) subject to Eq. (42). The problem can be solved by a calculus of variations approach rather than dynamic programing. The result is that there exist optimal arcs for which the control Λ lies in the interior of the admissible control region ($0 \leqq \Lambda \leqq \infty$) but not on its bounds. These are called "singular" arcs because along these arcs the variational Hamiltonian is not an explicit function of Λ.[14] Although this one-dimensional problem can be solved by application of Green's theorem, the general approach used in Ref. 50 is briefly described.

The variational Hamiltonian is defined as

$$H = \lambda\dot{p} + \dot{J} = (2p/\pi)^{1/2}\Lambda + \lambda[-2b\Lambda p + q^2V^{-1}] \tag{44}$$

where λ is a Lagrange multiplier associated with $\dot{p}$ in Eq. (42) and is propagated as

$$\dot{\lambda} = -\partial H/\partial p = \Lambda[2b\lambda - (2\pi p)^{-1/2}] \tag{45}$$

If Λ is on a singular arc, the only way the control can be in the interior of the control region ($0 < \Lambda < \infty$) is if

$$\partial H/\partial\Lambda = (2p/\pi)^{1/2} - 2bp\lambda = 0 \tag{46}$$

Equation (46) does not explicitly contain Λ. This is because the Hamiltonian is linear in Λ. Since it is assumed that the singular arc is of finite duration, the time derivative of Eq. (46) is also zero, i.e.,

$$d[(\partial H)/\partial\Lambda]/dt = 0 \tag{47}$$

This derivative is also *not* an explicit function of Λ and can be used to find λ on the singular arc as a function of p and the problem parameters. If λ is eliminated in Eq. (46), the value of p on the singular arc, p^*, is

$$p^*(t) = -bq^2/[2V(db/dt)] \tag{48}$$

Assuming the singular arc is minimizing over $t \in [0, T]$ (see Ref. 52 for necessary and sufficient conditions), then Eq. (48) forms a switching manifold. Initially, $p(0) = 0$ and climbs to $p(t_1) = p^*(t_1)$. During the interval $t \in [0, t_1]$, $\Lambda(t) = 0$. At t_1, $\Lambda(t_1)$ changes discontinuously to the gain which keeps $p(t)$ on $p^*(t)$. The gain on the singular arc is found from $d^2[\partial H/\partial \Lambda]/dt^2 = 0$. $p(t)$ remains on the singular arc until t_2 where Λ again discontinuously drops to zero. t_2 is chosen so that, with $\Lambda(t) = 0$ over $t \in (t_2, T)$, the terminal condition on $s(T)$ is satisfied.

This result differs considerably from the deterministic solution to the midcourse guidance problem which involves only impulsive corrections. Here, a tradeoff is made between the quality of the information as represented by the error variance q and the magnitude of the correction. For example, if $p(0) > p^*(0)$, then $\Lambda(t)$ would be infinite, forcing $p(0+)$ to $p^*(0+)$. A full correction to $p(0+) = 0$ is not made due to the measurement uncertainty. If q were zero, then an entire correction is made as soon as possible to save fuel. However, if the information is poor initially, then the correction should be made later when the information is better. The stochastic control law is optimally "hedging" the magnitude of the velocity correction against the measurement uncertainty.

From a dynamic programing solution in which the restriction to a linear controller is removed, Deyst[48] and Tung and Striebel[53] obtained improved results over the linear controller but indicated that linear controller is a reasonable approximation and captures the form of the optimal solution, i.e., the existence of switching manifolds.

