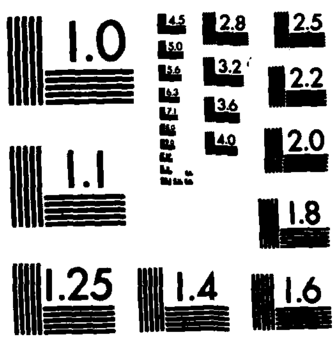
AD-A124 627 STOCHASTIC DIFFERENTIAL GAMES WITH COMPLEXITY CONSTRAINED STRATEGIES(U) CALIFORNIA UNIV LOS ANGELES SCHOOL OF ENGINEERING AND APPLIED SCIENCE D M STUART 1/2

UNCLASSIFIED MAR 82 DASG60-80-C-0007 F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

SE

AD A124627

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. [...] ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) STOCHASTIC DIFFERENTIAL GAMES WITH COMPLEXITY CONSTRAINED STRATEGIES	5. TYPE OF REPORT & PERIOD COVERED Final, 11/29/79-11/28/81
	6. PERFORMING ORG. REPORT NUMBER None
7. AUTHOR(s) Donald Macdonald Stuart, Jr. UCLA, School of Engineering & Applied Science	8. CONTRACT OR GRANT NUMBER(s) DASG-60-80-C-0007
9. PERFORMING ORGANIZATION NAME AND ADDRESS UCLA, School of Engineering and Applied Science Los Angeles, California 90024	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS None
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Technology Center U.S. Army Ballistic Missile Defense Command Huntsville, Alabama	12. REPORT DATE March, 1982
	13. NUMBER OF PAGES 119
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Same as No. 11	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is not limited.

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

None

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See attached

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102 LF 014 6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT

An examination is made of some problems encountered in the optimal control of a linear dynamic system by two independent controllers with noisy state observations, the controllers having either conflicting or concurring objectives. The question of what form the optimal controls should take is also discussed. By restricting consideration to linear forms, it is shown that the computational complexity of a general optimal linear strategy is considerable. Attention is further restricted to a particular linear form for the optimal controls: a matrix transformation of a vector which is the solution of a linear differential equation forced by the observations. Properties of certain forms of this type of control are analyzed, and it is shown that the parameters of these forms may be expressed in terms of solutions to a set of nonlinear differential equations with split boundary conditions. It is also demonstrated that these forms reduce, in a one-input case, to those specified by the separation principle of one-sided optimal control.

STOCHASTIC DIFFERENTIAL GAMES WITH COMPLEXITY CONSTRAINED STRATEGIES

By

Donald Macdonald Stuart, Jr.

March, 1982

Submitted under contract
DASG-60-80-C-0007
Advances In Technology Development
for Exoatmospheric Intercept Systems

U.S. Army Ballistic Missile
Defense Command
Advanced Technology Center
Huntsville, Alabama

School of Engineering and Applied Science
University of California
Los Angeles, California

PREFACE

With the development of advanced Army Ballistic Missile Defense systems there arises the requirement for increasingly sophisticated guidance and control techniques and systems. Fundamental to this area is the competitive situation wherein a target vehicle is attempting to avoid an intercept vehicle or to say the same thing another way an interceptor is attempting to hit an evasively maneuvering target. This competitive situation is referred to in the technical literature as the differential game problem. Results developed for this fundamental problem area obviously have applicability to a wide variety of situations not only military strategic but other competitive situations as well. This report is one of a companion set of reports issued on this broad research effort and it deals with continuous time differential games in a stochastic environment. One of its purposes is to develop techniques which can result in the simplest possible thoroughly effective systems.

ABSTRACT

An examination is made of some problems encountered in the optimal control of a linear dynamic system by two independent controllers with noisy state observations, the controllers having either conflicting or concurring objectives. The question of what form the optimal controls should take is also discussed. By restricting consideration to linear forms, it is shown that the computational complexity of a general optimal linear strategy is considerable. Attention is further restricted to a particular linear form for the optimal controls: a matrix transformation of a vector which is the solution of a linear differential equation forced by the observations. Properties of certain forms of this type of control are analyzed, and it is shown that the parameters of these forms may be expressed in terms of solutions to a set of nonlinear differential equations with split boundary conditions. It is also demonstrated that these forms reduce, in a one-input case, to those specified by the separation principle of one-sided optimal control.

CONTENTS

Chapter 1
INTRODUCTION

1.1 Preliminary Remarks

A wealth of practical problems arise out of natural engineering situations in which the control of the "system" is in the hands of more than a single controller. Such multiple controllers may have varying objectives, and these objectives may be wholly or partially conflicting or concurring. As examples we might cite pursuit and evasion situations with two vehicles, rendezvous in space of two vehicles, and control of an international economic system by several state governments. In many of these natural engineering situations the controllers are required to act with imperfect information as to the true state of the system; thus, in such cases the question of how to control in a manner which is in some sense optimal is usually difficult to answer. For this reason control theorists have often chosen to analyze abstract mathematical models which are thought to retain some important characteristics of their real-world counterparts, since such models yield more readily to analysis than the actual situations.

It is the object of this research to investigate the nature of a specific type of two-input control problem: one in which the controllers have conflicting objectives, the state of the system is described by a system of linear differential equations, the criterion functional is quadratic, and the controllers have available only state observations which are obscured by white Gaussian noise. This is a stochastic differential game situation. It is thought that a thorough analysis of this problem may reveal some interesting facts which will contribute to a greater understanding of more complicated problems.

In the realm of optimal control theory, systems which are described by linear differential equations and quadratic cost functionals have become classic objects of analysis. This is partially because they have the pedagogical advantage of yielding with relative ease solutions which illustrate theoretical principles in a simple framework. It is also because many real-world optimal control problems can be fitted into the linear-quadratic mathematical framework; hence we gain insight into the behavior of practical systems by studying the linear-quadratic models.

In the area of stochastic optimal control similar statements apply. Here the so-called "separation theorem" [9,15] enables us to combine our knowledge of the deterministic optimal control for linear-quadratic systems with the results of Kalman and Bucy [17,18] in the area of estimation and prediction of the state of stochastic dynamic systems to produce a control which is stochastically optimal in the sense that it minimizes the expected value of the cost functional. Specifically, the theory of deterministic optimal control when applied to linear-quadratic systems shows that the optimal control function can be expressed as linear state feedback; i.e., if we denote the control signal by $U(t)$ and the state of the system by $X(t)$, then

$$U(t)_{opt} = K(t)X(t)$$

where $K(t)$ is a feedback gain matrix determined by the parameters of the system. The separation theorem then shows that when the state is not directly observable the stochastically optimal control signal is

$$U(t)_{\text{stoc. opt}} = K(t)\hat{X}(t) .$$

where $\hat{X}(t)$ is the conditional mean of the state, based on all available knowledge and measurements, and $K(t)$ is the deterministically optimal feedback gain. This result is intuitively satisfying in that we simply use the best (mean square) estimate of the state in place of the actual value of the state to obtain the best realizable control function.

Differential games are natural objects for the application of optimal control theory, since in many cases the formulations of these problems are only slight modifications of ordinary optimal control problems with provisions for an extra control input to the plant. Indeed, differential games described by linear differential equations and quadratic payoff functionals yield under mild restrictions solutions which are not greatly different in nature from those of the analogous one-sided optimal control problems. To be specific, the optimal strategies for both players are linear state feedback control functions.

A natural conjecture then is that in the stochastic version of the linear-quadratic differential game, where the players are unable to observe the state directly, the stochastically optimal strategy would be to employ the conditional mean of the state in place of the state in the linear feedback. Unfortunately, this conjecture is false, as is easily shown by simple counterexamples. We shall, therefore, proceed to inquire about the nature of the optimal strategies and to analyze in particular the special cases where the controllers are restricted to the use of computationally feasible (practical) strategies.

To begin the development, we shall formulate the stochastic differential game problem which will be the underlying object of analysis for the remainder of this work.

1.2 The Deterministic Game Formulation

The differential game described by the system equation

$$\dot{X}(t) = F(t)X(t) - G_1(t)U_1(t) + G_2(t)U_2(t) \tag{1.1}$$

$\left(\text{where } X(0) = X_o, \; E\left\{X_o\right\} = \bar{X}_o, \text{ and } \mathrm{Cov}\left\{X_o\right\} = \Psi_{X_o}\right)$.

with payoff functional

$$J(U_1,U_2) = \frac{1}{2} E\left\{X^*(T)Q_3X(T) + \int_o^T U_1^*(t)Q_1(t)U_1(t)dt - \int_o^T U_2^*(t)Q_2(t)U_2^*(t)dt\right\} \tag{1.2}$$

(where $Q_1(t)$, $Q_2(t)$, and Q_3 are positive definite symmetric matrices, and the asterisk denotes vector or matrix transpose) and observation equations

$$
\begin{aligned}
Z_1(t) &= H_1(t)X(t) + \eta_1(t) \\
Z_2(t) &= H_2(t)X(t) + \eta_2(t)
\end{aligned}
\tag{1.3}
$$

(where η_1 and η_2 are zero mean white Gaussian noise, with

$$
E\{\eta_1(t)\eta_1^*(\tau)\} = R_1(t)\delta(t-\tau),\ E\{\eta_2(t)\eta_2^*(\tau)\} = R_2(t)\delta(t-\tau),
$$

$$
E\{\eta_1(t)\eta_2^*(\tau)\} = 0,
$$

and with $R_1(t)$, $R_2(t)$ continuously differentiable, $0 \leq T$) may be described in more efficient and general terms. First, note that since Q_1, Q_2, and Q_3 are positive definite and symmetric, they may be factored as

$$
Q_i = Q_i^{\frac{1}{2}*} Q_i^{\frac{1}{2}} \qquad i = 1,2,3
$$

where Q_i is triangular and non-singular. Then by the transformations

$$
X \to Q_3^{-\frac{1}{2}} X' \tag{1.4}
$$

$$
U_1 \to Q_1(t)^{-\frac{1}{2}} U_1' \tag{1.5}
$$

$$
U_2 \to Q_2(t)^{-\frac{1}{2}} U_2' \tag{1.6}
$$

the system equation becomes

$$\dot{X}'(t) = Q_3^{\frac{1}{2}}F(t)Q_3^{-\frac{1}{2}}X'(t) - Q_3^{\frac{1}{2}}G_1(t)Q_1^{-1}(t)U_1'(t) + Q_3^{\frac{1}{2}}G_2(t)Q_2^{-\frac{1}{2}}(t)U_2'(t) \tag{1.7}$$

Or, defining new matrices

$$F'(t) = Q_3^{\frac{1}{2}}F(t)Q_3^{-\frac{1}{2}} \tag{1.8}$$

$$G_1'(t) = Q_3^{\frac{1}{2}}G_1(t)Q_1^{-\frac{1}{2}}(t) \tag{1.9}$$

$$G_2(t) = Q_3^{\frac{1}{2}}G_2(t)Q_2^{-\frac{1}{2}}(t) \tag{1.10}$$

the system equation becomes

$$\dot{X}'(t) = F'(t)X'(t) - G_1'(t)U_1'(t) + G_2'(t)U_2'(t) \tag{1.1A}$$

the payoff functional becomes

$$J(U_1',U_2') = \frac{1}{2}E\left\{X'(T)^*X'(T) + \int_0^T U'^*(t)U'(t)dt - \int_0^T U_2'^*(t)U_2'(t)dt\right\} \tag{1.2A}$$

and the observation equations are

$$\begin{aligned} Z_1(t) &= H_1(t)Q_3^{-\frac{1}{2}}X'(t) + \eta_1(t) \overset{d}{=} H_1'(t)X'(t) + \eta_1(t) \\ Z_2(t) &= H_2(t)Q_3^{-\frac{1}{2}}X'(t) + \eta_2(t) \overset{d}{=} H_2'(t)X'(t) + \eta_2(t) \end{aligned} \tag{1.3A}$$

In view of the possibility of making these transformations, we may consider (1.1A), (1.2A), and (1.3A) to be a general problem formulation. Further generalization is possible, however.

Note that the solution to (1.1A) may be written (dropping the "prime" subscripts) as

$$X(t) = \phi(t)X_o - \int_o^t \phi(t)\phi^{-1}(\tau)G_1(\tau)U_1(\tau)d\tau + \int_o^t \phi(t)\phi^{-1}(\tau)G_2(\tau)U_2(\tau)d\tau \tag{1.11}$$

If we now define the integral operators T_1 and T_2 by

$$(T_1U_1)(t) = \int_o^t \phi(t)\phi^{-1}(\tau)G_1(\tau)U_1(\tau)d\tau \tag{1.12}$$

$$(T_2U_2)(t) = \int_o^t \phi(t)\phi^{-1}(\tau)G_2(\tau)U_2(\tau)d\tau \tag{1.13}$$

then equation (1.11) may be written as

$$X(t) = \phi(t)X_o - (T_1U_1)(t) + (T_2U_2)(t) \tag{1.14}$$

or

$$X(T) = \phi(T)X_o - (T_1U_1)(T) + (T_2U_2)(T) \tag{1.15}$$

Note that $\phi(T)X_o$ is the predicted miss distance under the condition of no control being applied by either player.

We shall henceforth drop the argument T whenever $t = T$, so it will be understood that when we write

$$X(T) = \phi X_o - T_1U_1 + T_2U_2 \tag{1.16}$$

this is equivalent to (1.15), and when the argument t is intended, we shall use the form (1.14).

The first term in the payoff functional (1.24) may thus be written as

$$X(T)^*X(T) = \left\langle \phi X_o - T_1U_1 + T_2U_2, \phi X_o - T_1U_1 + T_2U_2 \right\rangle \tag{1.17}$$

(where $\langle \cdot,\cdot \rangle$ here denotes the inner product in Euclidean space.) The other terms in the payoff functional may similarly be expressed as inner products in the Hilbert spaces formed as finite copies of the space $L^2(T)$. Hence, the payoff functional may be written as

$$\begin{aligned} J(U_1,U_2) = \frac{1}{2} E\Big\{&\left\langle \phi X_o - T_1U_1 + T_2U_2, \phi X_o - T_1U_1 + T_2U_2 \right\rangle \\ &+ \left\langle U_1,U_1 \right\rangle - \left\langle U_2,U_2 \right\rangle\Big\} \end{aligned} \tag{1.18}$$

We wish to find the U_1 and U_2 which minimaximize $J(U_1,U_2)$. We require that these be functions of only the observables Z_1 and Z_2 and the known statistics of X_o, η_1, η_2, with the specific functional forms to be determined.

To acquire some insight into the nature of the problem, we first solve the deterministic version;[1] i.e., we assume both players know precisely the initial condition X_o. We also assume both players are able to monitor the state continuously during the progress of the game. We may drop the expected value notation for the time being, and thus express the payoff as

$$J(U_1U_2) = \frac{1}{2}\left\{\left\langle \phi X_o - T_1U_1 + T_2U_2,\ \phi X_o - T_1U_1 + T_2U_2 \right\rangle \right.$$

$$\left. + \left\langle U_1, U_1 \right\rangle - \left\langle U_2, U_2 \right\rangle \right\} \tag{1.19}$$

To minimaximize this quantity with respect to U_1 and U_2, we form the functional derivative of $J(U_1,U_2)$ with respect to U_1 and U_2 and set these derivatives equal to zero. Thus

$$\frac{\partial}{\partial U_1} J(U_1U_2) = U_1 - T_1^*\phi X_o - T_1^*T_2U_2 + T_1^*T_1U_1 = 0 \tag{1.20}$$

$$\frac{\partial}{\partial U_2} J(U_1U_2) = -U_2 + T_2^*\phi X_o - T_2^*T_1U_1 + T_2U_2 = 0 \tag{1.21}$$

(where the asterisk here denotes the adjoint operator.) We see that for these equations to be true U_i must be in the range of T_i^*, and so we may write $U_i = T_i^*\lambda_i$. Substituting these expressions into the original equations (1.20) and 1.21), we have

[1]The derivation given here is due to Porter [22].

$$T_1^*\lambda_1 = T_1^*\phi X_o + T_1^*T_2U_2 - T_1^*T_1U_1 \tag{1.22}$$

$$T_2^*\lambda_2 = T_2^*\phi X_o + T_2^*T_2U_2 - T_2^*T_1U_1 \tag{1.23}$$

These equations will be satisfied if

$$\lambda_1 = \phi X_o + T_2U_2 - T_1U_1 = \lambda_2 \tag{1.24}$$

which implies

$$\lambda_1 = \phi X_o + T_2T_2^*\lambda_1 - T_1T_1^*\lambda_1 \tag{1.25}$$

Thus

$$\left[I + T_1T_1^* - T_2T_2^*\right]\lambda_1 = \phi X_o \tag{1.26}$$

or

$$\lambda_1 = \left[I + T_1T_1^* - T_2T_2^*\right]^{-1}\phi X_o \tag{1.27}$$

when the indicated inverse exists. Thus we may write

$$U_1 = T_1^*\left[I + T_1T_1^* - T_2T_2^*\right]^{-1}\phi X_o \tag{1.28}$$

$$U_2 = T_2^*\left[I + T_1T_1^* - T_2T_2^*\right]^{-1}\phi X_o \tag{1.29}$$

We note that the form of (1.19) is quite general and that the results above are valid for any abstract Hilbert space functional of this form.

When T_1 and T_2 are given by (1.12) and (1.13), then

$$T_1^*\left[I+T_1T_1^*-T_2T_2^*\right]^{-1} = G_1^*(t)\,\phi^*(T,t)\,\cdot \left[I + \int_o^T \phi(T,s)G_1(s)G_1^*(s)\phi^*(T,s)ds - \int_o^T \phi(T,s)G_2(s)G_2^*(s)\phi^*(T,s)ds\right]^{-1} \tag{1.30}$$

and

$$T_2^*\left[I+T_1T_1^*-T_2T_2^*\right]^{-1} = G_2^*(t)\,\phi^*(T,t)\,\cdot \left[I + \int_o^T \phi(T,s)G_1(s)G_1^*(s)\phi(T,s)ds - \int_o^T \phi(T,s)G_2(s)G_2^*(s)\phi^*(T,s)ds\right]^{-1} \tag{1.31}$$

If we define

$$K_1(t;t,T) = G_1^*(t)\phi^*(T,t)\left[I + \int_t^T \phi(T,s)G_1(s)G_1^*(s)\phi^*(T,s)ds - \int_t^T \phi(T,s)G_2(s)G_2^*(s)\phi^*(T,s)ds\right]^{-1} \phi(T,t) \tag{1.32}$$

and

$$K_2(t;t,T) = G_2^*(t)\phi^*(T,t)\left[I + \int_t^T \phi(T,s)G_1(s)G_1^*(s)\phi^*(T,s)ds - \int_t^T \phi(T,s)G_2(s)G_2^*(s)\phi^*(T,s)ds\right]^{-1} \phi(T,t) \tag{1.33}$$

then we may write

$$U_1(t) = K_1(t;0,T)X_o \tag{1.34}$$

$$U_2(t) = K_2(t;0,T)X_o \tag{1.35}$$

Here the arguments 0 and T of K_1 and K_2 indicate the initial and final times $t = 0$ and $t = T$, respectively.