4. Multiplicative Performance Criterion

The performance criteria in the previous section have been sums over all stages of functions evaluated at a particular stage [i.e., Eq. (33)]. Jacobson[54] considers minimizing the expected value of an exponential of a sum of quadratic functions of the state and control variables. This is a multiplicative cost function, since the exponential of a sum can be written as the product of the exponentials. The performance criterion to be minimized is

$$J = E[\mu \exp\{\tfrac{1}{2}\mu\psi\}] \tag{49}$$

where $\mu = \pm\gamma$; γ is a positive scalar; and ψ is a quadratic function of the form

$$\psi = \sum_{i=1}^{N} (x_i^T Q_i x_i + u_i^T R_i u_i) + x_{N+1}{}^T Q_{N+1} x_{n+1} \tag{50}$$

where $N+1$ is the final stage. If the positive exponential ($\mu = +\gamma$) is expanded in a Taylor series, then

$$J = \gamma\{1 + \gamma E[\psi]/2 + \gamma^2 E[\psi^2]/4\cdot 2! + \gamma^3 E[\psi^3]/8\cdot 3! + \cdots\} \tag{51}$$

The cost is a particular weighted sum of all the moments of ψ with γ as the only free parameter. This allows a degree of shaping of the probability density function of the quadratic function. If all the moments of ψ remain bounded as γ becomes arbitrarily close to zero, the first moment dominates the sequence and the optimization problem specializes to the standard LQG problem. If $\mu = -\gamma$, the cost resembles a likelihood function. Furthermore, the negative exponential criterion might represent, approximately, the probability that the entire state and control history lie within specified bounds so that the resulting feedback solution may be a control design for system to operate within specified bounds or limits.

4.1 Perfect Measurement Case

Jacobson[54] minimized Eq. (49) subject to Eq. (2) with deterministic state and control coefficients and perfect knowledge of the state at each stage. The control sequence is obtained from a backward recursion relationship for the expected cost to complete the problem from each stage, conditioned on the state at that stage. Minimizing this expected cost yields the optimal feedback control as a linear function of the current state as

$$u^* = -[(R_i + B_i^T \tilde{G}_{i+1} B_i)^{-1} B_i^T \tilde{G}_{i+1} A_i] x_i \tag{52}$$

where $\tilde{G}_i$ is propagated by recursion relationships backward from the terminal time $G_{N+1} = Q_{N+1}$

$$\tilde{G}_{i+1} = G_{i+1} + \mu G_{i+1}\Gamma_i(W_i^{-1} - \mu\Gamma_i^T G_{i+1}\Gamma_i)^{-1}\Gamma_i^T G_{i+1} \tag{53}$$

$$G_i = Q_i + A_i^T[\tilde{G}_{i+1} - \tilde{G}_{i+1}B_i(R_i + B_i^T\tilde{G}_{i+1}B_i)^{-1}B_i^T\tilde{G}_{i+1}]A_i \tag{54}$$

The gains in Eq. (52) are explicit functions of the process noise uncertainty which appear in the $\tilde{G}_i$ calculation of Eq. (53). For the negative exponential ($\mu = -\gamma$), as Q_{N+1} goes to infinity the effective weighting $\tilde{G}_{i+1}$ goes to $[\Gamma^T W \Gamma/\gamma]^{-1}$, which remains finite. As γ goes to ∞, the control gains go to infinity, which is the result obtained from the LQG problem.

As μ approaches zero, the gain Λ_k approaches that of the LQG problem. If $\mu = -\gamma$ and W_i tends toward infinity, the gains go to zero. This intuitively means that as the randomness increases little is achieved in decreasing the cost function by using control. On the other hand, if $\mu = +\gamma$ and if W_i tends toward infinity, the gains go to infinity. Furthermore, Jacobson[45] shows that the quadratic cost differential game, constructed by assuming that the noise is controlled by an intelligent player and is penalized in the cost function by a weighting equal to its variance, is equivalent to the $\mu = -\gamma$ exponential cost function if it is a cooperative game and to the $\mu = +\gamma$ exponential cost if it is a noncooperative game.

4.2 Noisy Measurement Case

Speyer, Deyst, and Jacobson[55] have considered the exponential cost function when only noisy observations of the state are available in the form of Eq. (3) where the nonrandom H_i is used and v_i is a gaussian white noise sequence. In the most general formulation the control law is a linear combination of the entire smoothed history of the state. In discrete time for a finite number of stages, this result is easily mechanized although a bit cumbersome. Two important special cases retain a finite constant dimensional controller. First, for the terminal control problem where only the terminal state is costed, the control law reduces to a linear combination of the elements of the present estimate of the state. The certainty equivalence principle does not hold because the feedback gains are a function of the estimation error covariance. In the second case, where the state is costed over the entire path, it is assumed that there is no process noise although there are initial condition uncertainty and noisy measurements. The control law becomes a linear combination of not only the present estimate of the state, but of an additional n-dimensional vector which is path-dependent.