Having seen this solution, we may state the problem somewhat differently: if we require that U_1 and U_2 be of the form $U_1 = K_1X_o$, then what are the transformations K_1 and K_2 which minimaximize the functional $J(U_1,U_2)$? In other words, we wish to find K_1 and K_2 which

$$\min_{K_1} \max_{K_2} J(K_1,K_2) \overset{d}{=} \frac{1}{2}\Big\{\Big\langle \phi X_o - T_1K_1X_o + T_2K_2X_o, \phi X_o - T_1K_1X_o + T_2K_2X_o\Big\rangle + \Big\langle K_1X_o,K_1X_o\Big\rangle - \Big\langle K_2X_o,K_2X_o\Big\rangle\Big\} \tag{1.36}$$

Now K_1 and K_2 are linear transformations from the Euclidean space containing X_o to the Hilbert spaces which are the domains of T_1 and T_2, respectively. We form the functional derivatives of $J(K_1,K_2)$ with respect to K_1 and K_2 as follows: let Δ_1 and Δ_2 be arbitrary linear transformations which have the same domains and ranges as K_1 and K_2, respectively, and let $s_1\Delta_1$ and $s_2\Delta_2$ be variations about K_1 and K_2, respectively, where s_1 and s_2 are scalars. Then, remembering the predicted miss distance $= \phi(T,0)X_o$,

$$J(K_1+s_1\Delta_1,K_2) = \frac{1}{2}\Big\{\Big\langle\Big[\phi-T_1(K_1+s_1\Delta_1)+T_2K_2\Big]X_o, \Big[\phi-T_1(K_1+s_1\Delta_1)+T_2K_2\Big]X_o\Big\rangle + \Big\langle(K_1+s_1\Delta_1)X_o, (K_1+s_1\Delta_1)X_o\Big\rangle - \Big\langle K_2X_o,K_2X_o\Big\rangle\Big\} \tag{1.37}$$

and, upon first expanding the above expression and then subtracting $J(K_1,K_2)$, we have

$$J(K_1+s_1\Delta_1,K_2) - J(K_1,K_2) = -s_1\Big\langle\Big[\phi-T_1K_1+T_2K_2\Big]X_o, T_1\Delta_1X_o\Big\rangle + \frac{1}{2}s_1^2\Big\langle T_1\Delta_1X_o,T_1\Delta_1X_o\Big\rangle + s_1\Big\langle K_1X_o,\Delta_1X_o\Big\rangle + \frac{1}{2}s_1^2\Big\langle\Delta_1X_o,\Delta_1X_o\Big\rangle \tag{1.38}$$

Dividing this expression by s_1 and letting s_1 approach zero, we have the functional derivative of $J(K_1,K_2)$ with respect to K_1, which is

$$\frac{\partial J(K_1,K_2)}{\partial K_1} = -\Big\langle\Big[\phi-T_1K_1+T_2K_2\Big]X_o,T_1\Delta_1X_o\Big\rangle + \Big\langle K_1X_o,\Delta_1X_o\Big\rangle \tag{1.39}$$

which may be written, using the properties of the adjoint operator T_1^* and combining terms, as

$$\frac{\partial J(K_1,K_2)}{\partial K_2} = -\Big\langle\Big(T_1^*\Big[\phi-T_1K_1+T_2K_2\Big] - K_1\Big)X_o,\Delta_1X_o\Big\rangle \tag{1.40}$$

A necessary condition then, that K_1 be a minimizing transformation, is

$$\left\langle \left(T_1^* \left[\phi - T_1 K_1 + T_2 K_2 \right] - K_1 \right) X_o , \Delta_1 X_o \right\rangle = 0 \tag{1.41}$$

Now, since Δ_1 was taken to be an <u>arbitrary</u> linear transformation, (1.41) implies that the vector

$$\left(T_1^* \left[\phi - T_1 K_1 + T_2 K_2 \right] - K_1 \right) X_o \tag{1.42}$$

is orthogonal to <u>any</u> linear transformation of X_o, which in turn implies that the transformation

$$T_1^* \left[\phi - T_1 K_1 + T_2 K_2 \right] - K_1 \tag{1.43}$$

is the null transformation. Thus for any vector X_o

$$T_1^* \left[\phi - T_1 K_1 + T_2 K_2 \right] X_o = K_1 X_o \tag{1.44}$$

This can be true only if $K_1 X_o$ is in the range of T_1^*; hence, we write $K_1 X_o = T_1^* \lambda_1$ for some λ_1 in the domain of T_1^*.

Following a line of reasoning similar to the above, after differentiating $J(K_1, K_2)$ with respect to K_2, we are led to

$$T_2^* \left[\phi - T_1 K_1 + T_2 K_2 \right] X_o = K_2 X_o \tag{1.45}$$

Hence, $K_2 X_o$ is in the range of T_2^*. We write $K_2 X_o = T_2^* \lambda_2$. Substituting $T_1^* \lambda_1$ and $T_2^* \lambda_2$ for $K_1 X_o$ and $K_2 X_o$, respectively, in the above pair of equations gives

$$T_1^*\phi X_o - T_1^*T_1T_1^*\lambda_1 + T_1^*T_2T_2^*\lambda_2 = T_1^*\lambda_1 \tag{1.46}$$

$$T_2^*\phi X_o - T_2^*T_1T_1^*\lambda_1 + T_2^*T_2T_2^*\lambda_2 = T_2^*\lambda_2 \quad . \tag{1.47}$$

which can be written

$$\phi X_o = T_1T_1^*\lambda_1 - T_2T_2^*\lambda_2 + \lambda_1 \tag{1.48}$$

$$\phi X_o = T_1T_1^*\lambda_1 - T_2T_2^*\lambda_2 + \lambda_2 \tag{1.49}$$

If $\lambda_1 = \lambda_2$, we have

$$\phi X_o = \left[I+T_1T_1^* - T_2T_2^*\right]\lambda_1 \tag{1.50}$$

or

$$\lambda_1 = \left[I+T_1T_1^* - T_2T_2^*\right]^{-1}\phi X_o \tag{1.51}$$

if the indicated inverse exists. Then, since $K_1X_o = T_1^*\lambda_1$, we have

$$K_1 = T_1^*\left[I+T_1T_1^* - T_2T_2^*\right]^{-1}\phi \tag{1.52}$$

and similarly

$$K_2 = T_2^*\left[I+T_1T_1^* - T_2T_2^*\right]^{-1}\phi \tag{1.53}$$

These expressions are what we expected: knowing that the overall optimal strategies are linear transformations of X_o; we are not surprised that when we ask which linear transformations of X_o are optimal we get as an answer the same (the overall optimal) transformations. However, the technique just employed can provide optimal <u>linear</u> strategies even when the form of the overall optimal strategies is not known.

Expressions (1.52) and (1.53) are open-loop optimal control strategies. Since we have temporarily assumed that the players are both able to monitor the state continuously, we may convert (1.52) and (1.53) to closed-loop or feedback type strategies by replacing ϕX_o with $\phi\phi^{-1}(t)X(t)$, $K_1(t;0,T)$ with $K_1(t;t,T)$, and $K_2(t;0,T)$ with $K_2(t;t,T)$. In this case, as the game progresses, the players constantly regard the present instant as the initial time of a new game and form their control functions accordingly.

Chapter 2
THE STOCHASTIC GAME PROBLEM

2.1 Preliminary Remarks

In some cases the players are not able to monitor the state continuously, but are able to make noisy observations of the state in the form given by (1.3A). If they are given only statistical information (1.1A) about the initial state X_o, then presumably they will be able to take advantage of their noisy measurements to improve the quality of their play over that of strictly open-loop strategies. Thus, we must find stochastically optimal stragegties, – methods by which the players process their observed data so that the expected value of the payoff functional is minimaximized with respect to the data processing methods. The players must find strategies which are optimal within the constraints of their limited information. This information includes the mean $\bar{X}_o$ and covariance Ψ_{X_o} of the initial state, plus the observations described by (1.3A). These quantities must be combined functionally to form the strategies U_1 and U_2. What the functional form should be will be determined by certain criteria of desirability, one of which is the so-called "certainty-coincidence" principle discussed by Willman [28]. This is simply a requirement that the stochastic strategies coincide with the deterministic strategies when the noise variances go to zero.

Other criteria are simplicity and physical realizability. Accordingly, we will require that the functional form of the strategies be a linear combination of the known quantities and the observables.

2.2 A Heuristic Justification for the Assumption of Linear Strategies

An interesting aspect of the selection of the form of the strategies for discrete-time games has been developed by K. Bley [6] and is extended here to the continuous-time case. We have hypothesized a criterion function of the form

$$J = E\left\{X^*(T)X(T) + \int_{t_o=0}^{T} (U_1^*U_1 - U_2^*U_2)dt\right\} \tag{2.1}$$

and a system equation which may be rewritten

$$dX = FXdt - G_1U_1dt + G_2U_2dt \tag{2.2}$$

We define

$$g = \min_{U_1(t)} \max_{U_2(t)} J \qquad t_o \le t \le T \tag{2.3}$$

For a given set of noise statistics and for minimax control strategies, the payoff will depend on $t_o = 0$ and $X(t_o) = X_o$; call this payoff $f\left(X(t_o), t\right)$. We write the minimax payoff as

$$g = E\left\{f\left(X(t_o), t_o\right)\right\} \tag{2.4}$$

Breaking the time interval $[0,T]$ into two sub-intervals $[0,\Delta]$ and $[\Delta,T]$, we may then write the criterion functional as

$$g = \min_{U_1} \max_{U_2} E\left\{ \int_0^\Delta (U_1^* U_1 - U_2^* U_2)dt + X^*(T)X(T) + \int_\Delta^T (U_1^* U_1 - U_2^* U_2)dt \right\}$$

$$= \min_{U_1} \max_{U_2} E\left\{ \int_0^\Delta (U_1^* U_1 - U_2^* U_2)dt + \min_{U_1} \max_{U_2} E\left\{X^*(T)X(T) + \int_\Delta^T (U_1^* U_1 - U_2^* U_2)dt\right\} \right\} \qquad (2.5)$$

Now

$$\min_{U_1} \max_{U_2} E\left\{X^*(T)X(T) + \int_\Delta^T (U_1^* U_1 - U_2^* U_2)dt\right\} = E\left\{f\left(X(0)+dX,\Delta\right)\right\} \qquad (2.6)$$

We expand f in a Taylor series about $\left(X(0),0\right)$

$$f\left(X(0) + dX,\Delta\right) = f\left(X(0),0\right) + \frac{\partial f}{\partial t_o}\Delta + \frac{\partial f}{\partial X}dX \qquad (2.7)$$

so

$$g = \min_{U_1} \max_{U_2} E\left\{ \int_0^\Delta (U_1^* U_1 - U_2^* U_2)dt + f\left(X(0),0\right) + \frac{\partial f}{\partial t_o}\Delta + \frac{\partial f}{\partial X}dX \right\} \qquad (2.8)$$

Since $\min_{U_1} \max_{U_2} E\left\{f\left(X(0),0\right)\right\} = g$, we may write the above as

$$0 = \min_{U_1} \max_{U_2} E\left\{\int_0^\Delta (U_1^* U_1 - U_2^* U_2)dt + \frac{\partial f}{\partial t_o}\Delta + \frac{\partial f}{\partial X}dX\right\} \tag{2.9}$$

Then, writing $dX = FX\Delta - G_1U_1\Delta + G_2U_2\Delta$ and substituting in the above,

$$0 = \min_{U_1} \max_{U_2} E\left\{\int_0^\Delta (U_1^* U_1 - U_2^* U_2)dt + \frac{\partial f}{\partial t_o}\Delta + \frac{\partial f}{\partial X}(FX - G_1U_1 + G_2U_2)\Delta\right\} \tag{2.10}$$

Approximating the integral by $(U_1^* U_1 - U_2^* U_2)\Delta$, we have

$$0 = \min_{U_1} \max_{U_2} E\left\{(U_1^* U_1 - U_2^* U_2)\Delta + \frac{\partial f}{\partial t_o}\Delta + \frac{\partial f}{\partial X}(FX - G_1U_1 + G_2U_2)\Delta\right\} \tag{2.11}$$

Dividing both sides by Δ, we have

$$0 = \min_{U_1} \max_{U_2} E\left\{U_1^* U_1 - U_2^* U_2 + \frac{\partial f}{\partial t_o} + \frac{\partial f}{\partial X}(FX - G_1U_1 + G_2U_2)\right\} \tag{2.12}$$

Since $\min_{U_1} \max_{U_2} E\left\{f\left(X(0),0\right)\right\} = g$,

$$\frac{\partial g}{\partial t_o} = \frac{\partial}{\partial t_o}\min_{U_1} \max_{U_2} E\left\{f\left(X(0),0\right)\right\} = \min_{U_1} \max_{U_2} E\left\{\frac{\partial f}{\partial t_o}\left(X(0),0\right)\right\} \tag{2.13}$$

Therefore,

$$\frac{\partial g}{\partial t_o} = \min_{U_1} \max_{U_2} E\left\{U_1^* U_1 - U_2^* U_2 + \frac{\partial f}{\partial X} (FX - G_1 U_1 + G_2 U_2)\right\} \tag{2.14}$$

We now must make an assumption about the form of $f(X,t)$; therefore we choose some general form, such as[1]

$$f(X,t) = X^* \lambda_0(t) X + U_1^* \lambda_1(t) X + U_2^* \lambda_2(t) + \phi(t) \tag{2.15}$$

where the $\lambda_i(t)$ $i = 0,1,2$ are unspecified matrices. Thus,

$$\frac{\partial f}{\partial X} = X^* \lambda_0(t) + U_1^* \lambda_1(t) + U_2^* \lambda_2(t) \tag{2.16}$$

Utilizing this expression for $\frac{\partial f}{\partial X}$ in (2.14) we have

$$\frac{\partial g}{\partial t_o} = \min_{U_1} \max_{U_2} E\Big\{U_1^* U_1 - U_2^* U_2 + (X^* \lambda_0 + U_1^* \lambda_1 + U_2^* \lambda_2) \cdot$$

$$(FX - G_1 U_1 + G_2 U_2)\Big\} \tag{2.17}$$

After collecting terms, the right-hand side of (2.17) may be rewritten as

$$\min_{U_1} \max_{U_2} E\Big\{X^* Q_0 X + U_1^* Q_1 U_1 + U_2^* Q_2 U_2 + X^* Q_4 U_1$$

$$+ X^* Q_5 U_2 + U_1^* Q_6 U_2\Big\} \tag{2.18}$$

---

[1] For a detailed examination of this subject see the dissertation of K.Bley [6] where the discrete-time version of the problem is analyz...

where

$$Q_0 = \lambda_0 F$$

$$Q_1 = \mathrm{SYM}\left\{I - \lambda_1' G_1\right\}$$

$$Q_2 = \mathrm{SYM}\left\{I + \lambda_2 G_2\right\}$$

$$Q_4 = F^* \lambda_1^* - \lambda_0 G_1$$

$$Q_5 = F^* \lambda_2^* + \lambda_0 G_2$$

$$Q_6 = \lambda_1 G_2 - G_1^* \lambda_2^* \tag{2.19}$$

and where $\mathrm{SYM}\left\{A\right\}$ denotes the symmetrized version of the positive definite matrix A. When differentiating the expression in brackets with respect to U_1, since U_1 is a minimizing control, we have

$$E\left\{U_1^* Q_1 + X^* Q_4 + U_2^* Q_6^*\right\} = 0 \tag{2.20}$$

Similarly,

$$E\left\{U_2^* Q_2 + X^* Q_5 + U_1^* Q_6\right\} = 0 \tag{2.21}$$

Now because U_1 is player 1's control, U_1 must be based only on the observation Z_1; similarly, U_2 must be based solely on Z_2. And since it is a property of conditional expectations that $E\left\{X\right\} = E\left\{E\left\{X \mid Z\right\}\right\}$ for random variables X and Z, we may write the above equations (2.20) and (2.21) as

$$E\left\{U_1^* Q_1 + X^* Q_4 + U_2^* Q_6^* \mid Z_1\right\} = 0 \tag{2.22}$$

$$E\left\{U_2^*Q_2 + X^*Q_5 + U_1^*Q_6 \mid Z_2\right\} = 0 \tag{2.23}$$

Furthermore, since $U_1 = U_1(Z_1)$ and $U_2 = U_2(Z_2)$ and because the taking of conditional expectations is a linear operation, we may write

$$U_1^*Q_1 + E\left\{X^* \mid Z_1\right\} Q_4 + E\left\{U_2^* \mid Z_1\right\} Q_6^* = 0 \tag{2.24}$$

$$U_2^*Q_2 + E\left\{X^* \mid Z_2\right\} Q_5 + E\left\{U_1^* \mid Z_2\right\} Q_6 = 0 \tag{2.25}$$

or

$$U_1 = -Q_1^{-1}Q_4^* \, E\left\{X \mid Z_1\right\} - Q_1^{-1}Q_6 \, E\left\{U_2 \mid Z_1\right\}. \tag{2.26}$$

and

$$U_2 = -Q_2^{-1}Q_5 \, E\left\{X \mid Z_2\right\} - Q_2^{-1}Q_6^* \, E\left\{U_1 \mid Z_2\right\} \tag{2.27}$$

Now if we denote by $T_1(\cdot)$ the linear operation

$$T_1(\cdot) = E\left\{\cdot \mid Z_1\right\} \tag{2.28}$$

and similarly for $T_2(\cdot)$

$$T_2(\cdot) = E\left\{\cdot \mid Z_2\right\} \tag{2.29}$$

the equations then read

$$U_1 = -Q_1^{-1}Q_4^*T_1X - Q_1^{-1}Q_6T_1U_2 \tag{2.30}$$

$$U_2 = -Q_2^{-1}Q_5T_2X - Q_2^{-1}Q_6^*T_2U_1 \tag{2.31}$$

and, substituting the second equation into the first,

$$U_1 = -Q_1^{-1}Q_4^*T_1X + Q_1^{-1}Q_6T_1Q_2^{-1}Q_5T_2X + Q_1^{-1}Q_6T_1Q_2^{-1}Q_6^*T_2U_1 \tag{2.32}$$

The expected value operators commute with the matrix operators Q_i, $i = 1,2,\ldots,6$, so

$$(I-Q_1^{-1}Q_6Q_2^{-1}Q_6^*T_1T_2)U_1 = Q_1^{-1}\left[Q_6Q_2^{-1}Q_5T_1T_2-Q_4^*T_1\right]X \tag{2.33}$$

If the norm of the operator $Q_1^{-1}Q_6Q_2^{-1}Q_6^*T_1T_2$ is less than unity, a Neumann expansion gives the inverse of $I - Q_1^{-1}Q_6Q_2^{-1}Q_6^*T_1T_2$, so

$$U_1 = (I-Q_1^{-1}Q_6Q_2^{-1}Q_6^*T_1T_2)^{-1}\, Q_1^{-1}\left[Q_6Q_2^{-1}Q_5T_1T_2-Q_4^*T_1\right]X \tag{2.34}$$

The above expression gives U_1 in terms of conditional expectations of the state vector X. A similar expression exists for U_2. We have considered only the starting point, but any point may be considered the starting point of a new game.

These expressions for the minimax strategies in terms of conditional expectations of the state indicate that when the process statistics are Gaussian the optimal strategies are linear (affine), since the conditional expectation of the state is a linear transformation of the observations. We might interpret this to mean that, when pitted against an opponent who is known to use linear strategies, the optimal counter-strategy is itself linear. However, the proof makes such essential use of the Gaussian character of the process statistics that it becomes invalid when either player at any instant uses a nonlinear control which would destroy the Gaussian probability distribution. When the Gaussian distribution is thought to be a reasonable approximation to the true distribution, the restriction to linear strategies is perhaps justifiable. Furthermore, since the solution to the deterministic problem is known to involve linear state feedback as control strategies, it is intuitively reasonable to believe that for small uncertainties in the state information the linear certainty-equivalent strategies cannot be too far from optimal. Thus, the class of general linear control strategies must contain strategies which, if not overall optimal, are at least bounded by the certainty-equivalent strategies in payoff. In practical situations, if the system designer has some confidence that a linear strategy will give nearly optimal performance, he can justify restriction of his design to linear strategies on the basis of computational feasibility considerations.

A final word about the form of the strategies: since any strategy which minimizes the expected value of the payoff must in some way depend on the probability distribution of the state variables, the task of selecting a strategy which is generally optimal against any form of opposing strategy is rather hopeless, since that opposing strategy may alter the probability distribution of the state variables in such a way as to give each player a different notion of what that probability distribution is. It is with a view to the futility of searching for the perfectly optimal strategy that we gladly restrict our attention to the task of finding an optimal linear strategy. We shall soon see that even this restriction is not sufficient to insure that the resulting control functionals are computationally feasible.

We have required that the strategies be simple, linear, and physically realizable. A pair of general expressions meeting these requirements is

$$U_1 = M_1\hat{X}_1 + N_1Z_1 \tag{2.35}$$

$$U_2 = M_2\hat{X}_2 + N_2Z_2 \tag{2.36}$$

A straightforward approach to the game problem might be to assume strategies of this form, to substitute these expressions in the payoff functional, and to proceed with the optimization over the class of linear functionals M_1, N_1, M_2, and N_2. Then, if the certainty-equivalent strategy were optimal, we would expect to find that $M_1 = K_1$ and $M_2 = K_2$, while N_1 and N_2 are zero. While the proposed approach is in fact a poor one if useful solutions to the game problem are desired, some revealing facts are brought to light by taking it, and we shall therefore do so.

But before proceeding, we point out two facts:

i) We have tacitly assumed that the conditional means $\hat{X}_1$ and $\hat{X}_2$ are computable by the players, but we have not specified how the computation would be done.

ii) We have asked that the "certainty-coincidence principle" be satisfied. Thus, in terms of the forms we have assumed, we require

$$M_1 \to K_1 \qquad M_2 \to K_2$$

$$N_1 \to 0 \qquad N_2 \to 0 \tag{2.37}$$

as the observation noise covariances go to zero. In view of this requirement, we may rearrange (2.35) and (2.36) into a more convenient form. We write

$$\begin{aligned} U_1 &= M_1\hat{X}_1 + N_1Z_1 = (M_1+N_1H_1)\,\hat{X}_1 + N_1\,(Z_1-H_1\hat{X}_1) \\ &= K_1\left[\hat{X}_1+L_1(Z_1-H_1\hat{X}_1)\right] \\ &= K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] \quad . \end{aligned} \tag{2.38}$$

where $K_1 = M_1 + N_1H_1$, $\hat{Z}_1 = H_1\hat{X}_1$, and $K_1L_1 = N_1$. Similarly, we write

$$U_2 = K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right] \tag{2.39}$$

and require K_1 and K_2 to approach the deterministic feedback gain as the observation noises go to zero. Thus, our assumed strategies have the form of linear transformations of the conditional mean plus linear operations on the residuals. The payoff functional becomes

$$\begin{aligned} J(K_1,L_1,K_2,L_2) = \frac{1}{2}\,E\Big\{ &\phi X - T_1K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] + T_2K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right], \\ &\phi X - T_1K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] + T_2K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right] \\ &+ K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right], K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] \\ &- K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right], K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right] \Big\} \end{aligned} \tag{2.40}$$

and we wish to find K_1, L_1, K_2, and L_2, which provide a saddle point of the functional (2.40).

However, before proceeding further, it is desirable to pause and develop the techniques we will need for handling such problems, i.e., finding functional derivatives of expected value functionals. To illustrate these techniques, we will derive some well-known relationships which will be found useful later in this exposition.

2.3 Minimum Variance Estimation

The first example we treat is that of minimum-variance estimation. We wish to estimate a vector X on the basis of our observation of another vector Z. We assume knowledge of the mean $\bar{X}$ and the variance ψ_{XX} of X and of the covariance of X and Z, ψ_{XZ}. We also assume knowledge of the mean $\bar{Z}$ and variance ψ_{ZZ} of Z. We ask that our estimator be linear and realizable and that it obey the certainty-coincidence principle. We thus assume the estimator has the form

$$\hat{X}_1 = \bar{X} + L(Z-\bar{Z}) \tag{2.41}$$

where $\hat{X}_1$ denotes the estimate and L is a linear operator. The estimation error is

$$\epsilon_1 = X - \hat{X}_1 = X - \bar{X} - L(Z-\bar{Z}) \tag{2.42}$$

and the variance of the error may be written as a functional of L

$$J(L) = \frac{1}{2} E\left\{\left\langle X-\bar{X}-L(Z-\bar{Z}),\ X-\bar{X}-L(Z-\bar{Z})\right\rangle\right\} \tag{2.43}$$

Forming the functional derivative of this functional with respect to the operator L and setting this equal to zero, we have

$$E\left\{\left\langle X-\bar{X}-L(Z-\bar{Z}),\ \Delta(Z-\bar{Z})\right\rangle\right\} = 0 \tag{2.44}$$

where, as before, Δ is any arbitrary linear operation on the observation $Z-\bar{Z}$. We interpret (2.44) to mean that the expression

$$X - \bar{X} - L(Z-\bar{Z}) \tag{2.45}$$

is statistically orthogonal to any linear transformation of the observation $Z-\bar{Z}$. In order for this to be true, $Z-\bar{Z}$ must be uncorrelated with (2.45); i.e.,

$$E\left\{\left[X-\bar{X}-L(Z-\bar{Z})\right]\left[Z-\bar{Z}\right]^*\right\} = \Psi_{XZ} - L\Psi_{ZZ} = 0 \tag{2.46}$$

This is an abstract form of the Wiener-Hopf equation describing the linear estimate which is optimal in the mean square sense. It is well known that when the random variables are normally distributed the linear estimate is over-all optimal. Furthermore, since the optimal mean square estimate is the conditional mean of the random variable to be estimated, we see that (2.41) provides us with the conditional mean of X when X and Z are normally distributed and L satisfies (2.46). We note that when the form of L is specified as

$$L(Z-\bar{Z}) = \int_{-\infty}^{t} W(t,\tau)\left[Z(\tau) - \bar{Z}(\tau)\right] \tag{2.47}$$

then the Wiener-Hopf equation takes its familiar form

$$\psi_{XZ}(t,\sigma) = \int_{-\infty}^{t} W(t,\tau)\ \psi_{ZZ}(t,\sigma)d\tau \tag{2.48}$$

2.4 A Stochastic Optimal Regulator

We next treat a more complicated problem: the stochastic optimal regulator. This is the one-player version of our stochastic differential game. We first look at the deterministic case, which when cast in abstract Hilbert space form appears as the following minimization problem:

$$\min_{U}\ J(U) = \frac{1}{2}\left\{\left\langle \phi X-TU,\ \phi X-TU\right\rangle + \left\langle U,U\right\rangle\right\} \tag{2.49}$$

Differentiating J(U) with respect to U and setting the derivative equal to zero, we have

$$-T^{*}\phi X + T^{*}TU + U = 0 \tag{2.50}$$

This equation has a solution only if U is in the range of T^{*}, or $U = T^{*}\lambda$ for some λ . Substituting this into (2.50), we have

$$-T^{*}\phi X + T^{*}TT^{*}\lambda + T^{*}\lambda = 0 \tag{2.51}$$

Equation (2.51) will be satisfied if

$$\left[TT^* + I\right]\lambda = \phi X \tag{2.52}$$

or $\lambda = \left[TT^* + I\right]^{-1} \phi X$, if the indicated inverse exists, which would imply

$$U = T^*\left[TT^* + I\right]^{-1} \phi X \tag{2.53}$$

When the system under consideration is a continuous-time dynamical system described by the differential equation

$$\dot{X}(t) = F(t)X(t) + G(t)U(t) \qquad X(0) = X_o \tag{2.54}$$

then TU takes the form

$$TU = \int_0^T \phi(T)\, \phi^{-1}(\tau)G(\tau)U(\tau)d\tau \tag{2.55}$$

and

$$TT^* = \int_0^T \phi(T)\, \phi^{-1}(\tau)G(\tau)\, G^*(\tau)\phi^*(T)\phi^{-1^*}(\tau)d\tau \tag{2.56}$$

This is recognized as the controllability matrix of the system. Thus, if the system is controllable, TT^* is positive definite and the existence of $\left[TT^* + I\right]^{-1}$ is assured.

We might now ask the question: "Of all controls of the form $U = KX$, which linear transformation K minimizes the functional $J(K)$ where

$$J(K) = \frac{1}{2} \left\{ \left\langle (\Phi - TK)\, X, (\Phi - TK) X \right\rangle + \left\langle KX, KX \right\rangle \right\} \text{?"} \tag{2.57}$$

By a procedure similar to that used with the differential game, we would find that the optimal K has the form

$$K = T^{*} \left[TT^{*} + I \right]^{-1} \Phi \tag{2.58}$$

This is not a surprising answer in view of the previous result.

We may now consider the stochastic version of this problem. Assume that we do not know X exactly, but do know its conditional mean $\hat{X}$ and its conditional covariance ψ_{XX}, these quantities being conditioned on the observation of a correlated random variable Z. The correlation between X and Z is denoted ψ_{XZ}. Random variable Z has conditional mean $\hat{Z}$ and conditional covariance ψ_{ZZ}, these quantities being conditioned on the observed history of Z. We invoke the certainty-coincidence principle and the criteria of simplicity and realizability to postulate the form of U as

$$U = K\left[\hat{X} + L(Z - \hat{Z})\right] \tag{2.59}$$

We thus ask for the values of K and L which minimize the functional

$$J(K,L) \doteq \frac{1}{2} E\Big\{\Big\langle \phi X - TK\big[\hat{X}+L(Z-\hat{Z})\big],\ \phi X - TK\big[\hat{X}+L(Z-\hat{Z})\big]\Big\rangle$$

$$+\Big\langle K\big[\hat{X}+L(Z-\hat{Z})\big],\ K\big[\hat{X}+L(Z-\hat{Z})\big]\Big\rangle\Big\} \tag{2.60}$$

Forming the derivative of this functional with respect to L and setting this equal to zero, we have

$$\frac{\partial}{\partial L} J(K,L) = K^*T^*TK\big[\hat{X}+L(Z-\hat{Z})\big] - K^*T^*\phi X + K^*K\big[\hat{X}+L(Z-\hat{Z})\big] = 0 \tag{2.61}$$

which will be satisfied if

$$\big[T^*T + I\big] K\big[\hat{X}+L(Z-\hat{Z})\big] - T^*\phi X = 0 \tag{2.62}$$

We interpret this to mean that the expression (2.62) above is orthogonal to any linear transformation of the quantity $Z - \hat{Z}$. In particular, (2.62) is orthogonal to $\big[(T^*T+I)\ K - T^*\phi\big]\ L(Z-\hat{Z})$, and we may express this by

$$E\Big\{\Big\langle \big[T^*T+I\big] K\big[\hat{X}+L(Z-\hat{Z}\big] - T^*\phi X,\ \big[(T^*T+I)\ K - T^*\phi\big]\ L(Z-\hat{Z})\Big\rangle\Big\} = 0 \tag{2.63}$$

We may also differentiate (2.60) with respect to the transformation K and set this equal to zero. Doing this, we have

$$\frac{\partial}{\partial K} J(K,L) = T^*TK\big[\hat{X}+L(Z-\hat{Z})\big] - T^*\phi X + K\big[\hat{X}+L(Z-\hat{Z})\big] = 0 \tag{2.64}$$

Again we interpret this to mean that (2.64) is orthogonal to any linear transformation of the quantity $\hat{X} + L(Z-\hat{Z})$, in particular to $\left[(T^*T+I)\,K - T^*\phi\right]\left[\hat{X}+L(Z-\hat{Z})\right]$. This we express as

$$E\left\{\left\langle\left[T^*T+I\right] K\left[\hat{X}+L(Z-\hat{Z})\right] - T^*\phi X, \left[(T^*T+I)\,K - T^*\phi\right]\left[\hat{X}+L(Z-\hat{Z})\right]\right\rangle\right\} = 0 \tag{2.65}$$

Combining (2.63) and (2.65), we have

$$E\left\{\left\langle\left[T^*T+I\right] K\left[\hat{X}+L(Z-\hat{Z})\right] - T^*\phi X, \left[(T^*T+I)\,K - T^*\phi\right] \hat{X}\right\rangle\right\} = 0 \tag{2.66}$$

We may write $X = \hat{X} + \epsilon$; and then, using the fact that estimation error is orthogonal to any linear transformation of the conditional mean (for normal random variables), we rewrite (2.66) as

$$E\left\{\left\langle\left[T^*T+I\right] K\left[\hat{X}+L(Z-\hat{Z})\right] - T^*\phi\hat{X}; \left[(T^*T+I)\,K - T^*\phi\right] \hat{X}\right\rangle\right\} = 0 \tag{2.67}$$

or, defining $\Lambda = (T^*T+I)K$, we write (2.67) as

$$E\left\{\left\langle(\Lambda-T^*\phi)\hat{X} + \Lambda L(Z-\hat{Z}), (\Lambda-T^*\phi)\,\hat{X}\right\rangle\right\} = 0 \tag{2.68}$$

Proper interpretation of (2.68) implies that

$$E\left\{\left\langle\Lambda\left[\hat{X}+L(Z-\hat{Z})\right] - T^*\phi X,\ \Lambda L(Z-\hat{Z})\right\rangle\right\} = 0 \tag{2.69}$$

or, again writing $X = \hat{X} + \epsilon$ and noting that the estimation error is orthogonal to all linear transformations of the observables (for normal random variables), we may write (2.69) as

$$E\left\{\left\langle (\Lambda - T^*\phi)\,\hat{X} + \Lambda L(Z-\hat{Z}),\ \Lambda L(Z-\hat{Z})\right\rangle\right\} = 0 \tag{2.70}$$

Subtracting (2.68) from (2.70), we have

$$E\left\{\left\langle \Lambda L(Z-\hat{Z}),\ \Lambda L(Z-\hat{Z})\right\rangle - \left\langle (\Lambda - T^*\phi)\,\hat{X}, (\Lambda - T^*\phi)\,\hat{X}\right\rangle\right\} = 0 \tag{2.71}$$

Now, since the first term on the left depends only on the covariances of observation noise and initial values of X and the second term depends on the mean initial value, for (2.71) to be satisfied we must have

$$\Lambda - T^*\phi = (T^*T+I)\,K - T^*\phi = 0 \tag{2.72}$$

which implies

$$\Lambda L = (T^*T+I)\,KL = 0 \tag{2.73}$$

Equation (2.72) is the relation which described the feedback gain K for the deterministic regulator problem, so the solution of (2.72) is known to be

$$K = T^*\left[TT^* + I\right]^{-1}\phi \tag{2.74}$$

Substituting this expression into (2.73), we have after some manipulation

$$T^*\phi' L = 0 \tag{2.75}$$

This will be satisfied if $L = 0$. Thus the stochastic optimal controller is

$$U_{\text{STOCH OPT}} = K\hat{X} = T^*\left[TT^*I\right]^{-1} \phi\hat{X} \tag{2.76}$$

We have obtained a weakened version of the separation theorem; i.e., we assumed a control function of the form

$$U = K\left[\hat{X}+L(Z-\hat{Z})\right] \; ; \; \hat{X} = \text{Best linear estimate of } X \tag{2.77}$$

and found that $L = 0$, and K is equal to the feedback gain matrix of the deterministically optimal control.

2.5 A Stochastic Differential Game - Special Case

We now return to the stochastic game problem, having developed some techniques and insights which will prove useful. The functional we wish to minimaximize is given by:

$$\begin{aligned} J(K_1,L_1,K_2,L_2) = \tfrac{1}{2} E\Big\{ & \Big\langle \phi X - T_1K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] + T_2K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right], \\ & \phi X - T_1K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] + T_2K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right]\Big\rangle \\ & + \left\langle K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right], K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right]\right\rangle \\ & - \left\langle K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right], K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right]\right\rangle\Big\} \end{aligned} \tag{2.78}$$

In some special cases it is possible to simplify this problem by decoupling, so that each player solves an independent stochastic optimal regulator problem. Therefore, before proceeding to the general problem, we examine one of these special cases. By making the following changes of variables,

$$T_2 \to (I+T_1T_1^*)T_2' \, (I+T_2'^*T_1T_1^*T_2')^{-\frac{1}{2}} \tag{2.79}$$

$$K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] \to K_1'\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] + T_1^*T_2'K_2'\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right] \tag{2.80}$$

$$K_2 \to (I+T_2'^*T_1T_1^*T_2')^{\frac{1}{2}}K_2' \tag{2.81}$$

the payoff functional becomes

$$\begin{aligned}
J(K_1',L_1,K_2',L_2) = \tfrac{1}{2} E\Big\{ & \Big\langle \phi X - T_1K_1'\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] + T_2'K_2'\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right], \\
& \phi X - T_1K_1'\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] + T_2'K_2'\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right]\Big\rangle \\
& + \Big\langle K_1'\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right], K_1'\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] \Big\rangle \\
& + \Big\langle T_1^*T_2'K_2'\left[\hat{X}_2+L_2(Z_1-\hat{Z}_1)\right], T_1^*T_2'K_2'\left[\hat{X}_2+L_2(Z_1-\hat{Z}_1)\right]\Big\rangle \\
& + 2\Big\langle K_1'\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right], T_1^*T_2'K_2'\left[\hat{X}_2+L_2(Z_1-\hat{Z}_1)\right]\Big\rangle \\
& - \Big\langle (I+T_2'^*T_1T_1^*T_2')^{\frac{1}{2}} K_2'\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right], \\
& (I+T_2'^*T_1T_1^*T_2')^{\frac{1}{2}} K_2'\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right]\Big\rangle\Big\}
\end{aligned} \tag{2.82}$$

Differentiating this functional with respect to L_1 and L_2 and setting the results equal to zero, we have

$$\frac{\partial J}{\partial L_1} = K_1^* T_1^* T_1 K_1' \left[\hat{X}_1 + L_1(Z_1 - \hat{Z}_1)\right] - K_1'^* T_1^* \phi X$$

$$+ K_1'^* K_1' \left[\hat{X}_1 + L_1(Z_1 - \hat{Z}_1)\right] = 0 \qquad (2.83)$$

$$\frac{\partial J}{\partial L_2} = K_2'^* T_2'^* T_2' K_2' \left[\hat{X}_2 + L_2(Z_2 - \hat{Z}_2)\right] + K_2'^* T_2' \phi X$$

$$- K_2'^* K_2' \left[\hat{X}_2 + L_2(Z_2 - \hat{Z}_2)\right] = 0 \qquad (2.84)$$

Differentiating with respect to K_1' and K_2', we have .

$$\frac{\partial K}{\partial K_1'} = T_1^* T_1 K_1' \left[\hat{X}_1 + L_1(Z_1 - \hat{Z}_1)\right] - T_1^* \phi X + K_1' \left[\hat{X}_1 + L_1(Z_1 - \hat{Z}_1)\right] = 0 \quad (2.85)$$

$$\frac{\partial J}{\partial K_2'} = T_2' T_2' K_2' \left[\hat{X}_2 + L_2(Z_2 - \hat{Z}_2)\right] + T_2' \phi X - K_2' \left[\hat{X}_2 + L_2(Z_2 - \hat{Z}_2)\right] = 0 \quad (2.86)$$

These equations are seen to be independent and identical in form to those of the stochastic optimal regulator problem. Thus, the two players play the transformed game using minimum-variance type state estimators, transforming their strategies back to the original game by use of the transformation equations. However, this solution is limited in usefulness in that it requires player 1 to know the quantity $Z_2 - \hat{Z}_2$, his opponent's observation, a circumstance which would rarely be true. This result is essentially that of Behn and Ho [3] but is a slight generalization in that neither player need have exact knowledge of the initial state.