The problem is to minimize Eq. (49) subject to the stochastic system, Eqs. (2) and (3); using the average value of the coefficients, w_i and v_i are zero mean gaussian white noise processes with covariances as defined in Eq. (28), and the initial condition uncertainty is normal with zero mean and variance defined in Eq. (28). To solve this from a dynamic programing approach, define the optimal return function as the cost from stage i to stage $N+1$, given the measurement sequence Z_i and the optimal control policy from stage i to N, i.e.,

$$J(Z_i) = \min_{u_i, \ldots, u_N} \mu E[\exp(\mu\psi/2)/Z_i] \tag{55}$$

The recursion formula after using the fundamental lemma of Ref. 29 becomes simply

$$J(Z_i) = \min_{u_i} \mu E[J(Z_{i+1})/Z_i] \tag{56}$$

This recursion formula is more difficult to use than that given by Eq. (9) since the entire state history, costed in ψ, is affected by the expectation over the measurement z_{i+1}.

Consider the terminal cost problem first ($Q_i = 0$ $i \in [1, N]$). This gives results similar to those of the previous section. The control law is now

$$u_i = -\Lambda_i \hat{x}_i \tag{57}$$

where $\hat{x}$ is the conditional mean generated by the Kalman filter [Eqs. (26–31)] and where Λ_i has the same functional form as the gains in Eq. (52). However, the recursion formulas, Eqs. (53) and (54), have $Q_i = 0$ for $i \in [1, N]$; Γ_i is replaced by K_{i+1}, the Kalman gain, and W_i is replaced by

$$E[s_{i+1} s_{i+1}{}^T] = S_{i+1} = H_{i+1} M_{i+1} H_{i+1} + V_{i+1} \tag{58}$$

where the residue s_{i+1} is defined in Eq. (27). The terminal condition on G_{N+1} is

$$G_{N+1} = Q_{N+1} + \mu Q_{N+1}(P_{N+1}{}^{-1} - \mu Q_{N+1})^{-1} Q_{N+1} \tag{59}$$

For the negative exponential, if $Q_{N+1} \to \infty$, then $G_{N+1} \to P_{N+1}{}^{-1}/\gamma$, and the effective weighting for the control gains using Eq. (53) is

$$\tilde{G}_{N+1} \to [P_{i+1} + K_{N+1}{}^T S_{N+1} K_{N+1}]^{-1}/\gamma$$

The general exponential cost problem ($Q_i \neq 0$ for $i \in [1, N+1]$ can be cast into a terminal cost problem by defining a state as the composite vector of the state history up to the current stage. That is,

$$\mathbf{y}_i{}^T = [x_1{}^T, x_2{}^T, \ldots, x_i{}^T] \tag{60}$$

where $\mathbf{y}_i$ is an $(n \cdot i)$-dimensional vector. Note that at each step the vector $\mathbf{y}_i$ increases in dimension by n. Define the $[(N+1) \cdot (N+1) \times (N+1) \cdot (N+1)]$ composite cost weighting matrix as

$$D_{N+1} = \begin{bmatrix} Q_1 & & & 0 \\ & Q_2 & & \\ & & \ddots & \\ 0 & & & Q_{N+1} \end{bmatrix} \tag{61}$$

So D_{N+1} is block diagonal with the ith block equal to Q_i.

The cost Eq. (55), becomes

$$J = E\left[\mu \exp \frac{\mu}{2}\left(\sum_{i=1}^{N} u_i{}^T R_i u_i + \mathbf{y}_{N+1}{}^T D_{N+1} \mathbf{y}_{N+1}\right)\right] \tag{62}$$

which is in the form of a terminal control problem. The problem is to minimize Eq. (62) subject to the stochastic system, in terms of $\mathbf{y}_i$

$$\mathbf{y}_{i+1} = \tilde{A}_i \mathbf{y}_i + \tilde{B}_i u_i + \tilde{\Gamma}_i w_i \tag{63}$$

$$z_i = \tilde{H}_i \mathbf{y}_i + v_i \tag{64}$$

where

$$\tilde{A}_i = \begin{bmatrix} I_i \\ 0 \;\vdots\; A_i \end{bmatrix}, \quad \tilde{B}_i = \begin{bmatrix} 0 \\ B_i \end{bmatrix}, \quad \tilde{\Gamma}_i = \begin{bmatrix} 0 \\ \Gamma_i \end{bmatrix}, \quad \tilde{H}_i{}^T = \begin{bmatrix} 0 \\ H_i{}^T \end{bmatrix} \tag{65}$$

are $(i+1) \cdot n \times i \cdot n$, $(i+1) \cdot n \times p$, $(i+1) \cdot n \times m$, and $q \times (i+1) \cdot n$ matrices, respectively. At each stage $\mathbf{y}_i$ increases in dimension by n. The controller analogous to the terminal controller, Eq. (52), is

$$u_i{}^* = -\tilde{\Lambda}_i \hat{\mathbf{y}}_i \tag{66}$$

where $\hat{\mathbf{y}}_i$ is the conditional mean of $\mathbf{y}_i$. Since Eqs. (63) and (64) form a linear stochastic system, $p(\mathbf{y}_i/Z_i)$ is gaussian. In this enlarged state space, answers corresponding to the terminal cost problem result. The complication is the growing memory of the controller.

If there is no process noise entering the system, the preceding general problem can be transformed into a terminal cost by projecting all the states to the terminal stage in terms of an undefined control sequence. After a good deal of algebraic manipulation, the performance index becomes an exponential function of two n-dimensional vector η_i propagated as

$$\eta_{i+1} = \eta_i + c_i \Phi_{N+1, i+1} B_i u_i; \qquad \eta_1 = 0 \tag{67}$$

where

$$c_i = \sum_{k=1}^{i} \Phi_{N+1,k}{}^{-1} Q_k \Phi_{N+1,k}{}^{-1}, \quad \Phi_{k,l} = \prod_{j=l}^{k=1} A_j, \quad \Phi_{kk} = I \tag{68}$$

The new state, η_{i+1}, is path-dependent since it depends upon u_i, which is a function of the measurement sequence Z_i. The control is found to be a linear function of the composite vector $\hat{y}_i{}^T = [\hat{x}_i{}^T, \eta_i{}^T]$ as

$$u_i{}^* = -\Lambda_i \hat{y}_i \tag{69}$$

where Λ_i is a $p \times 2n$ matrix. The form of the gain is similar to that of Eq. (52) with recursion relations similar to Eqs. (53, 54, and 59), c.f. Ref. 55. The significance of Eq. (69) is that it is a constant dimensional controller. Both this case and the general problem show that improved performance is attainable if additional states are constructed, besides the original state estimate. This is reminiscent of integral compensation in classical control theory.

5. Conclusions

Surveyed are results obtained by removing some of the restrictions used in the LQG problem. Unfortunately, when a restriction is removed, other restrictions need be imposed in order to obtain implementable controllers. If white multiplicative state and control noise is added to the LQG problem, only the perfect estimation problem leads to tractable results. When the number of stages is large in minimizing the exponential cost functional, only the terminal controller, no measurement noise or no process noise cases lead to tractable controllers. In certain cases, approximations to the general problem lead to implementable and elegant solutions; however, no bounds on the goodness of the approximation are available.

The significant aspect of these extensions, which also makes the stochastic optimal control so difficult in general, is that more information is used in the solution. These extensions produce controllers which "hedge" in the presence of uncertainty. Since these controllers operate differently according to the noise intensity (as measured by the noise variance), the Certainty Equivalence Principle does not hold. This intuitive notion is lacking in the LQG problem. It must be noted that, for the stochastic control problem, the quadratic cost function is represented by a probability density function and that the variance of this cost might be just as significant as its mean. This density function is obtained in a few special cases in Ref. 56.