2.6 A More General Stochastic Differential Game

In the case of the general stochastic differential game, it can be shown by techniques similar to those used in equations (2.61) through (2.74) that the minimax values of K_1 and K_2 are given by

$$K_1 = T_1^*\left[I+T_1T_1^*-T_2T_2^*\right]^{-1} \phi(T,to) \tag{2.87}$$

$$K_2 = T_2^*\left[I+T_1T_1^*-T_2T_2^*\right]^{-1} \phi(T,to) \tag{2.88}$$

These are seen to be the deterministically optimal feedback gains. Analogous to (2.75), but considerably more complicated, are the equations describing L_1 and L_2

$$\left[\phi+T_2K_2\right]L_1 \psi_{Z_1Z_1} - T_2K_2L_2 \psi_{Z_2Z_1} = -T_2K_2 \psi_{\epsilon_2Z_1} \tag{2.89}$$

$$T_1K_1L_1\psi_{Z_2Z_1} + \left[\phi-T_1K_1\right] L_2\psi_{Z_2Z_2} = T_1K_1\psi_{\epsilon_1Z_2} \tag{2.90}$$

The above equations are necessary conditions which must be satisfied by linear operations on noisy state observations which make up part of the strategies assumed in (2.38) and (2.39). The derivation of equations (2.87) through (2.90) is given in the Appendix.

We may make the following observations at this point:

i) $L_1 = L_2 = 0$ is not a solution to this set of equations; hence, we see that the optimal linear strategy is not a certainty-equivalent strategy.

ii) As $\psi_{Z_1Z_1}$, $\psi_{Z_2Z_2}$, $\psi_{\epsilon_1Z_2}$, and $\psi_{\epsilon_2Z_1}$ become small, $L_1 = L_2 = 0$ tends to more nearly satisfy (2.89) and (2.90); hence, our solution satisfies the certainty coincidence principle.

We may illustrate the use of the theory just developed by a simple example due to Willman [28].

2.7 A Simple Example

Example: Discrete-time, one-stage scalar game

Transition equation: $Y = X + U - V$

Payoff functional: $J = \frac{1}{2} E\left\{aY^2 + U^2 - cV^2\right\} \quad c > a > 0$

Observation equations:

$$Z_1 = X + \eta_1$$

$$Z_2 = X + \eta_2$$

$$\begin{bmatrix}\eta_1 \\ \eta_2\end{bmatrix} \text{ is normal; } \left(\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}R_1 & 0 \\ 0 & R_2\end{bmatrix}\right) \text{ ; } X \text{ is normal; } (0,P)$$

Making the following definitions and changes of variables,

$$U_1 = U \qquad T_1U_1 = -\sqrt{a}\,U_1 \qquad \phi X = \sqrt{a}\,X$$

$$U_2 = \sqrt{c}\,V \qquad T_2U_2 = -\sqrt{\frac{a}{c}}\,U_2$$

under these transformations the problem statement becomes:

<u>Transition equation:</u> $Y = X + U_1 - \frac{1}{\sqrt{c}} U_2$

<u>Payoff functional:</u> $J = \frac{1}{2} E\left\{(\phi X - T_1 U_1 + T_2 U_2)^2 + U_1^2 - U_2^2\right\}$

<u>Observation equations:</u> $Z_1 = X + \eta_1$

$$Z_2 = X + \eta_2$$

$\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$ is normal; $\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}$; X is normal; (0,P)

We first derive expressions for the feedback gains K_1 and K_2

$$K_1 = T_1^*\left[I + T_1T_1^* - T_2T_2^*\right]^{-1} \phi = -\sqrt{a}\left[1 + a - \frac{a}{c}\right]^{-1}\sqrt{a} = -\frac{ac}{c(1+a)-a}$$

$$K_2 = T_2^*\left[I + T_1T_1^* - T_2T_2^*\right]^{-1} \phi = -\sqrt{\frac{a}{c}}\left[1 + a - \frac{a}{c}\right]^{-1}\sqrt{a} = -\frac{a\sqrt{c}}{c(1+a)-a}$$

Then

$$T_1K_1 = \frac{\sqrt{aac}}{c(1+a)-a} \; ; \; \phi - T_1K_1 = \frac{\sqrt{a}\,(c-a)}{c(1+a)-a}$$

$$T_2K_2 = \frac{\sqrt{aa}}{c(1+a)-a} \; ; \; \phi + T_2K_2 = \frac{\sqrt{a}\,c(a+1)}{c(1+a)-a}$$

The filter equations then become

$$\frac{\sqrt{a}\,c(a+1)}{c(a+1)-a} L_1 \psi_{Z_1Z_1} - \frac{\sqrt{aa}}{c(a+1)-a} L_2 \psi_{Z_2Z_1} = -\frac{\sqrt{aa}}{c(a+1)-a} \psi_{\epsilon_2 Z_1} + \frac{\sqrt{a}\,c(a+1)}{c(a+1)-a} \psi_{\epsilon_1 Z_1}$$

$$\frac{\sqrt{a}\,ac}{c(a+1)-a} L_1\psi_{Z_1Z_2} + \frac{\sqrt{a}\,(c-a)}{c(a+1)-a} L_2\psi_{Z_2Z_2} = \frac{\sqrt{a}\,ac}{c(a+1)-a}\psi_{\epsilon_1Z_2} + \frac{\sqrt{a}\,(c-a)}{c(a+1)-a}\psi_{\epsilon_2Z_2}$$

which simplify to

$$c(a+1)\, L_1\psi_{Z_1Z_1} - a\, L_2\psi_{Z_2Z_1} = -a\,\psi_{\epsilon_2Z_1} + c(a+1)\,\psi_{\epsilon_1Z_1}$$

$$a\,c\,L_1\,\psi_{Z_1Z_2} + (c-a)\,L_2\,\psi_{Z_2Z_2} = a\,c\,\psi_{\epsilon_1Z_2} + (c-a)\,\psi_{\epsilon_2Z_2}$$

We must now derive expressions for the conditional means and covariances:

$$\hat{X}_1 = \frac{P}{P+R_1} Z_1 \,;\; \hat{Z}_1 = \hat{X}_1 \,;\; Z_1 - \hat{Z}_1 = Z_1 - \frac{P}{P+R_1} Z_1 = \frac{R_1}{P+R_1} Z_1$$

$$\hat{X}_2 = \frac{P}{P+R_2} Z_2 \,;\; \hat{Z}_2 = \hat{X}_2 \,;\; Z_2 - \hat{Z}_2 = Z_2 - \frac{P}{P+R_2} Z_2 = \frac{R_2}{P+R_2} Z_2$$

$$\epsilon_1 = X - \hat{X}_1 = X - \frac{P}{P+R_1}(X+\eta_1) = \frac{R_1X - P\eta_1}{P+R_1}$$

$$\epsilon_2 = X - \hat{X}_2 = X - \frac{P}{P+R_2}(X+\eta_2) = \frac{R_2X - P\eta_2}{P+R_2}$$

$$\psi_{Z_1Z_1} = E\left\{(Z_1-\hat{Z}_1)^2\right\} = \frac{R_1^2}{(P+R_1)^2} E\left\{(X+\eta_1)^2\right\} = \frac{R_1^2}{P+R_1}$$

$$\psi_{Z_1Z_2} = E\left\{(Z_1-\hat{Z}_1)(Z_2-\hat{Z}_2)\right\} = \frac{R_1R_2}{(P+R_1)(P+R_2)} E\left\{(X+\eta_1)(X+\eta_2)\right\}$$

$$= \frac{PR_1R_2}{(P+R_1)(P+R_2)} = \psi_{Z_2Z_1}$$

$$\psi_{Z_2Z_2} = E\left\{(Z_2-\hat{Z}_2)^2\right\} = \frac{R_2^2}{(P+R_2)^2} E\left\{(X+\eta_2)^2\right\} = \frac{R_2^2}{(P+R_2)}$$

$$\psi_{\epsilon_1 Z_1} = E\left\{(X-\hat{X}_1)(Z_1-\hat{Z}_1)\right\} = E\left\{ \frac{R_1X - P\eta_1}{(P+R_1)} \cdot \frac{R_1}{(P+R_1)} (X+\eta_1)\right\}$$

$$= \frac{R_1^2P - PR_1^2}{(P+R_1)^2} = 0$$

$$\psi_{\epsilon_1 Z_2} = E\left\{(X-\hat{X}_1)(Z_2-\hat{Z}_2)\right\} = E\left\{ \frac{R_1X - P\eta_1}{(P+R_1)} \cdot \frac{R_2}{(P+R_2)} (X+\eta_2)\right\}$$

$$= \frac{PR_1R_2}{(P+R_1)(P+R_2)}$$

$$\psi_{\epsilon_2 Z_1} = E\left\{(X-\hat{X}_2)(Z_1-\hat{Z}_1)\right\} = E\left\{\frac{(R_2X-P\eta_2)}{(P+R_2)} R_1 \frac{(X+\eta_1)}{(P+R_1)}\right\} = \frac{PR_1R_2}{(P+R_1)(P+R_2)}$$

$$\psi_{\epsilon_2 Z_2} = E\left\{(X-\hat{X}_2)(Z_2-\hat{Z}_2)\right\} = E\left\{ \frac{R_2X - P\eta_2}{(P+R_2)} \cdot \frac{R_2}{(P+R_2)} \cdot (X+\eta_2)\right\}$$

$$= \frac{R_2^2P - PR_2^2}{(P+R_2)^2} = 0$$

Substituting these expressions into the equations describing L_1 and L_2, we have

$$c(a+1)\, L_1 \frac{R_1^2}{P+R_1} - a\, L_2 \frac{PR_1R_2}{(P+R_1)(P+R_2)} = \frac{-aPR_1R_2}{(P+R_1)(P+R_2)}$$

$$acL_1 \frac{PR_1R_2}{(P+R_1)(PR_1R_2)} + (c-a)\, L_2 \frac{R_2^2}{P+R_2} = ac \frac{PR_1R_2}{(P+R_1)(P+R_2)}$$

which simplify to

$$cR_1(a+1)(P+R_2)\, L_1 - aPR_2L_2 = -aPR_2$$

$$acPR_1L_1 + R_2(c-a)(P+R_1)L_2 = acPR_1$$

which have solutions

$$L_1 = \frac{aP[acPR_1 - (c-a)(+R_1)R_2]}{cR_1[P+R_1)(P+R_2)(a+1)(c-a) + a^2P^2]}$$

$$L_2 = \frac{aP[aPR_2 + (a+1)(P+R_2)cR_1]}{R_2[(P+R_1)(P+R_2)(a+1)(c-a) + a^2P^2]}$$

We may now derive expressions for the control functions:

$$U_1 = K_1\left[\hat{X}_1 + L_1(Z_1-\hat{Z}_1)\right] = \frac{-ac}{(c(a+1)-a)}$$

$$\cdot\left[\frac{P}{P+R_1}Z_1 + \frac{aP[acPR_1 - (c-a)(P+R_1)R_2]}{cR_1[(P+R_1)(P+R_2)(a+1)(c-a) + a^2P^2]}\cdot\frac{R_1Z_1}{(P+R_1)}\right]$$

$$= \frac{-aP}{(P+R_1)(c(a+1)-a)}\left[1 + \frac{a[acPR_1 - (c-a)(P+R_1)R_2}{[(P+R_1)(P+R_2)(a+1)(c-a) + a^2P^2]}\right]Z_1$$

After some manipulation, this becomes

$$U_1 = \frac{-aP[(c-a)(P+R_2) + aP]}{[(P+R_1)(P+R_2)(a+1)(c-a) + a^2P^2]}Z_1$$

Similarly,

$$U_2 = K_2\left[\hat{X}_2 + L_2(Z_2-\hat{Z}_2)\right] = \frac{-ac}{c(a+1)-a}$$

$$\cdot\left[\frac{P}{P+R_2}Z_2 + \frac{aP[aPR_2 + (a+1)(P+R_2)cR_1]}{R_2[(P+R_1)(P+R_2)(a+1)(c-a) + a^2P^2]}\cdot\frac{R_2Z_2}{(P+R_2)}\right]$$

$$= \frac{-aP\sqrt{c}}{(P+R_2)[c(a+1)-a]}\left[1+\frac{a[aPR_2 + (a+1)(P+R_2)cR_1]}{[(P+R_1)(P+R_2)(a+1)(c-a) + a^2P^2]}\right]Z_2$$

which after some manipulation becomes

$$U_2 = \frac{-aP\sqrt{c}[(P+R_1)(a+1) - aP]}{[(P+R_1)(P+R_2)(a+1)(c-a) + a^2P^2]}Z_2$$

These answers are the same as those obtained by Willman when they are retransformed to the original problem.

We note that the problem was solved in three parts:

i) The feedback gain was derived.

ii) The conditional means and covariances were derived.

iii) The expressions for L_1 and L_2 were derived.

Of these steps, (i) is relatively straightforward and would be done in the course of solving the deterministic game. Furthermore, the procedure is not altered essentially when higher-dimensional multi-stage or continuous-time games are considered. Steps (ii) and (iii) are simplified immensely when one-stage discrete-time games are considered, because the problem of obtaining the conditional statistics is isolated from that of obtaining L_1 and L_2; i.e., steps (ii) and (iii) may be taken separately. In multi-stage or continuous-time games the covariance of the state depends on L_1 and L_2, and vice-versa. The result of this is that the conditional statistics and L_1 and L_2 must be obtained simultaneously.

No attempt to perform this computation will be made, since the ensuing analysis will show that no computationally feasible solution exists. In Chapter 3 the problem of computing conditional statistics is taken up under the simplifying assumption that $L_1 = L_2 = 0$. It is shown there that, even under this assumption, computation of the conditional mean of the state requires that each controller retain the entire past history of his observations. This data storage requirement is impractical; thus, in Chapter 4 a different approach is taken which requires that the strategies be optimized over a set of computationally feasible control functionals.

Chapter 3

PROBLEMS OF STATE ESTIMATION IN TWO-INPUT COOPERATIVE AND COMPETITIVE CONTROL SITUATIONS

3.1 Discrete-Time Case

To illustrate the various considerations affecting the problem of estimating the state of a linear system controlled by two or more inputs derived from independently made state observations, we begin with a discrete-time example.

Suppose we have a system described by the difference equation

$$X(i+1) = \phi(i+1,i)\, X(i) - G_1(i)\, U_1(i) + G_2(i)\, U_2(i) \qquad (3.1)$$

where $X(\cdot)$ is an n-vector; $E\left\{X(0)\right\} = \bar{X}_o$; Cov $\left\{X_o\right\} = \psi_{X_o}$ and $\phi(i+1,i)$ is a state transition matrix and thus satisfies relations such as

$$\begin{aligned} \phi(i,i) &= I \quad ; \quad I = \text{Identity Matrix} \\ \phi(i+1,i) &= F(i)\, \phi(i,i) \end{aligned} \qquad (3.2)$$

It was pointed out at the end of the previous chapter that the conditional statistics and the optimal L_1 and L_2 must be obtained simultaneously. Since here we are primarily interested in providing an expository development, we initially treat a simplified version of the problem: we shall assume that L_1 and L_2 are known by both players, so that we have only to deal with the state estimation problem. Furthermore, we shall assume that controller number 2 is restricted to $L_2 = 0$. Thus,

$$U_2(i) = K_2(i)\, \hat{X}_2(i) \qquad (3.3)$$

where

$$\hat{X}_2(i) = E\left\{X(i)|z_2(i)\right\} \tag{3.4}$$

$$z_2(i) = Z_2(0),\ Z_2(1),\ ---,\ Z_2(i) \tag{3.5}$$

$$Z_2(j) = H_2(j)\ X(j) + \eta_2(j) \quad ; \quad j = 0,1,\ ---,\ N \tag{3.6}$$

where $H_2(j)$ is an $m_2$xn matrix and η_2 is white, Gaussian, and

$$E\left\{\eta_2(i)\ \eta_2^*(j)\right\} = R_2(i)\ \delta_{ij} \tag{3.7}$$

and where δ_{ij} is the Kronecker delta. The problem then is to compute

$$\hat{X}_1(i) = E\left\{X(i)|z_1(i)\right\} \tag{3.8}$$

where

$$z_1(i) = Z_1(0),\ Z_1(1),\ ---,\ Z_1(i) \tag{3.9}$$

$$Z_1(j) = H_1(j)\ X(j) + \eta_1(j) \tag{3.10}$$

$$\eta_1 \text{ white, Gaussian, } E\left\{\eta_1(i)\ \eta_1^*(j)\right\} = R_1(i)\ \delta_{ij} \tag{3.11}$$

and η_1 and η_2 are independent

The following relations hold

$$\begin{aligned}\bar{X}_1(i+1) \stackrel{d}{=} E\left\{X(i+1)|z_1(i)\right\} &= \phi(i+1,i)\ \hat{X}_1(i)\\ &+ G_2(i)\ K_2(i)\ \hat{X}_{21}(i) - G_1(i)\ U_1(i)\end{aligned} \tag{3.12}$$

where

$$\hat{X}_{21}(i) = E\left\{\hat{X}_2(i) \mid z_1(i)\right\} \tag{3.13}$$

We know that $X(i+1)$ and $Z_1(i+1)$ are correlated Gaussian random vectors, that $X(i+1)$ has conditional mean $\bar{X}_1(i+1)$, and that $Z_1(i+1)$ has conditional mean $H_1(i+1)\,\bar{X}(i+1) \overset{d}{=} Z(i+1)$. Thus, by a well-known property of Gaussian random vectors [8, p.32], we may write

$$\hat{X}_1(i+1) = \bar{X}_1(i+1) + \Lambda(i+1)\left[Z_1(i+1) - \bar{Z}_1(i+1)\right] \tag{3.14}$$

where

$$\Lambda_1(i+1) = \Psi_{XZ_1}(i+1)\,\Psi_{Z_1Z_1}^{-1}(i+1) \tag{3.15}$$

where

$$\Psi_{XZ_1}(i+1) = E\left\{\left[X(i+1) - \bar{X}(i+1)\right]\left[Z_1(i+1) - \bar{Z}(i+1)\right]^*\right\} \tag{3.16}$$

and

$$\Psi_{Z_1Z_1}(i+1) = E\left\{\left[Z_1(i+1) - \bar{Z}_1(i+1)\right]\left[Z_1(i+1) - \bar{Z}_1(i+1)\right]^*\right\} \tag{3.17}$$