The notion of the Separation Theorem is important in that the stochastic optimal control problem is greatly simplified. Separation occurs because the control only affects the conditional mean of the state and none of the higher moments. If the control affected the variance, for example, control effort would be needed to improve the estimate. (In general, estimation and control are coupled together. This is called the "dual control" problem.[20]) However, the definition of separation given in Refs. 7 and 21 is applicable to all the cases given here when the state space is enlarged, as need be in the exponential criterion problem. The additional information needed for the controller over just the estimation problem is called by Striebel[39] an "information" statistic. This separation is a one way result, since the estimation process does affect the controller design.

In Ref. 54, Jacobson showed the relationship of a deterministic differential game composed of a quadratic cost and linear dynamics with a stochastic control problem, using an exponential cost function but with perfect knowledge of the state. A similar relationship has not been found for the case of noisy state information with exponential cost given here and a differential game. However, there seem to be similarities between the LQG game with noisy state measurement given by Willman[46] discussed in Sec. 2.8 and the general solution given in Sec. 4.2, since they both depend upon the smoothed history of the state and, in the limit to continuous time, result in an infinite dimensional controller.

References

[1] Kalman, R. E., "Contributions to the Theory of Optimal Control," *Boletín de la Sociedad Mat. Mex.*, Vol. 5, 1960, pp. 102–119.

[2] Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME, Journal of Basic Engineering*, Vol. 82, March 1960, pp. 34–45.

[3] Kalman, R. E. and Bucy, R. S., "New Results in Linear Filtering

and Prediction Theory," *Transactions of the ASME, Journal of Basic Engineering*, Vol. 83, Dec. 1961, pp. 95–107.

[4] Gunckel, T. L. and Franklin, G. F., "A General Solution for Linear Sampled Data Control," *Transactions of the ASME, Journal of Basic Engineering*, Vol. 85D, June 1963, pp. 197–201.

[5] Joseph, P. D. and Tou, J. T., "On Linear Control Theory," *AIEE Transactions (Appl. Ind.)*, Vol. 80, Sept. 1961, pp. 193–196.

[6] Potter, J. E., "A Guidance-Navigation Separation Theorem," Rept. RE-11, 1964, Experimental Astronomy Lab., MIT, Cambridge, Mass.

[7] Wonham, W. M., "On the Separation Theorem of Stochastic Control," *SIAM Journal Cont.*, Vol. 6, No. 2, 1968, pp. 312–326.

[8] Athans, M., "Special Issue on the Linear-Quadratic-Gaussian Estimation and Control Problem," *IEEE Transactions on Automatic Control*, Vol. AC-16, No. 6, Dec. 1971.

[9] Mendel, J. M. and Gieseking, D. L., "Bibliography on the Linear-Quadratic-Gaussian Problem," *IEEE Transactions on Automatic Control*, Vol. AC-16, No. 6, Dec. 1971, pp. 847–869.

[10] Breakwell, J. V., Speyer, J. L., and Bryson, A. E., "Optimization and Control of Nonlinear Systems Using the Second Variation," *Journal of the Society of Industrial and Applied Mathematics*, Vol. 1, No. 2, 1963, pp. 193–223.

[11] Kelley, H. J., "An Optimal Guidance Approximation Theory," *IEEE Transactions on Automatic Control*, Oct. 1964, pp. 375–380.

[12] Speyer, J. L. and Bryson, A. E., "A Neighboring Optimum Feedback Control Scheme Based on Estimated Time-to-Go with Applications to Re-Entry Flight Paths," *AIAA Journal*, Vol. 6, No. 5, May 1968, pp. 769–776.

[13] Powers, W. F., "A Method for Computing Trajectories in Optimum Linear Perturbation Guidance Schemes," *AIAA Journal*, Vol. 6, No. 12, Dec. 1968, pp. 2451–2452.

[14] Bryson, A. E. and Ho, Y.-C., *Applied Optimal Control*, Blaisdell, Waltham, Mass., 1969.

[15] Athans, M., "The Role and Use of the Stochastic Linear-Quadratic-Gaussian Problem in Control System Design," *IEEE Transactions on Automatic Control*, Vol. AC-16, No. 6, Dec. 1971, pp. 529–552.

[16] Speyer, J. L., "Instability of Optimal Stochastic Control Systems Under Parameter Variations," *AIAA Journal*, Vol. 9, No. 3, March 1971, pp. 382–396.

[17] Denham, W. F., "Choosing the Nominal Path for a Dynamic System with Random Forcing Function to Optimize Statistical Performance," TR 449, 1964, Div. of Engineering and Applied Physics, Harvard University, Cambridge, Mass.

[18] Fitzgerald, R., "A Gradient Method for Optimizing Stochastic Systems," Ph.D. thesis, May 1964, MIT, Cambridge, Mass.

[19] Vander Stoep, D. R., "Trajectory Shaping for the Minimization of State-Variable Estimation Errors," *IEEE Transactions on Automatic Control*, Vol. AC-13, No. 6, June 1968.

[20] Fel'dbaum, A. A., *Optimal Control Systems*, Academic Press, New York, 1967.

[21] Witsenhausen, H. S., "Separation of Estimation and Control for Discrete Time Systems," *Proceedings of the IEEE*, Vol. 59, No. 11, Nov. 1971, pp. 1557–1566.

[22] Fleming, H. W., "Optimal Continuous Parameter Stochastic Control," *SIAM Review*, Vol. 11, 1969, pp. 470–509.

[23] Wonham, W. M., "Optimal Stochastic Control," *Automatica*, Vol. 5, No. 1, Jan. 1969, pp. 113–118.

[24] Aoki, M., *Optimization of Stochastic Systems—Topics in Discrete-Time Systems*, Academic Press, New York, 1967.

[25] Tou, J. T., *Optimum Design of Digital Control Systems*, Academic Press, New York, 1963.

[26] McLane, P. J., "Optimal Stochastic Control of Linear Systems with State-and-Control-Dependent Disturbances," *IEEE Transactions on Automatic Control*, Vol. AC-16, No. 6, Dec. 1971, pp. 793–798.

[27] Gustafson, D. E. and Speyer, J. L., "Linear Minimum-Variance Filters Applied to Phase-Lock Loops," *Proceedings of the National Telecommunications Conference*, Atlanta, Ga., Nov. 1973.

[28] Wong, E., *Stochastic Processes in Information and Dynamical Systems*, McGraw-Hill, New York, 1971.

[29] Meier, L., III, Lawson, R. E., and Tether, A. J., "Dynamic Programming for Stochastic Control of Discrete Systems," *IEEE Transactions on Automatic Control*, Vol. AC-16, No. 6, Dec. 1971.

[30] Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[31] Kalman, R. E. and Englar, T. S., "A User's Manual for the Automatic Synthesis Program," CR-475, June 1966, NASA.

[32] Frost, P. A. and Kailath, T., "An Innovations Approach to Least-Squares Estimation—Part III: Nonlinear Estimation in White Gaussian Noise," *IEEE Transactions on Automatic Control*, Vol. AC-16, No. 3, June 1971.

[33] Wonham, W. M., "Random Differential Equations in Control Theory," in *Probabilistic Methods in Applied Mathematics*, edited by A. T. Bharucha-Reid, Academic Press, New York, 1969, Chap. 2.

[34] McLane, P. J., "The Optimal Regulator Problem for a Stationary Linear System with State-Dependent Noise," *Transactions of the ASME, Journal of Basic Engineering*, Vol. 92, June 1970, pp. 363–367.

[35] Kleinman, D. L., "Optimal Stationary Control of Linear Systems with Control-Dependent Noise," *IEEE Transactions on Automatic Control*, Vol. AC-14, Dec. 1969, pp. 673–677.

[36] Gustafson, D. E., "On Optimal Estimation and Control of Linear Systems with State-Dependent and Control Dependent Noise," Doctoral Dissertation, June 1972, MIT, Cambridge, Mass.

[37] Alspach, D. L. and Sorenson, H. W., "Stochastic Optimal Control for Linear but Non-Gaussian Systems," *International Journal of Control*, Vol. 13, No. 6, 1971, pp. 1169–1181.