$\Lambda_1(i+1)$ is conventionally given in another form. If we define the error covariance matrix P_{11} by the equation

$$P_{11}(i+1) \overset{d}{=} E\left\{\left[X(i+1) - \hat{X}_1(i+1)\right]\left[X(i+1) - \hat{X}(i+1)\right]^*\right\}$$

$$= E\left\{\epsilon_1(i+1)\,\epsilon_1^*(i+1)\right\} \tag{3.18}$$

where

$$\epsilon_1(i+1) = \text{estimation error} \tag{3.19}$$

then since

$$\begin{aligned}\epsilon_1(i+1) &= X(i+1) - \hat{X}(i+1) \\ &= X(i+1) - \bar{X}(i+1) - L_1(i+1)\left[Z(i+1) - \bar{Z}(i+1)\right] \\ &= X(i+1) - \bar{X}(i+1) - L_1(i+1)\left[H_1(i+1)\right. \\ &\quad \left. \cdot \left(X(i+1) - \bar{X}(i+1)\right) + \eta_1(i+1)\right]\end{aligned} \tag{3.20}$$

and because $\eta_1(i+1)$ is independent of $X(i+1) - \bar{X}(i+1)$, we have

$$\begin{aligned}P_{11}(i+1) &= \psi_{XX}(i+1) - \psi_{XX}(i+1)\, H_1^*(i+1)\, \psi_{Z_1Z_1}^{-1}(i+1) \\ &\quad \cdot H_1(i+1)\, \psi_{XX}(i+1)\end{aligned} \tag{3.21}$$

Furthermore, since $Z_1(i+1) = H_1(i+1)\, X(i+1) + \eta_1(i+1)$, we have

$$\psi_{Z_1Z_1}(i+1) = H_1(i+1)\, \psi_{XX}(i+1)\, H_1(i+1)^* + R_1(i+1) \tag{3.22}$$

We may thus write

$$\begin{aligned}R_1(i+1) &= \psi_{Z_1Z_1}(i+1) - H_1(i+1)\, \psi_{XX}(i+1)\, H_1^*(i+1) \\ &= \psi_{Z_1Z_1}(i+1)\left[I - \psi_{Z_1Z_1}^{-1}(i+1)\, H_1(i+1)\, \psi_{XX}(i+1)\right. \\ &\quad \left. \cdot H_1^*(i+1)\right]\end{aligned} \tag{3.23}$$

So

$$\psi_{Z_1Z_1}^{-1}(i+1) = \left[I - \psi_{Z_1Z_1}^{-1}(i+1)\ H_1(i+1)\ \psi_{XX}(i+1)\ H_1^*(i+1)\right] \cdot R_1(i+1)^{-1} \tag{3.24}$$

and

$$\begin{aligned}\Lambda_1(i+1) &= \psi_{XZ_1}(i+1)\ \psi_{Z_1Z_1}^{-1}(i+1) \\ &= \psi_{XX}(i+1)\ H_1^*(i+1)\Big[I - \psi_{Z_1Z_1}^{-1}(i+1)\ H_1(i+1) \\ &\quad \cdot \psi_{XX}(i+1)\ H_1^*(i+1)\Big]R_1^{-1}(i+1) \\ &= \Big[\psi_{XX}(i+1) - \psi_{XX}(i+1)\ H_1^*(i+1)\ \psi_{Z_1Z_1}^{-1}(i+1) \\ &\quad \cdot H_1(i+1)\ \psi_{XX}(i+1)\Big]H_1^*(i+1)\ R_1^{-1}(i+1) \\ &= P_{11}(i+1)\ H_1^*(i+1)\ R_1^{-1}(i+1) \end{aligned} \tag{3.25}$$

Combining (3.12), (3.14), and (3.25), we have

$$\begin{aligned}\hat{X}_1(i+1) &= \phi(i+1,i)\ \hat{X}_1(i) + P_{11}(i+1)\ H_1^*(i+1)\ R_1^{-1}(i+1) \\ &\quad \cdot \left[Z_1(i+1) - H_1(i+1)\ \bar{X}_1(i+1)\right] + G_2(i)\ K_2(i)\ \hat{X}_{21}(i) \\ &\quad - G_1(i)\ U_1(i)\end{aligned} \tag{3.26}$$

This equation involves the quantity $\hat{X}_{21}(i)$. To compute $\hat{X}_{21}(i)$, let us initially assume that controller number 2 uses a state estimate which has the form

$$\hat{X}_2(i) = \theta(i,0)\ \bar{X}(0) + \sum_{j=0}^{i} W(i,j)\ Z_2(j)\ . \tag{3.27}$$

with no restrictions for the moment on the matrices $\theta(i,0)$ or $W(i,j)$. However, we do assume that these matrices are known to controller number 1. Since $\bar{X}(0)$ is known to both players, we may write

$$\hat{X}_{21}(i) = \theta(i,0)\ \bar{X}(0) + \sum_{j=0}^{i} W(i,j)\ E\left\{Z_2(j)|z_1(i)\right\} \tag{3.28}$$

and because $Z_2(j)$ and $Z_1(i)$ are correlated Gaussian random variables, we may write

$$E\left\{Z_2(j)|z_1(i)\right\} = E\left\{Z_2(j)|z_1(i-1)\right\} + M(j,i)\left[Z_1(i) - \bar{Z}_1(i)\right] \tag{3.29}$$

where

$$M(j,i) = \Phi_{Z_2Z_1}(j,i-1)\ \Phi_{Z_1Z_1}^{-1}(i) \tag{3.30}$$

$$\Phi_{Z_2Z_1}(j,i-1) = E\left\{\left[Z_2(j) - E\left\{Z_2(j)|z_1(i-1)\right\}\right]\left[Z_1(i) - \bar{Z}_1(i)\right]^*\right\} \tag{3.31}$$

$$\Phi_{Z_1Z_1}(i) = E\left\{\left[Z_1(i) - \bar{Z}_1(i)\right]\left[Z_1(i) - \bar{Z}_1(i)\right]^*\right\} \tag{3.32}$$

$$\bar{Z}_1(i) = H_1(i)\ \bar{X}_1(i) \tag{3.33}$$

Notice that (3.29) is a difference equation whose solution may be written

$$E\left\{Z_2(j)|z_1(i)\right\} = E\left\{Z_2(j)\right\} + \sum_{k=0}^{i} M(j,k)\left[Z_1(k) - Z_1(k)\right] \tag{3.34}$$

Thus, (3.28) may be written

$$\hat{X}_{21}(i) = \theta(i,0)\,\bar{X}(0) + \sum_{j=0}^{i} W(i,j)\,E\left\{Z_2(j)\right\}$$

$$+ \sum_{j=0}^{i} W(i,j) \sum_{k=0}^{i} M(j,k)\left[Z_1(k) - \bar{Z}_1(k)\right] \tag{3.35}$$

Let us assume that $E\left\{Z_2(j)\right\} = D(j)\,\bar{X}(0)$ (3.36)

Then, defining $T(i,0) = \theta(i,0) + \sum_{j=0}^{i} W(i,j)\,D(j)$, we may write (3.37)

$$\hat{X}_{21}(i) = T(i,0)\,X(0) + \sum_{j=0}^{i} W(i,j) \sum_{k=0}^{i} M(j,k)\left[Z_1(k) - \bar{Z}_1(k)\right] \tag{3.38}$$

We observe at this point that in order to calculate $\hat{X}_1(i+1)$ one must know $\hat{X}_{21}(i)$, which in turn requires the preservation of the observations $Z_1(k), k = 0,1,---,i$.

3.2 Continuous-Time Case

We are interested mainly in the continuous-time version of the equations so far derived. The continuous-time equations are obtained by the familiar process of writing $\phi(i+1,i)$ as $\phi(t+\Delta,t)$ and expanding $\phi(t+\Delta,t)$ in a Taylor series as

$$\phi(t+\Delta,t) = \phi(t,t) + F(t)\Delta + O(\Delta^2) = I + F(t)\Delta + O(\Delta^2) \qquad (3.39)$$

We also modify the forcing terms in (3.1), so that this equation becomes

$$X(t+\Delta) = \left[I + F(t)\Delta + O(\Delta)^2\right] X(t) - G_1(t)\, U_1(t)\Delta + G_2(t)U_2(t)\Delta \qquad (3.40)$$

After we have subtracted X(t) from both sides, divided both sides by Δ, and taken the limit as Δ approaches zero, (3.40) becomes

$$\dot{X}(t) = F(t)\, X(t) - G_1(t)\, U_1(t) + G_2(t)\, U_2(t) \qquad (3.41)$$

By a similar procedure, (3.12) becomes, upon substituting (3.39) and modifying the forcing terms,

$$\bar{X}_1(t+\Delta) = \left[I + F(t)\Delta + O(\Delta^2)\right] \hat{X}_1(t) + G_2(t)\, K_2(g)\, \hat{X}_{21}(t)\Delta - G_1(t)U_1(t)\Delta \qquad (3.42)$$

Letting Δ approach zero, we see that $\bar{X}(t) \to \hat{X}(t)$. Likewise, (3.26) may be written

$$\hat{X}_1(t+\Delta) = \left[I + F(t)\Delta + O(\Delta^2)\right] \hat{X}_1(t) + P_{11}(t+\Delta)\, H_1^*(t+\Delta)\, R_1^{-1}(t+\Delta) \left[Z_1(t+\Delta) - H_1(t+\Delta)\, \bar{X}_1(t+\Delta)\right] \Delta + G_2(t)\, K_2(t)\, \hat{X}_{21}(t)\Delta - G_1(t)\, U_1(t)\Delta \qquad (3.43)$$

Subtracting $\hat{X}(t)$ from both sides, dividing both sides by Δ, taking limits as Δ approaches zero, and using (3.43), we have

$$\dot{\hat{X}}_1(t) = F(t)\,\hat{X}_1(t) + F_{11}(t)\,H_1^*(t)\,R_1^{-1}(t)\left[Z_1(t) - H_1(t)\,\hat{X}_1(t)\right]$$

$$+ G_2(t)\,K_2(t)\,\hat{X}_{21}(t) - G_1(t)\,U_1(t) \tag{3.44}$$

Here the spectral properties of $\eta_1(t)$ must be modified so that over a unit time interval the additive noise has the same corruptive influence as in the discrete-time case. Specifically, $\eta_1(t)$ is taken to be a white noise process with spectral density

$$E\left\{\eta_1(t)\,\eta_1^*(\tau)\right\} = R_1(t)\,\delta(t-\tau) \tag{3.45}$$

where $\delta(t-\tau)$ is a Dirac delta function.

Using these same techniques, (3.29) becomes

$$E\left\{Z_2(\tau_j)\,|\,z_1(t)\right\} = E\left\{Z_2(\tau_j)\,|\,z_1(t-\Delta)\right\} + M(\tau_j,t)\left[Z_1(t) - \bar{Z}_1(t)\right]\Delta \tag{3.46}$$

Subtracting $E\left\{Z_2(t)\,|\,z_1(t-\Delta)\right\}$ from both sides, dividing by Δ, and taking limits as Δ approaches zero, we have

$$\frac{d}{dt}E\left\{Z_2(t)\,|\,z_1(t)\right\} = M(t,t)\left[Z_1(t) - \bar{Z}_1(t)\right] \tag{3.47}$$

Equation (3.47) has solution

$$E\left\{Z_2(t)\,|\,z_1(t)\right\} = E\left\{Z_2(t)\right\} + \int_0^t M(t,\sigma)\left[Z_1(\sigma) - \bar{Z}_1(\sigma)\right]$$

$$= D(t)\bar{X}(0) + \int_0^t M(t,\sigma)\left[Z_1(\sigma) - Z_1(\sigma)\right]d\sigma \tag{3.48}$$

The similarity to (3.34) is obvious.

If we assume that controller number 2 uses a state estimate of the form

$$\hat{X}_2(t) = \theta(t,0)\,\bar{X}(0) + \int_0^t W(t,\tau)\, Z_2(\tau)d\tau \tag{3.49}$$

then the continuous-time analog of (3.38) is

$$\hat{X}_{21}(t) = T(t,0)\,\bar{X}(0) + \int_0^t W(t,\tau) \int_0^\tau M(\tau,\sigma)$$

$$\cdot \left[Z_1(\sigma) - \bar{Z}_1(\sigma)\right] d\sigma d\tau \tag{3.50}$$

Thus, calculation of $\hat{X}_{21}(t)$ appears to require storage of $Z_1(\sigma), 0 \leq \sigma \leq t$. We now prove this to be true; i.e., (3.50) can be obtained in no simpler form.

To this point we have made no restrictive assumptions about $W(t,\tau)$ or $\theta(t,0)$. We shall now do so, showing that in order for each controller to compute $E\{X(t)|Z_k(t)\}, k = 1,2$, he must store all past observations.

We first assume that $W(t,\tau)$ is of the form

$$W(t,\tau) = C(t)\, Q(t,\tau)\, N(t) \tag{3.51}$$

where $Q(t,\tau)$ is a p x p matrix which satisfies

$$Q(t,\tau) = I \tag{3.52}$$

$$\frac{d}{dt}' Q(t,\tau) = \Gamma(t)\, Q(t,\tau) \tag{3.53}$$

and $N(t)$ is a $p \times m_2$ matrix. We also assume that $\theta(t,0)$ is of the form

$$\theta(t,0) = C(t)\, Q(t,0)\, \gamma(0) \tag{3.54}$$

where $C(t)$ is a differentiable $n \times p$ matrix. These assumptions are equivalent to requiring that $\hat{X}_2(t)$ be given by

$$\hat{X}_2(t) = C(t)\, q(t) \tag{3.55}$$

where $q(t)$ satisfies the pth order differential equation

$$\frac{d}{dt} q(t) = \Gamma(t)\, q(t) + N(t)\, Z_2(t) \tag{3.56}$$

$$q(0) = \gamma(0)\, \bar{X}(0) \tag{3.57}$$

We call $\hat{X}_2(t)$ a "p dimensional state estimator." Under these assumptions, we see that $\hat{X}_{21}(t)$, given by (3.50), can be written

$$\hat{X}_{21} = \left[C(t)\, Q(t,0)\, \gamma(0) + \int_0^t C(t)\, Q(t,\tau)\, N(\tau)\, D(\tau) d\tau\right] \bar{X}(0) + \int_0^t C(t)\, Q(t,\tau) \int_0^t M(\tau,\sigma) \left[Z_1(\sigma) - \bar{Z}_1(\sigma)\right] d\sigma\, d\tau \tag{3.58}$$

Defining a new variable $\hat{q}(t)$ by

$$\hat{q}(t) = \left[Q(t,0)\, \gamma(0) + \int_0^t Q(t,\tau)\, N(\tau)\, D(\tau)\right] \bar{X}(0) + \int_0^t Q(t,\tau)\, N(\tau) \int_0^t M(\tau,\sigma) \left[Z_1(\sigma) - Z_1(\sigma)\right] d\sigma\, d\tau \tag{3.59}$$

we see that $\hat{X}_{21}(t) = C(t)\,\hat{q}(t)$ (3.60)

and that $\hat{X}_{21}(t)$ satisfies the integro-differential equation

$$\frac{d}{dt}\hat{X}_{21}(t) = \left[\frac{d}{dt}(t)\right]\hat{q}(t) + C(t)\,\frac{d}{dt}\hat{q}(t) \tag{3.61}$$

$$\begin{aligned}\frac{d}{dt}\hat{q}(t) &= \Gamma(t)\,\hat{q}(t) + N(t)\,D(t)\,\bar{X}(0) \\ &\quad + N(t)\int_0^t M(t,\sigma)\left[Z_1(\sigma) - \bar{Z}_1(\sigma)\right] d\sigma \\ &\quad + \int_0^t Q(t,\tau)\,N(\tau)\,M(\tau,t)\left[Z_1(t) - \bar{Z}_1(t)\right] d\tau\end{aligned} \tag{3.62}$$

At this point we define a new matrix

$$\pi(t) \overset{d}{=} \int_0^t Q(t,\tau)\,N(\tau)\,M(\tau,t)\,d\tau \tag{3.63}$$

Also, we note that because of (3.48),

$$\int_0^t M(t,\sigma)\left[Z_1(\sigma) - \bar{Z}_1(\sigma)\right] d\sigma + D(t)\,\bar{X}(0) = E\left\{Z_2(t)\,|\,z_1(t)\right\} \tag{3.64}$$

and because of the independence of $\eta_2(t)$ and $Z_1(t)$,

$$\begin{aligned}E\left\{Z_2(t)\,|\,Z_1(\tau),\ 0 \le \tau \le t\right\} &= H_2(t)\,E\left\{X(t)\,|\,Z_1(\tau),\ 0 \le \tau \le t\right\} \\ &= H_2(t)\,\hat{X}_1(t)\end{aligned} \tag{3.65}$$

We may thus write (3.56) as

$$\frac{d}{dt}\hat{q}(t) = \Gamma(t)\,\hat{q}(t) + N(t)\,H_2(t)\,\hat{X}_1(t) + \pi(t)\left[Z_1(t) - \bar{Z}_1(t)\right] \tag{3.66}$$

and (3.61) becomes

$$\frac{d}{dt}\hat{X}_{21}(t) = \left[\frac{d}{dt}C(t) + C(t)\,\Gamma(t)\right]\hat{q}(t) + N(t)\,H_2(t)\,\hat{X}_1(t) + \pi(t)\left[Z_1(t) - \bar{Z}_1(t)\right] \tag{3.67}$$

Repeating (3.44), we have

$$\frac{d}{dt}\hat{X}_1(t) = F(t)\,\hat{X}_1(t) + P_{11}(t)\,H_1^*(t)\,R_1^{-1}(t) \cdot \left[Z_1(t) - H_1(t)\,\hat{X}_1(t)\right] + G_2(t)\,K_2(t)\,\hat{X}_{21}(t) - G_1(t)\,U_1(t) \tag{3.68}$$

Equations (3.66), (3.67), and (3.68) taken together constitute a system of n+p first order differential equations whose solution gives the state estimate $\hat{X}_1(t)$. This result is intuitively reasonable: if controller number 2 is constrained to use a "p dimensional" state estimator, then controller number 1 must use an "n+p dimensional" state estimator. Furthermore, because of the restrictive assumptions we have made, we are actually able to solve the game problem, i.e., obtain L_1. This is done as follows: since we have assumed $L_2 = 0$, we may write $U_2 = K_2(t)\,\hat{X}_2$; using (3.55), this may be written $U_2 = K_2(t)\,C(t)\,q(t)$;

and so (3.41) becomes

$$\frac{d}{dt} X = F(t)\, X(t) + G_2(t)\, K_2(t)\, C(t)\, q(t) - G_1(t)\, U_1(t) \tag{3.69}$$

These equations may be written

$$\begin{bmatrix} \dot{x} \\ \dot{q} \end{bmatrix} = \begin{bmatrix} F & G_2K_2C \\ NH_2 & \Gamma \end{bmatrix} \begin{bmatrix} X \\ q \end{bmatrix} + \begin{bmatrix} 0 \\ N \end{bmatrix} \eta_2 + \begin{bmatrix} -G_1 \\ 0 \end{bmatrix} U_1 \tag{3.70}$$

3.3 Control Applications of the State Estimation Procedure

The original criterion functional may be written in terms of this augmented system as

$$J = E\left\{ \begin{bmatrix} X(T) \\ q(T) \end{bmatrix}^* \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X(T) \\ q(T) \end{bmatrix} + \int_0^T U_1^*(t) U_1(t) dt - \int_0^T \begin{bmatrix} X(t) \\ q(t) \end{bmatrix}^* \begin{bmatrix} 0 & 0 \\ 0 & C^* K_2^* K_2 C \end{bmatrix} \begin{bmatrix} X(t) \\ q(t) \end{bmatrix} dt \right\} \tag{3.71}$$

This is a classical one-sided stochastic optimal control problem of the linear-quadratic type, and the solution is well known to be of the form

$$U_1 = K \begin{bmatrix} \hat{X}_1 \\ \hat{q} \end{bmatrix} \tag{3.72}$$

We have already observed that $\hat{X}_1$ satisfies (3.44), which may be combined with (3.60) to read

$$\hat{X}_1' = F(t)\hat{X}_1(t) + P_{11}(t) H_1^*(t) R_1^{-1} \left[Z_1(t) - H_1(t)\, \hat{X}_1(t) \right] + G_2(t)\, K_2(t)\, C(t)\, \hat{q}(t) - G_1(t)\, U_1(t) \tag{3.73}$$

Furthermore, $\hat{q}$ is given as the solution of (3.66), which is

$$\dot{\hat{q}} = \Gamma(t)\hat{q}(t) + N(t)\hat{X}_1 + \pi(t)\left[Z_1(t) - H_1\hat{X}_1(t)\right] \tag{3.74}$$

Behn and Ho [3] have solved this problem for the case in which $\eta_1 = 0$, and their result is that

$$U_1 = K\begin{bmatrix} X \\ \epsilon_2 \end{bmatrix} \tag{3.75}$$

where K may be written $K = \left[K_1 : D_p\right]$ and K_1 is the deterministic optimal feedback gain derived in Chapter 1. Since we may write $\epsilon_2 = X - Cq$, Behn and Ho's solution may also be written

$$U_1 = \left[K_1 + D_p \,\vdots\, -D_pC\right]\begin{bmatrix} X \\ q \end{bmatrix} \tag{3.76}$$

Then if we fix controller number 2's strategy, i.e., require that he continue to play as if $\eta_1 = 0$, the problem is simply a stochastic optimal control problem. We may apply the separation principle to obtain

$$U_1 = \left[K_1 + D_p \,\vdots\, -D_pC\right]\begin{bmatrix} \hat{X}_1 \\ \hat{q} \end{bmatrix} \tag{3.77}$$

Thus, for this special case we have solved the game problem. The result may be written

$$\begin{aligned} U_1 &= K_1\hat{X}_1 + D_p\hat{X}_1 - D_p C\hat{q} \\ &= K_1\left[\hat{X}_1 + L_1(Z_1 - \hat{Z}_1)\right] \end{aligned} \tag{3.78}$$

Therefore, L_1 satisfies

$$K_1 L_1 (Z_1 - \hat{Z}_1) = D_P \left[\hat{X}_1 - C\hat{q} \right] \tag{3.79}$$

From a computational standpoint, such a requirement is unreasonable; thus, a game strategy which incorporates the conditional mean of the state, the conditioning being done on all past observations, is not satisfactory from an engineering viewpoint unless the opposing strategy is known to be dimensionally restricted. If the opposing strategy is in fact dimensionally restricted, the resulting game situation is unsymmetrical.

3.4 A Suboptimal Estimation Procedure

An interesting suboptimal state-estimation procedure has been developed by Rhodes and Luenberger [24], the significance of which will be shown in Chapter 4. The method uses state estimates generated by differential equations which are of the same order as the controlled system. This procedure is a compromise between estimation error and computational difficulty. We have already developed a differential equation (3.44) describing the conditional mean $\hat{X}_1$ of the state:

$$\begin{aligned}\dot{\hat{X}}_1 &= F(t)\,\hat{X}_1(t) + P_{11}(t)\,H_1^*(t)\,R_1^{-1}(t) \\ &\quad \cdot \left[Z_1(t) - H_1(t)\,\hat{X}_1(t) \right] + G_2\,K_2\,\hat{X}_{21}(t) \\ &\quad - G_1(t)\,U_1(t)\end{aligned} \tag{3.80}$$

We recall that the problem of dimensionality enters the picture in the calculation of $\hat{X}_{21}(t)$. As a simplifying assumption, let us take $\hat{X}_2(t)$ to be approximated by some linear transformation of $\hat{X}_1(t)$; i.e., let

$$\hat{X}_{21}(t) = \Omega_1(t)\,\hat{X}_1(t) \tag{3.81}$$

where $\Omega_1(t)$ is to be chosen according to some criterion of optimality. For reasons which we shall see later, a desirable criterion is mean square error; i.e., we choose Ω_1 to minimize

$$\begin{aligned}
\frac{1}{2}\,\mathrm{tr}\left[\mathrm{Cov}\,(\hat{X}_2 - \Omega_1\hat{X}_1)\right] &= \frac{1}{2}\,\mathrm{tr}\left[E\left\{\left[\hat{X}_2 - \Omega_1\hat{X}_1\right]\left[\hat{X}_2 - \Omega_1\hat{X}_1\right]^*\right\}\right] \\
&= \frac{1}{2}\,\mathrm{tr}\left[E\left\{X_2X_2^* - \Omega_1X_1X_2^* - X_2X_1^*\Omega_1^* + \Omega X_1X_1^*\Omega_1^*\right\}\right] \\
&= \frac{1}{2}\,\mathrm{tr}\left[E\left\{X_2X_2^* - 2X_2X_1^*\Omega_1^* + \Omega_1X_1X_1^*\Omega_1^*\right\}\right]
\end{aligned} \tag{3.82}$$

Taking the gradient with respect to Ω_1 and setting the resulting expression equal to zero, we have

$$E\left\{-\hat{X}_2\hat{X}_1^* + \Omega_1\hat{X}_1\hat{X}_1^*\right\} = 0 \tag{3.83}$$

or

$$\Omega_1 = E\left\{\hat{X}_2\hat{X}_1^*\right\}\left[E\left\{\hat{X}_1\hat{X}_1^*\right\}\right]^{-1} \tag{3.84}$$

Now, if $\epsilon_1 = X - \hat{X}_1$ and $\epsilon_2 = X - \hat{X}_2$ and we define the vector ρ by

$$\rho = \begin{bmatrix} X \\ \epsilon_1 \\ \epsilon_2 \end{bmatrix} \tag{3.85}$$

and let

$$P = \text{Cov}\,\rho = E\left\{\rho\rho^*\right\} = \begin{bmatrix} P_{00} & P_{01} & P_{02} \\ P_{10} & P_{11} & P_{12} \\ P_{20} & P_{21} & P_{22} \end{bmatrix}. \tag{3.86}$$

then

$$\hat{X}_2\hat{X}_1^* = (X-\epsilon_2)(X-\epsilon_1)^* \tag{3.87}$$

where Cq is controller number 2's p-dimensional state estimator, which is designed with the assumption that $\eta_1 = 0$. Behn and Ho have shown that under this assumption p = n and C = I; i.e., controller number 2 needs only an n-dimensional state estimate, in this case generated by a Kalman filter.

This problem is not, however, a true game problem, since all of the parameters of controller number 2's strategy are fixed. But, since the purpose of this chapter is to analyze the problem of state estimation, with game theoretic considerations suppressed temporarily, we proceed in that vein.

Because controller number 2's p(=n) dimensional state estimate is based on erroneous assumptions, it is not certain how good a state estimate it is. Even within the class of n-dimensional estimators, it may not be optimal either as an estimator or as a strategic variable. Clearly, from controller number 2's viewpoint, the "p-dimensional" state estimate is inferior to a "2n + p-dimensional" estimator, which he would use were he not constrained. Inductively, we conclude that no state estimates generated by finite-ordered differential equations can make optimum use of all of the information contained in the observations.

The reason for the difficulty encountered in making the state estimates is that the "state" of the system includes the "state" of each controller's estimate. When a system is described by a differential equation, then its state estimate is also described by a differential equation of the same order. When a system is described by a integral expression, its state estimate is also an integral expression; and to compute this integral, all past values of observations must be retained. So

$$E\left\{\hat{X}_2\hat{X}_1^*\right\} = P_{11} - P_{01} - P_{20} + P_{21} \tag{3.88}$$

and

$$E\left\{\hat{X}_1\hat{X}_1^*\right\} = E\left\{(X-\epsilon_1)\right\}\hat{X}_1^* = E\left\{X\,\hat{X}_1^*\right\} - E\left\{\epsilon_1\hat{X}_1^*\right\} \tag{3.89}$$

It is a property of optimal estimates [8, pp 38-43] that $E\left\{\epsilon_1\hat{X}_1^*\right\} = 0$. For the moment, we shall assume this to be true for our estimate also and verify the fact later. Thus, (3.89) may be written

$$E\left\{\hat{X}_1\hat{X}_1^*\right\} = E\left\{X\,\hat{X}_1^*\right\} = E\left\{X(X-\epsilon_1)^*\right\} = P_{00} - P_{01} \tag{3.90}$$

Using (3.88) and (3.90), (3.84) may be written

$$\begin{aligned} \Omega_1 &= \left[P_{00} - P_{01} - P_{20} + P_{21}\right]\left[P_{00} - P_{01}\right]^{-1} \\ &= I - \left[P_{20} - P_{21}\right]\left[P_{00} - P_{01}\right]^{-1} \end{aligned} \tag{3.91}$$

Substituting this expression into (3.80), we have

$$\dot{\hat{X}}_1 = F(t)\,\hat{X}_1 + P_{11}(t)\,H_1^*(t)\,R_1^{-1}(t)\left[Z_1(t) - H_1(t)\,\hat{X}_1(t)\right]$$

$$+ G_2(t)\,K_2(t)\left[I - \left[P_{20} - P_{21}\right]\left[P_{00} - P_{01}\right]^{-1}\right]\hat{X}_1(t)$$

$$- G_1(t)\,U_1(t) \tag{3.92}$$

or

$$\dot{\hat{X}}_1 = \left[F(t) + G_2(t)\,K_2(t)\left[I - (P_{20} - P_{21})\,(P_{00} - P_{01})^{-1}\right]\right]\hat{X}_1(t)$$

$$+ P_{11}(t)\,H_1^*(t)\,R_1^{-1}(t)\left[Z_1(t) - H_1(t)\,\hat{X}_1(t)\right]$$

$$- G_1(t)\,U_1(t) \tag{3.93}$$

We have thus derived an "n-dimensional" state estimator for controller number 1. This estimator is given in terms of the covariances P_{20}, P_{21}, P_{00}, P_{01}, and P_{11}, quantities which must be calculated separately. Note that as $\eta_2 \to 0$, $\Omega \to I$.

It is impossible to calculate these covariances, however, unless we have some knowledge of the form of $\hat{X}_2(t)$. We therefore assume that $\hat{X}_2(t)$ is an n-dimensional estimator of the same form as $\hat{X}_1(t)$ and is thus described by the differential equation

$$\frac{d}{dt}\hat{X}_2(t) = F(t)\,\hat{X}_2(t) + P_{22}(t)\,H_2^*(t)\,R_2^{-1}(t)$$

$$\left[Z_2(t) - H_2(t)\,\hat{X}_2(t)\right] - G_1(t)\,K_1(t)\,\hat{X}_{12}(t)$$

$$+ G_2(t)\,U_2(t) \tag{3.94}$$

Again, we approximate $\hat{X}_1(t)$ by $\Omega_2(t)\ \hat{X}_2(t)$ and by an analogous manipulation obtain

$$\Omega_2(t) = I - (P_{10} - P_{12})(P_{00} - P_{02})^{-1} \tag{3.95}$$

Thus, (3.94) becomes

$$\frac{d}{dt}\hat{X}_2(t) = \left[F(t) - G_1(t)K_1(t)\left[I - (P_{10} - P_{12})(P_{00} - P_{02})^{-1}\right]\right]\hat{X}_2(t)$$
$$+ P_{22}(t)H_2^*(t)R_2^{-1}(t)\left[Z_2(t) - H_2(t)\hat{X}_2(t)\right] + G_2(t)\ U_2(t) \tag{3.96}$$

We are now in a position to calculate the covariances. We begin with system equation

$$\frac{d}{dt}X(t) = F(t)\ X(t) - G_1(t)K_1(t)\hat{X}_1(t) + G_2(t)K_2(t)\hat{X}_2(t) \tag{3.97}$$

This may be rewritten, dropping the "t" argument, as

$$\dot{X} = \left[F - G_1K_1 + G_2K_2\right] X + G_1K_1\epsilon_1 - G_2K_2\epsilon_2 \tag{3.98}$$

We may express $\dot{\epsilon}_1 = \dot{X} - \dot{\hat{X}}_1$ by

$$\dot{\epsilon}_1 = +G_2K_2(P_{20} - P_{21})(P_{00} - P_{01})^{-1}\ X$$
$$+ \left[F + G_2K_2\left[I - (P_{20} - P_{21})(P_{00} - P_{01})^{-1}\right] - P_{11}H_1^*R_1^{-1}H_1\right]\epsilon_1$$
$$- G_2K_2\epsilon_2 - P_{11}H_1^*R_1^{-1}\eta_1 \tag{3.99}$$

Similarly,

$$\dot{\epsilon}_2 = -G_1K_1(P_{10}-P_{12})(P_{00}-P_{02})^{-1}X + G_1K_1\epsilon_1$$
$$+\Big[F - G_1K_1\big[I - (P_{10}-P_{12})(P_{00}-P_{02})^{-1}\big]$$
$$- P_{22}H_2^*R_2^{-1}H_2\Big]\epsilon_2 - P_{22}H_2^*R_2^{-1}\eta_2 \qquad (3.100)$$

We may write a differential equation describing the vector $\rho = \begin{bmatrix} X \\ \epsilon_1 \\ \epsilon_2 \end{bmatrix}$ as follows:

$$\dot{\rho} = \Gamma\rho + B\eta \qquad (3.101)$$

where

$$\eta = \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \qquad (3.102)$$

$$B = \begin{bmatrix} 0 & 0 \\ -P_{11}H_1^*R_1^{-1} & 0 \\ 0 & -P_{22}H_2^*R_2^{-1} \end{bmatrix} \qquad (3.103)$$

$$\Gamma = \left[\begin{array}{c|c|c} F-G_1K_1+G_2K_2 & G_1K_1 & -G_2K_2 \\ \hline +G_2K_2(P_{20}-P_{21}) & F+G_2K_2\big[I-(P_{20}-P_{21}) & -G_2K_2 \\ \cdot(P_{00}-P_{01})^{-1} & \cdot(P_{00}-P_{01})^{-1}\big] & \\ & -P_{11}H_1^*R_1^{-1}H_1 & \\ \hline -G_1K_1(P_{10}-P_{12}) & G_1K_1 & F-G_1K_1\big[I-(P_{10}-P_{12}) \\ \cdot(P_{00}-P_{02})^{-1} & & \cdot(P_{00}-P_{02})^{-1}\big] \\ & & -P_{22}H_2^*R_2^{-1}H_2 \end{array}\right]$$

(3.104)

Then $P = \mathrm{Cov}\ (\rho)$ satisfies the differential equation

$$\dot{P} = \Gamma P + P\Gamma^* + BRB^* \tag{3.105}$$

where

$$R = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix} \tag{3.106}$$

Denoting the submatrices of Γ and B by

$$\Gamma = \begin{bmatrix} \Gamma_{00} & \Gamma_{01} & \Gamma_{02} \\ \Gamma_{10} & \Gamma_{11} & \Gamma_{12} \\ \Gamma_{20} & \Gamma_{21} & \Gamma_{22} \end{bmatrix} \tag{3.107}$$

$$B = \begin{bmatrix} 0 & 0 \\ -B_1 & 0 \\ 0 & -B_2 \end{bmatrix} \tag{3.108}$$

and expanding (3.105), we have

$$\begin{bmatrix} \dot{P}_{00} & \dot{P}_{01} & \dot{P}_{02} \\ \dot{P}_{10} & \dot{P}_{11} & \dot{P}_{12} \\ \dot{P}_{20} & \dot{P}_{21} & \dot{P}_{22} \end{bmatrix} = \begin{bmatrix} \Gamma_{00} & \Gamma_{01} & \Gamma_{02} \\ \Gamma_{10} & \Gamma_{11} & \Gamma_{12} \\ \Gamma_{20} & \Gamma_{21} & \Gamma_{22} \end{bmatrix} \begin{bmatrix} P_{00} & P_{01} & P_{02} \\ P_{10} & P_{11} & P_{12} \\ P_{20} & P_{21} & P_{22} \end{bmatrix}$$
$$+ \begin{bmatrix} P_{00} & P_{01} & P_{02} \\ P_{10} & P_{11} & P_{12} \\ P_{20} & P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} \Gamma_{00}^* & \Gamma_{10}^* & \Gamma_{20}^* \\ \Gamma_{01}^* & \Gamma_{11}^* & \Gamma_{21}^* \\ \Gamma_{02}^* & \Gamma_{12}^* & \Gamma_{22}^* \end{bmatrix}$$
$$+ \begin{bmatrix} 0 & 0 \\ -B_1 & 0 \\ 0 & -B_2 \end{bmatrix} \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix} \begin{bmatrix} 0 & -B_1^* & 0 \\ 0 & 0 & -B_2^* \end{bmatrix} \tag{3.109}$$

Sorting out the expressions for $\dot{P}_{11}$ and $\dot{P}_{10}$, we have

$$\dot{P}_{11} = \Gamma_{10}P_{01} + \Gamma_{11}P_{11} + \Gamma_{12}P_{21} + P_{10}\Gamma_{10}^{*} + P_{11}\Gamma_{11}^{*} + P_{12}\Gamma_{12}^{*} + B_1R_1B_1^{*} \tag{3.110}$$

$$\dot{P}_{10} = \Gamma_{10}P_{00} + \Gamma_{11}P_{10} + \Gamma_{12}P_{20} + P_{10}\Gamma_{00}^{*} + P_{11}\Gamma_{01}^{*} + P_{12}\Gamma_{02}^{*} \tag{3.111}$$

Then, subtracting (3.111) from (3.110), we have

$$\begin{aligned}\dot{P}_{11} - \dot{P}_{10} = {} & \Gamma_{10}(P_{01}-P_{00}) + \Gamma_{11}(P_{11}-P_{10}) + \Gamma_{12}(P_{21}-P_{20}) \\ & + B_1R_1B_1^{*} + P_{10}(\Gamma_{10}-\Gamma_{00})^{*} + P_{11}(\Gamma_{11}-\Gamma_{01})^{*} \\ & + P_{12}(\Gamma_{12}-\Gamma_{02})^{*}\end{aligned} \tag{3.112}$$

Now $\Gamma_{12} = -G_2K_2$ and $\Gamma_{02} = -G_2K_2$, so the last term of (3.112) may be dropped. Also,

$$\Gamma_{10} - \Gamma_{00} = +G_2K_2(P_{20}-P_{21})(P_{00}-P_{01})^{-1} - F + G_1K_1 - G_2K_2 \tag{3.113}$$

and

$$\begin{aligned}\Gamma_{11} - \Gamma_{01} = {} & F - G_2K_2(P_{20}-P_{21})(P_{00}-P_{01})^{-1} - P_{11}H_1^{*}R_1^{-1}H_1 \\ & + G_2K_2 - G_1K_1 = -(\Gamma_{10}-\Gamma_{00}) - B_1H_1\end{aligned} \tag{3.114}$$

So the fifth and sixth terms of the right side of (3.112) may be written

$$(P_{11}-P_{10})(\Gamma_{10}-\Gamma_{00})^{*} - P_{11}H_1^{*}R_1^{-1}H_1P_{11} \tag{3.115}$$

Finally, note that

$$\Gamma_{10}(P_{01}-P_{00}) = -G_2K_2(P_{20}-P_{21})) = -\Gamma_{12}(P_{21}-P_{20}) \tag{3.116}$$

so (3.112) may be written

$$\dot{P}_{11} - \dot{P}_{10} = \Gamma_{11}(P_{11}-P_{10}) + (P_{11}-P_{10})(\Gamma_{10}-\Gamma_{00})^* \tag{3.117}$$

Since

$$P_{10} = E\left\{\epsilon_1 X^*\right\} = E\left\{\epsilon_1(\hat{X}_1+\epsilon_1)^*\right\} = E\left\{\epsilon_1\hat{X}_1^*\right\} + P_{11} \tag{3.118}$$

we may, by choosing $P_{11}(0) = P_{10}(0) = \text{Cov}\left[X(0)\right]$, insure that $P_{11}(t) = P_{10}(t)$ for all time; this forces

$$E\left\{\epsilon_1\hat{X}_1^*\right\} = 0 \quad \text{for all time.} \tag{3.119}$$

This condition was assumed in the derivation of $\Omega_1(t)$, and is now verified. A parallel development will show that, by choosing $P_{22}(0) = P_{20}(0) = \psi_{X_o}$, we can guarantee that

$$P_{22}(t) = P_{20}(t) \quad \text{for all time;} \tag{3.120}$$

thus,

$$E\left\{\epsilon_2\hat{X}_2^*\right\} = 0 \quad \text{for all time.} \tag{3.121}$$

Note that (3.119) is true regardless of the form assumed for $\hat{X}_2$.