[38] Curry, R., "A Separation Theorem for Nonlinear Measurements," *IEEE Transactions on Automatic Control*, Vol. AC-14, No. 5, Oct. 1969, pp. 561–564.

[39] Striebel, C. T., "Sufficient Statistics in the Optimal Control of Stochastic Systems," *Journal of Mathematical Analysis and Applications*, Vol. 12, 1965, pp. 576–592.

[40] Ho, Y.-C., "Linear Stochastic Singular Control Problems," *Journal of Optimization Theory and Applications*, Vol. 9, No. 1, Jan. 1972, pp. 24–31.

[41] Denham, W. F. and Speyer, J. L., "Optimal Measurement and Velocity Corrections Programs for Midcourse Guidance," *AIAA Journal*, Vol. 2, No. 5, 1964, pp. 896–907.

[42] Athans, M., "The Matrix Minimum Principle," *Information and Control*, Vol. II, 1968, pp. 592–606.

[43] Gustafson, D. E. and Speyer, J. L., "On the Design of Linear Regulators for Nonlinear Stochastic Systems," AIAA Paper 74-127, Washington, D.C., 1974.

[44] Ho, Y.-C., Bryson, A. E., Jr., and Baron, S., "Differential Games and Optimal Pursuit—Evasion Strategies," *IEEE Transactions on Automatic Control*, Vol. AC-10, Oct. 1965, pp. 385–389.

[45] Behn, R. and Ho, Y.-C., "On a Class of Linear Stochastic Differential Games," *IEEE Transactions on Automatic Control*, Vol. AC-13, No. 3, June 1968, pp. 227–240.

[46] Willman, W., "Formal Solutions for a Class of Stochastic Pursuit-Evasion Games," *IEEE Transactions on Automatic Control*, Vol. AC-14, No. 5, Oct. 1969, pp. 504–509.

[47] Sworder, D., "Feedback Control of a Class of Linear Systems with Jump Parameters," *IEEE Transactions on Automatic Control*, Vol. AC-14, No. 1, Feb. 1969, pp. 9–15.

[48] Deyst, J. J., "Optimal Control in the Presence of Measurement Uncertainties," Sc.D. thesis, 1967, MIT, Cambridge, Mass.

[49] Breakwell, J. V. and Striebel, C. T., "Minimum Effort Control in Interplanetary Guidance," IAS Paper 63-80, New York, Jan. 1963.

[50] Breakwell, J. V. and Tung, F., "Minimum Effort Control of Several Terminal Components," *SIAM Journal on Control*, Ser. A, Vol. 3, No. 1, 1965, pp. 295–316.

[51] Breakwell, J. V., "Stochastic Optimization Problems in Space Guidance," in *Stochastic Optimization and Control*, edited by H. F. Karreman, Wiley, New York, 1968.

[52] Speyer, J. L. and Jacobson, D. H., "Necessary and Sufficient Conditions for Optimality for Singular Control Problems: A Transformation Approach," *Journal of Mathematical Analysis and Applications*, Vol. 33, 1971, pp. 163–187.

[53] Tung, F. and Striebel, C. T., "A Stochastic Optimal Control Problem and Its Applications," *Journal of Mathematical Analysis and Applications*, Vol. 12, 1965, pp. 350–359.

[54] Jacobson, D. H., "Optimal Stochastic Linear Systems with Exponential Performance Criterion and their Relation to Deterministic Differential Games," *IEEE Transactions on Automatic Control*, Vol. AC-18, No. 2, April 1973, pp. 124–131.

[55] Speyer, J. L., Deyst, J. J., and Jacobson, D. H., "Optimization of Stochastic Linear Systems with Additive Measurement and Process Noise Using Exponential Performance Criteria," *Proceedings of the Eleventh Annual Allerton Conference on Circuit and System Theory*, Oct. 1973; also *IEEE Transactions on Automatic Control*, Vol. AC-19, No. 4, Aug. 1974.

[56] Sain, M. K. and Liberty, S. R., "Performance-Measure Densities for a Class of LQG Control Systems," *IEEE Transactions on Automatic Control*, Vol. AC-16, No. 5, Oct. 1971, pp. 431–439.