Thus far, we have assumed a specific form for $\hat{X}_2$. We will now relax this assumption and assume that $\hat{X}_2$ is obtained by an arbitrary function of the observed data Z_2. Then (3.100) becomes

$$\dot{\epsilon}_2 = \left[F - G_1K_1 + G_2K_2\right] X + G_1K_1\epsilon_1 - G_2K_2\epsilon_2 - \frac{d}{dt}\hat{X}_2 \tag{3.122}$$

and (3.101) becomes

$$\dot{\rho} = \Gamma\rho + B\eta_1 + C\dot{\hat{X}}_2 \tag{3.123}$$

where Γ is the same as before except for the third row, which becomes

$$\Gamma_{20} = \left[F - G_1K_1 + G_2K_2\right] \tag{3.124}$$

$$\Gamma_{21} = G_1K_1 \tag{3.125}$$

$$\Gamma_{22} = -G_2K_2 \tag{3.126}$$

and where

$$B = \begin{bmatrix} 0 \\ -P_{11}H_1^*R_1^{-1} \\ 0 \end{bmatrix} \tag{3.127}$$

and

$$C = \begin{bmatrix} 0 \\ 0 \\ -I \end{bmatrix} \tag{3.128}$$

Now $P = E\{\rho\rho^*\}$ satisfies the differential equation

$$\dot{P} = \Gamma P + P\Gamma^* + BR_1B^* + C\dot{\hat{X}}_2P^* + P\dot{\hat{X}}_2C^* \tag{3.129}$$

Since $C\dot{\hat{X}}_2\rho^*$ is of the form

$$C\dot{\hat{X}}_2'\rho^* = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -\dot{\hat{X}}_2x^* & -\dot{\hat{X}}_2\epsilon_1^* & -\dot{\hat{X}}_1\epsilon_2^* \end{bmatrix} \tag{3.130}$$

and $\rho\dot{\hat{X}}_2^* C^*$ is of the form

$$\rho\hat{X}_2^* C^* = \begin{bmatrix} 0 & 0 & -\chi\hat{X}_2^* \\ 0 & 0 & -\epsilon_1\hat{X}_2^* \\ 0 & 0 & -\epsilon_2\hat{X}_2^* \end{bmatrix} \tag{3.131}$$

the equations for $\dot{P}_{11}$ and $\dot{P}_{10}$ are unchanged from (3.110) and (3.111); thus equation (3.119) is valid regardless of the form of $\hat{X}_2$. In order to calculate $\hat{X}_1$, however, player number 1 must make some assumption about the form of $\hat{X}_2$.

One need be no more general in his assumptions about the form of $\hat{X}_2$ than to assume that $\hat{X}_2$ is generated by a 2nth order differential equation, because from player number 2's viewpoint the system is described by the set of differential equations

$$\dot{X} = FX - G_1K_1\hat{X}_1 + G_2U_2 \tag{3.132}$$

$$\begin{aligned}\dot{\hat{X}}_1 = {} & \left[F + G_2K_2\left[I - (P_{20}-P_{21})(P_{00}-P_{01})^{-1}\right]\right]\hat{X}_1 \\ & + P_{11}H_1^*R_1^{-1}\left[Z_1 - H_1\hat{X}_1\right] - G_1K_1\hat{X}_1\end{aligned} \tag{3.133}$$

and observation equation

$$Z_2 = H_2X + \eta_2 \tag{3.134}$$

We note that (3.132) and (3.133) may be written

$$\begin{bmatrix} \dot{X} \\ \dot{\hat{X}}_1 \end{bmatrix} = A\begin{bmatrix} X \\ \hat{X}_1 \end{bmatrix} + B\eta_1 + GU_2 \tag{3.135}$$

where

$$A = \left[\begin{array}{c|c} F & -G_1K_1 \\ \hline P_{11}H_1^*R_1^{-1}H_1 & F-G_1K_1+G_2K_2\left[I-(P_{20}-P_{21})(P_{00}-P_{01})^{-1}\right]-P_{11}H_1^*R_1^{-1}H_1 \end{array}\right] \tag{3.136}$$

$$B = \left[\begin{array}{c} 0 \\ \hline -P_{11}H_1^*R_1^{-1} \end{array}\right]; \quad G = \left[\begin{array}{c} G_2 \\ \hline 0 \end{array}\right] \tag{3.137}$$

and that (3.134) may be written

$$Z_2 = \left[\begin{array}{c|c} H_2 & 0 \end{array}\right]\begin{bmatrix} X \\ \hat{X}_1 \end{bmatrix} + \eta_2 \tag{3.138}$$

The problem of estimating X then becomes a standard linear state estimation problem. Thus, $\hat{X}_2$ satisfies

$$\hat{X}_2 = \left[\begin{array}{c|c} H_2 & 0 \end{array}\right]\hat{Y} \tag{3.139}$$

where

$$Y = \begin{bmatrix} X \\ \hat{X}_1 \end{bmatrix} \tag{3.140}$$

and $\hat{Y}$ satisfies

$$\dot{\hat{Y}} = AY + K\left[Z_2 - H_{20}\hat{Y}\right] \tag{3.141}$$

$$\hat{Y}(0) = \begin{bmatrix} \bar{X}(0) \\ \bar{X}(0) \end{bmatrix} \tag{3.142}$$

and

$$K = P_2 H_{20}^{*} R_2^{-1} \tag{3.143}$$

where

$$P_2 = \mathrm{Cov}\,(Y - \hat{Y}) \tag{3.144}$$

and

$$\dot{P}_2 = AP_2 + P_2A^{*} + BR_1B^{*} - P_2H_{20}^{*}R_2^{-1}H_{20}P_2 \tag{3.145}$$

Furthermore, the separation principle asserts that the optimal control is given by $U_2 = K_2\hat{X}_2$. This points up an important fact about the game problem: if one player is constrained to using an n-dimensional control strategy, the opposing player's unconstrained optimal control strategy, if it exists, is no more than 2n-dimensional.

Chapter 4

THE DIFFERENTIAL GAME PROBLEM WITH DIMENSIONALLY CONSTRAINED CONTROL STRATEGIES

4.1 Introduction

In Chapter 3 it was shown that when the two controllers were not constrained dimensionally they could not generate the conditional mean of the state with finite dimensional computing methods. In Chapter 2 it was shown that the optimal linear strategies can be written in terms of a conditional mean of the state plus some additional terms. It would thus appear that an overall optimal linear control strategy could not be generated unless the controller retains all of his past observations for use in computing the control. In many real engineering situations, however, such a requirement may not be practically met. Thus, we may wish to specify control strategies which are, first, computationally practical and, second, optimal within the class of strategies satisfying whatever computational efficiency criterion we select.

4.2 The Dimensionality Constraint

We shall examine here the nature of control strategies which are optimal within the class of strategies which can be written in the form

$$U_1(t) = K_1(t)\,\hat{X}_1(t) \tag{4.1}$$

$$U_2(t) = K_2(t)\,\hat{X}_2(t) \tag{4.2}$$

where $\hat{X}_1(t)$ and $\hat{X}_2(t)$ are in some sense n-dimensional "estimates" of the state which satisfy the differential equations

$$\dot{\hat{X}}_1 = (A_1-G_1K_1)\hat{X}_1 + B_1(Z_1-H_1\hat{X}_1) \; ; \; \hat{X}_1(0) = \bar{X}_0 \tag{4.3}$$

$$\dot{\hat{X}}_2 = (A_2+G_2K_2)\hat{X}_2 + B_2(Z_2-H_2\hat{X}_2) \; ; \; \hat{X}_2(0) = \bar{X}_0 \tag{4.4}$$

where $K_i(t)$, $A_i(t)$, and $B_i(t)$, $i = 1,2$, are unspecified and must be chosen in a manner which will optimize the criterion functional. A restriction of this problem which we may also wish to consider is that in which part of the parameters are specified and only the remaining unspecified quantities must be selected.

This approach has been considered in problems of both state estimation and stochastic optimal control [14]. In these cases its appeal is in its potential as a computationally efficient suboptimal estimation/control scheme. In the two-input situation the dimensionality constraint appears to be motivated more by necessity than by mere economy.

4.3 A Specialized Relationship

Rhodes and Luenberger [23] have taken the above approach to a problem closely related to the one under consideration here and have derived the following result, presented here without proof.

Theorem 4.3a

For the stochastic differential game problem described by (1.1A), (1.2A), and (1.3A) with controls given by

$$U_1^o = K_1 \hat{X}_1 \tag{4.5}$$

$$U_2^o = K_2 \hat{X}_2 \tag{4.6}$$

where K_1 and K_2 are given by (1.52) and (1.53), respectively, and $\hat{X}_1$ and $\hat{X}_2$ satisfy equations of the form (4.3) and (4.4), with

$$A_1 = F - G_1K_1 + G_2K_2\left[I - (P_{20}-P_{21})(P_{00}-P_{01})^{-1}\right] \tag{4.7}$$

$$A_2 = F - G_1K_1\left[I - (P_{10}-P_{12})(P_{00}-P_{02})^{-1}\right] + G_2K_2 \tag{4.8}$$

$$B_1 = P_{11} H_1^* R_1^{-1} \tag{4.9}$$

$$B_2 = P_{22} H_2^* R_2^{-1} \tag{4.10}$$

and with P_{ij} as defined in Chapter 3, the following inequalities hold:

$$E\left\{J(U_1^o, U_2^o)|\hat{X}_1\right\} \le E\left\{J(U_1, U_2^o)|\hat{X}_1\right\} \tag{4.11}$$

$$E\left\{J(U_1^o, U_2^o)|\hat{X}_2\right\} \ge E\left\{J(U_1^o, U_2)|\hat{X}_2\right\} \tag{4.12}$$

This result appears to be stronger than it is: (4.11) and (4.12) merely say that if the state estimate derived in Chapter 3 is used then the control strategy which optimizes the conditional expected value of the payoff functional is the certainty-equivalent strategy when the conditioning is done on the value of the state estimate. Equations (4.11) and (4.12) do not imply that the certainty-equivalent strategy optimizes the conditional expected value of the payoff when the conditioning is done on all past observations, nor do they say anything about the overall (unconditional) expected value.

4.4 Generalized Relationships for n-Dimensional Control Strategies

We wish to derive some necessary conditions for control strategies of the form described by (4.1) through (4.4) to satisfy the following saddle point conditions:

$$E\left\{J(U_1^o, U_2^o)\right\} \leq E\left\{J(U_1, U_2^o)\right\} \tag{4.13}$$

$$E\left\{J(U_1^o, U_2^o)\right\} \geq E\left\{J(U_1^o, U_2)\right\} \tag{4.14}$$

In order to put the problem in a format more suitable to our needs, we shall reformulate it somewhat. First, we define the state estimation errors ϵ_1 and ϵ_2 by

$$\epsilon_1 = X - \hat{X}_1 \tag{4.15}$$

$$\epsilon_2 = X - \hat{X}_2 \tag{4.16}$$

where $\hat{X}_1$ and $\hat{X}_2$ are generated by estimators of the form (4.3) and (4.4). Then, using (4.1), (4.2), (4.15), and (4.16), the system equation (1.1A) may be rewritten as

$$\dot{X} = (F - G_1K_1 + G_2K_2)\, X + G_1K_1\epsilon_1 - G_2K_2\epsilon_2 \tag{4.17}$$

Combining (4.17) with (4.3), (4.4), (4.15), (4.16) and (1.3A), we see that the estimation errors ϵ_1 and ϵ_2 satisfy

$$\dot{\epsilon}_1 = (F-A_1+G_2K_2)X + (A_1-B_1H_1)\epsilon_1 - G_2K_2\epsilon_2 - B_1\eta_1 \tag{4.18}$$

$$\dot{\epsilon}_2 = (F-A_2-G_1K_1)X + G_1K_1\epsilon_1 + (A_2-B_2H_2)\epsilon_2 - B_2\eta_2 \tag{4.19}$$

Therefore, (4.17), (4.18), and (4.19) taken together may be written

$$\dot{\rho} = \Gamma\rho + B\eta \tag{4.20}$$

where

$$\rho = \begin{bmatrix} X \\ \epsilon_1 \\ \epsilon_2 \end{bmatrix} \tag{4.21}$$

$$\eta = \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$$

$$\Gamma = \begin{bmatrix} \Gamma_{00} & \Gamma_{01} & \Gamma_{02} \\ \Gamma_{10} & \Gamma_{11} & \Gamma_{12} \\ \Gamma_{20} & \Gamma_{21} & \Gamma_{22} \end{bmatrix} = \left[\begin{array}{c|c|c} F-G_1K_1 + G_2K_2 & G_1K_1 & -G_2K_2 \\ \hline F-A_1 + G_2K_2 & A_1 - B_1H_1 & -G_2K_2 \\ \hline F-A_2 - G_1K_1 & G_1K_1 & A_2 - B_2H_2 \end{array}\right] \tag{4.22}$$

$$B = \begin{bmatrix} 0 & 0 \\ B_1 & 0 \\ 0 & B_2 \end{bmatrix} \tag{4.23}$$

Note that the quantity $U_1^*U_1 - U_2^*U_2$ may be written in terms of the vector ρ as

$$U_1^*U_1 - U_2^*U_2 = \rho^*Q\rho \tag{4.24}$$

where

$$Q = \left[\begin{array}{c:c:c} K_1^*K_1 - K_2^*K_2 & -K_1^*K_1 & K_2^*K_2 \\ \hdashline -K_1^*K_1 & K_1^*K_1 & 0 \\ \hdashline K_2^*K_2 & 0 & -K_2^*K_2 \end{array}\right] \tag{4.25}$$

Note also that $X^*(T)X(T)$ may be written in terms of the vector ρ as

$$X^*(T)X(T) = \rho^*(T)Q_T\rho(T) \tag{4.26}$$

where

$$Q_T = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{4.27}$$

In view of these relationships, we may write the payoff functional as

$$\begin{aligned} J &= E\left\{\rho^*(T)Q_T\rho(T) + \int_0^T \rho^*(\tau)Q(\tau)\rho(\tau)d\tau\right\} \\ &= \mathrm{tr}\left[E\left\{\rho(T)\ \rho^*(T)\ Q_T + \int_0^T \rho(\tau)\rho^*(\tau)Q(\tau)d\tau\right\}\right] \end{aligned} \tag{4.28}$$

And, defining $P(t)$ by

$$P(t) = E\left\{\rho(t)\rho^*(t)\right\} = \begin{bmatrix} P_{00} & P_{01} & P_{02} \\ P_{10} & P_{11} & P_{12} \\ P_{20} & P_{21} & P_{22} \end{bmatrix} \tag{4.29}$$

(4.28) may be written

$$J = \mathrm{tr}\left[P(T)Q_T + \int_0^T P(\tau)Q(\tau)d\tau\right] \tag{4.30}$$

Thus, the stochastic game problem has been converted to a deterministic game to which classical deterministic optimal control techniques may be applied.

In applying these classical techniques, we first note that the matrix P(t) satisfies the differential equation

$$\dot{P}(t) = \Gamma(t)\, P(t) + P(t)\, \Gamma^*(t) + B(t)\, R(t)\, B^*(t) \tag{4.31}$$

where

$$R(t) = \begin{bmatrix} R_1(t) & 0 \\ 0 & R_2(t) \end{bmatrix} \tag{4.32}$$

We then form the Hamiltonian corresponding to the payoff functional (4.30) and the differential equation constraint (4.31), which is

$$H(A_1,A_2,B_1,B_2,K_1,K_2) = -\mathrm{tr}\left[PQ\right] + \mathrm{tr}\left[\lambda(\Gamma P+P\Gamma^*+BRB^*)\right] \tag{4.33}$$

where λ is a Lagrange multiplier matrix, which satisfies the canonical Euler-Lagrange equation

$$\dot{\lambda} = -\frac{\partial H}{\partial P} = -\lambda\Gamma - \Gamma^*\lambda + Q$$

$$\lambda(T) = Q_T \tag{4.34}$$

where the gradient operation is as defined in Chapter 1 and thus

$$\left[\frac{\partial H}{\partial P}\right]_{ij} = \frac{\partial H}{\partial P_{ji}} \tag{4.35}$$

According to the Maximum Principle, we wish to select A_1, B_1, and K_1 so as to minimize H and to select A_2, B_2, and K_2 so as to maximize H. We shall see that the order of maximization and minimization does not matter. Since A_1, A_2, B_1, B_2, K_1, and K_2 are incorporated in various submatrices of Γ, G, and Q, we may partition the expression for the Hamiltonian in order to isolate those submatrices of interest for optimization with respect to a particular quantity. Thus, since the matrix B_1 appears as a part of Γ and G, we may write as a necessary condition

$$\frac{\partial H}{\partial B_1} = \frac{\partial}{\partial B_1} \operatorname{tr} \left[\lambda\Gamma P + \lambda P\Gamma^* + \lambda BRB^*\right] = \frac{\partial}{\partial B_1} \operatorname{tr} \left[2P\lambda\Gamma + \lambda BRB^*\right]$$

$$= \frac{\partial}{\partial B_1} \operatorname{tr} \left(2P\lambda \begin{bmatrix} 0 & 0 & 0 \\ 0 & -B_1H_1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 0 & 0 \\ 0 & B_1R_1B_1^* & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) = 0 \tag{4.36}$$

It is convenient at this point to partition the P and λ matrices by

$$P = \begin{bmatrix} P_0 \\ \hdashline P_1 \\ \hdashline P_2 \end{bmatrix} \; ; \; \lambda = \left[\lambda_0 \mid \lambda_1 \mid \lambda_2\right] \tag{4.37}$$

Here P_0, P_1, and P_2 are n x 3n matrices, and λ_0, λ_1, and λ_2 are 3n x n matrices. These matrices may be further partitioned when convenient; e.g.,

$$\lambda_1 = \begin{bmatrix} \lambda_{01} \\ \hdashline \lambda_{11} \\ \hdashline \lambda_{21} \end{bmatrix} \; ; \; P_1 = \left[P_{10} \,\vdots\, P_{11} \,\vdots\, P_{12}\right] \tag{4.38}$$

where the λ_{ii} and P_{ii} , $i = 0,1,2$, are n x n matrices.

Using this notation, we may write (4.36) as

$$\frac{\partial H}{\partial B_1} = \frac{\partial}{\partial B_1} \mathrm{tr}\left[-2P_1\lambda_1 B_1 H_1 + \lambda_{11} B_1 R_1 B_1^*\right]$$

$$= -2H_1 P_1 \lambda_1 + 2R_1 B_1^* \lambda_{11} = 0$$

or

$$\lambda_{11}^* B_1 = \lambda_1^* P_1^* H_1^* R_1^{-1} \tag{4.39}$$

Equation (4.39) is a necessary condition for minimization with respect to the matrix B_1. Completely analogous arguments regarding the matrix B_2 lead to the expression

$$\lambda_{22}^* B_2 = \lambda_2^* P_2^* H_2^* R_2^{-1} \tag{4.40}$$

The Hamiltonian (4.33) is also quadratic in K_1, so we write

$$\frac{\partial H}{\partial K_1} = \frac{\partial}{\partial K_1} \mathrm{tr}\left[-PQ + \lambda\Gamma P + \lambda P\Gamma^*\right] = \frac{\partial}{\partial K_1} \mathrm{tr}\left[-PQ + 2P\lambda\Gamma\right] = 0 \tag{4.41}$$

Equation (4.41) may also be written

$$\frac{\partial}{\partial K_1} \mathrm{tr}\left(-P\begin{pmatrix} K_1^*K_1 & -K_1^*K_1 & 0 \\ -K_1^*K_1 & K_1^*K_1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 2P\lambda \begin{pmatrix} -G_1K_1 & G_1K_1 & 0 \\ 0 & 0 & 0 \\ -G_1K_1 & G_1K_1 & 0 \end{pmatrix} \right) = 0 \tag{4.42}$$

Now note that we may write

$$\operatorname{tr}\left(P\begin{bmatrix} K_1^*K_1 & -K_1^*K_1 & 0 \\ -K_1^*K_1 & K_1^*K_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \operatorname{tr}\left((P_1-P_0)\begin{bmatrix} -I \\ I \\ 0 \end{bmatrix} K_1^*K_1\right) \qquad (4.43)$$

where P_1 and P_0 are as defined in (4.37). Also note that we may write

$$\operatorname{tr}\left(P\lambda\begin{bmatrix} -G_1K_1 & G_1K_1 & 0 \\ 0 & 0 & 0 \\ -G_1K_1 & G_1K_1 & 0 \end{bmatrix}\right) = \operatorname{tr}\left[(P_1-P_0)(\lambda_0+\lambda_2)\; G_1K_1\right] \qquad (4.44)$$

Substituting (4.43) and (4.44) into (4.42), we have

$$\frac{\partial}{\partial K_1}\operatorname{tr}\left\{-(P_1-P_0)\begin{bmatrix} -I \\ I \\ 0 \end{bmatrix}K_1^*K_1 + 2(P_1-P_0)(\lambda_0+\lambda_2)G_1K_1\right\} =$$

$$-2(P_1-P_0)\begin{bmatrix} -I \\ I \\ 0 \end{bmatrix}K_1^* + 2(P_1-P_0)(\lambda_0+\lambda_2)\; G_1 \qquad = 0 \quad (4.45)$$

or

$$(P_1-P_0)\left((\lambda_0+\lambda_2)G_1 - \begin{bmatrix} -I \\ I \\ 0 \end{bmatrix}K_1^*\right) = 0 \qquad (4.46)$$

Again, analogous arguments apply to the feedback matrix K_2 and produce the expression

$$(P_2'-P_0)\left((\lambda_0+\lambda_1)G_2 - \begin{bmatrix} -I \\ I \\ 0 \end{bmatrix} K_2^*\right) = 0 \qquad (4.47)$$

As opposed to the case for the B and K matrices, the Hamiltonian is linear in the A matrices; consequently, the maximum principle dictates in the case of the minimizing matrix A_1

$$\frac{\partial}{\partial a_{1ij}} \begin{cases} > 0, & a_{1ij} = a_{1ij} \max \\ < 0, & a_{1ij} = a_{1ij} \min \end{cases} \tag{4.48}$$

where

$$A_1 = \begin{bmatrix} a_{111} & a_{112} & -- & a_{11n} \\ a_{121} & a_{122} & -- & a_{12n} \\ a_{1n1} & a_{1n2} & ---- & a_{1nn} \end{bmatrix} \tag{4.49}$$

For matrix A_2,

$$\frac{\partial H}{\partial a_{2ij}} \begin{cases} > 0, & a_{2ij} = a_{2ij} \min \\ < 0, & a_{2ij} = a_{2ij} \max \end{cases} \tag{4.50}$$

and an expression analogous to (4.49) defines the elements of A_2.

Singular cases exist where neither inequality is satisfied in (4.48) or (4.50), i.e., where the derivative is equal to zero. In such cases, if the condition can be sustained, some higher-order test, such as the Kelley necessary condition [19], may be applied in an attempt to determine the values of the elements. It will now be shown that, if the necessary conditions for K_1 and B_1 are satisfied, the entire trajectory lies on a singular surface for A_1.

4.5 A Singular Surface

It was mentioned in section 4.2 that in some restrictions of the game problem there defined some of the parameters might be specified and thus not available for optimization. We shall see that it is only under these conditions that the optimal A coefficient matrices would be chosen by (4.48) or (4.50), i.e., be bang-bang. Otherwise, the gradient of the Hamiltonian with respect to the A matrices is zero during the entire interval [0, T]. This is shown for the case of the A_1 matrix as follows:

$$\frac{\partial H}{\partial A_1} = \frac{\partial}{\partial A_1} \operatorname{tr}\left[2P\lambda\dot{\Gamma}\right] = \frac{\partial}{\partial A_1} \operatorname{tr}\left(2P\lambda \begin{bmatrix} 0 & 0 & 0 \\ -A_1 & A_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)$$

$$= 2\frac{\partial}{\partial A_1} \operatorname{tr}\left[(P_1-P_0)\lambda_1 A_1\right] = 2(P_1-P_0)\lambda_1 \qquad (4.51)$$

From the boundary condition given in (4.34), we see that $\lambda_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ at $t = T$ and, thus, a singular condition exists at the boundary. We shall now show that

$$\frac{d}{dt}(P_1-P_0)\lambda_1 \equiv 0 \qquad 0 \le t \le T \qquad (4.52)$$

whenever the optimality conditions (4.39) and (4.46) for B_1 and K_1, respectively, are satisfied.

Consider

$$\frac{d}{dt}(P_1-P_0)\lambda_1 = (P_1-P_0)\dot{\lambda}_1 + (\dot{P}_1-\dot{P}_0)\lambda_1 \qquad (4.53)$$

Since the matrix λ satisfies (4.34), the submatrix λ_1 satisfies

$$\dot{\lambda}_1 = -\Gamma^*\lambda_1 - \lambda \begin{pmatrix} \Gamma_{01} \\ \Gamma_{11} \\ \Gamma_{21} \end{pmatrix} + \begin{pmatrix} -K_1^*K_1 \\ K_1^*K_1 \\ 0 \end{pmatrix}$$

$$= -\Gamma^*\lambda_1 - \lambda_1\Gamma_{11} - (\lambda_0+\lambda_2)G_1K_1 + \begin{pmatrix} -I \\ I \\ I \end{pmatrix} K_1^*K_1$$

$$\lambda_1(T) = 0 \tag{4.54}$$

Then, because of relationship (4.46), the expression $(P_1-P_0)\dot{\lambda}_1$ becomes

$$(P_1-P_0)\dot{\lambda}_1 = -(P_1-P_0)\Gamma^*\lambda_1 - (P_1-P_0)\lambda_1\Gamma_{11} \tag{4.55}$$

Since the matrix P satisfies (4.31), the submatrix P_1 satisfies

$$\dot{P}_1 = P_1\Gamma^* + \begin{bmatrix} \Gamma_{10} & \Gamma_{11} & \Gamma_{12} \end{bmatrix} P + \begin{bmatrix} 0 & B_1R_1B_1^* & 0 \end{bmatrix} \tag{4.56A}$$

and the submatrix P_0 satisfies

$$\dot{P}_0 = P_0\Gamma^* + \begin{bmatrix} \Gamma_{00} & \Gamma_{01} & \Gamma_{02} \end{bmatrix} P \tag{4.56B}$$

Thus,

$$\dot{P}_1 - \dot{P}_0 = (P_1-P_0)\Gamma^* + \begin{bmatrix} \Gamma_{10}-\Gamma_{00} & \Gamma_{11}-\Gamma_{01} & \Gamma_{12}-\Gamma_{02} \end{bmatrix} P$$

$$+ \begin{bmatrix} 0 & B_1R_1B_1^* & 0 \end{bmatrix} \tag{4.57}$$

Now note that $\Gamma_{12} = \Gamma_{02} = -G_2K_2$; therefore,

$$\Gamma_{12} - \Gamma_{02} = 0 \tag{4.58}$$

Also note that

$$\Gamma_{10} - \Gamma_{00} = -(\Gamma_{11}-\Gamma_{01}) - B_1H_1 \tag{4.59}$$

Substituting (4.58) and (4.59) into (4.57), we have

$$\begin{aligned}\dot{P}_1 - \dot{P}_0 = {} & (P_1-P_0)\Gamma^* + (\Gamma_{11}-\Gamma_{01})(P_1-P_0) - B_1H_1P_0 \\ & + \begin{bmatrix} 0 & B_1R_1B_1^* & 0 \end{bmatrix} .\end{aligned} \tag{4.60}$$

Therefore,

$$\begin{aligned}(\dot{P}_1-\dot{P}_0)\lambda_1 = {} & (P_1-P_0)\Gamma^*\lambda_1 + (\Gamma_{11}-\Gamma_{01})(P_1-P_0)\lambda_1 \\ & - B_1H_1P_0\lambda_1 + \begin{bmatrix} 0 & B_1R_1B_1^* & 0 \end{bmatrix}\lambda_1\end{aligned} \tag{4.61}$$

Using equation (4.39), we see that the last term of equation (4.61) may be written

$$\begin{bmatrix} 0 & B_1R_1B_1^* & 0 \end{bmatrix}\lambda_1 = B_1H_1P_1\lambda_1 \tag{4.62}$$

so (4.61) becomes

$$\begin{aligned}(\dot{P}_1-\dot{P}_0)\lambda_1 = {} & (P_1-P_0)\Gamma^*\lambda_1 + (\Gamma_{11}-\Gamma_{01})(P_1-P_0)\lambda_1 \\ & + B_1H_1(P_1-P_0)\lambda_1\end{aligned} \tag{4.63}$$

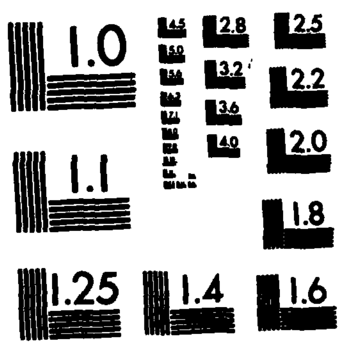
AD-A124 627 STOCHASTIC DIFFERENTIAL GAMES WITH COMPLEXITY CONSTRAINED STRATEGIES(U) CALIFORNIA UNIV LOS ANGELES SCHOOL OF ENGINEERING AND APPLIED SCIENCE D M STUART MAR 82 DASG60-80-C-0007 2/2

UNCLASSIFIED F/G 12/1 NL

END

FILMED

DTIC


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

Adding (4.55) and (4.63) and combining terms, we have

$$\frac{d}{dt}(P_1-P_0)\lambda_1 = -(P_1-P_0)\lambda_1\Gamma_{11} + (\Gamma_{11}-\Gamma_{01} + B_1H_1)(P_1-P_0)\lambda_1 \tag{4.64}$$

Since this equation is linear in $(P_1-P_0)\lambda$ and is homogeneous, and since $(P_1-P_0)\lambda_1 = 0$ at $t = T$, we must have

$$\frac{\partial H}{\partial A_1} = (P_1-P_0)\lambda_1 \equiv 0 \qquad 0 \le t \le T \tag{4.65}$$

Similar relations apply for $\frac{\partial H}{\partial A_2}$. Thus, if we choose the B and K matrices on the basis of the maximum principle, we must look beyond the maximum principle for help in specifying the elements of the A matrices.

4.6 Specifying the A Matrices[1]

Equation (4.65) indicates that, if the optimal values for B_1 and K_1 are employed, the state trajectory lies in a surface in state-space on which the Hamiltonian is first-order independent of variations in A_1; an analogous condition exists with regard to A_2. In some control situations of this type, we may make use of higher-order necessary conditions on variations of A_1 or A_2. A well-known second-order necessary condition is the Legendre necessary condition, expressed as

$$\frac{\partial^2 H}{\partial A_1^{\,2}} \ge 0 \tag{4.66}$$

[1] This section is based in part on material presented by Johansen [14].

This condition is, of course, trivially satisfied for the game problem under consideration here because of (4.65). In cases where (4.66) obtains with equality, another necessary condition, the Kelley necessary condition, is sometimes applied. This condition is expressed as

$$(-1)^k \frac{\partial H}{\partial A_1}\left\{\frac{d^{2k}}{dt^{2k}}\frac{\partial H}{\partial A_1}\right\} \geq 0 \quad k=0,1,2,\ldots \tag{4.67}$$

However, since the differential equation describing $\frac{\partial H}{\partial A_1}$ is seen to be linear in $\frac{\partial H}{\partial A_1}$ and is homogeneous, and since $\frac{\partial H}{\partial A_1}$ is zero on the boundary, all time derivatives of $\frac{\partial H}{\partial A_1}$ are zero on the singular surface and (4.67) is also trivially satisfied.

The reason for the apparent paradox is that the problem has been given too may degrees of freedom: if the B and K matrices are chosen optimally, the payoff is actually independent of the A matrices. This aspect of the problem is related to the non-uniqueness of optimal control strategies of the form given by (4.1) through (4.4). As an illustration of this non-unique characteristic, we may consider the strategy of controller number 1, which may be written in the form

$$U_1 = K_1 \hat{X}_1 \tag{4.68}$$

$$\dot{\hat{X}}_1 = (A_1 - G_1K_1 - B_1H_1)\hat{X}_1 + B_1Z_1 \; ; \; \hat{X}_1(0) = \bar{X}_0 \tag{4.69}$$

Assume for the moment that A_1, B_1, and K_1 have been specified. As a preliminary step, for notational convenience, we shall define a new matrix A_0:

$$A_0 = A_1 - G_1K_1 - B_1H_1 \tag{4.70}$$

Then (4.68) and (4.69) become

$$U_1 = K_1\hat{X}_1 \tag{4.71}$$

$$\dot{\hat{X}}_1 = A_0\hat{X}_1 + B_1Z_1;\ \hat{X}_1(0) = \bar{X}_0 \tag{4.72}$$

We shall now show that we may arbitrarily change A_0 to a new matrix A_0' and that, by adjusting the matrices B_1 and K_1, we can obtain the same control strategy U_1.

We first define a new variable $\hat{X}_1'$ by

$$\hat{X}_1' = D\hat{X}_1 \tag{4.73}$$

where D is a differentiable nonsingular matrix to be specified. Then (4.71) and (4.72) may be written

$$U_1 = K_1D^{-1}\hat{X}_1' \tag{4.74}$$

$$\dot{\hat{X}}_1' = (\dot{D}+DA_0)D^{-1}\hat{X}_1' + DB_1Z_1\ ; \quad \hat{X}_1'(0) = D(0)\ \bar{X}_0 \tag{4.75}$$

We then adjust the matrices B_1 and K_1 by the relationships

$$B_1' = DB_1 \tag{4.76}$$

$$K_1' = K_1D^{-1} \tag{4.77}$$

Next we choose the matrix D, requiring that

$$\dot{D} + DA_0 = A_0' \tag{4.78}$$

This may be done by defining two matrices ϕ_1 and ϕ_2 which satisfy the differential equations

$$\phi_1 = -\phi_1 A_0' \qquad \phi_1(0) = I \tag{4.79}$$

$$\phi_2 = -\phi_2 A_0' \qquad \phi_2(0) = I \tag{4.80}$$

Then, by direct substitution into (4.78), we verify that

$$D = \phi_1^{-1} D_0 \phi_2 \tag{4.81}$$

where D_0 is a nonsingular constant matrix, which we may choose in such a manner that

$$\hat{X}_1(0) = \bar{X}_0 \tag{4.82}$$

Then from (4.75) we infer that

$$D_0 = I \tag{4.83}$$

Thus, the control strategy U_1 may be written

$$U_1 = B_1' \hat{X}_1' \tag{4.84}$$

$$\dot{\hat{X}}_1' = A_0' \hat{X}_1' + B_1' Z_1 \; ; \quad \hat{X}_1(0) = \bar{X}_0 \tag{4.85}$$

We conclude that only in cases where special restrictions apply to the form of the B or K matrices are we unable to arbitrarily specify the A matrices.

This is not to say that specification of particular values for the A matrices can be done without concern over the implications, since fixing these values also fixes the values of the B and K matrices and may lead to excessively large or impractical values for them. In some cases, careful selection of the A matrices can lead to considerable simplication of the computation leading to the B and K matrices. A case in point is the one-sided problem.

4.7 Relationships with the One-Sided Case and the Separation Principle

When examining (4.30) in detail, one observes a certain similarity between it and the expression for the Kalman filter gain, which is

$$B_1 = \bar{P}_{11} H_1^* R_1^{-1} \tag{4.86}$$

Upon expanding,

$$\lambda_1^* P_1^* = \lambda_{01}^* P_{10}^* + \lambda_{11}^* P_{11}^* + \lambda_{21}^* P_{12}^*$$

so (4.39) may be written

$$\lambda_{11}^* B_1 = \left[\lambda_{01}^* P_{10}^* + \lambda_{11}^* P_{11}^* + \lambda_{21}^* P_{12}^*\right] H_1^* R_1^{-1} \tag{4.87}$$

Then, if $\lambda_{01}^* P_{10}^* + \lambda_{21}^* P_{12}^* = 0$, (4.87) would be satisfied by (4.86). By examining the differential equations describing $\lambda_{01}^* P_{10}^*$ and $\lambda_{21}^* P_{12}^*$, however, it can be seen that their sum is not identically zero, $0 \leq t \leq T$. This is an example of the non-separability of the problem: the filter gain B_1 depends explicitly on the elements of the λ matrix, which in turn depend on the feedback gain K_1.

A similar situation is encountered when we examine (4.46), describing K_1. If a value for K_1 could be found satisfying

$$(\lambda_0+\lambda_2)G_1 = \begin{bmatrix} -I \\ I \\ 0 \end{bmatrix} K_1^* \tag{4.88}$$

this value would also satisfy (4.46) and would be explicitly independent of the elements of the P matrix. For (4.88) to be satisfied, however, it would be necessary that $(\lambda_{20}+\lambda_{22})G_1 = 0$; and this condition is not generally true. Another condition which would render K_1 independent of the P matrix is

$$(P_{12}-P_{02})(\lambda_{20}+\lambda_{22})G_1 = 0 \tag{4.89}$$

Again, however, examination of the differential equations describing P and λ shows that (4.89) is not generally true.

Therefore, as a result of the above situations, the solutions to (4.39) and (4.46) are

$$B_1 = P_{11}H_1^*R_1^{-1} + \lambda_{11}^{*-1}\left[\lambda_{01}^* P_{10}^* + \lambda_{21}^* P_{12}^*\right] H_1^* R_1^{-1} \tag{4.90}$$

$$K_1 = G_1^* (\lambda_0+\lambda_2)^* (P_1-P_0)^* \left[P_{00}-P_{10}+P_{11}-P_{01}\right]^{-1} \tag{4.91}$$

assuming that the indicated inverses exist.

Now notice that, since we may choose the A matrix arbitrarily, as shown in section 4.6, a particularly good choice is A = F, which, as can be easily demonstrated, results in

$$P_{01} = P_{11} = P_{10} \tag{4.92}$$

i.e., the estimation error ϵ_1 is uncorrelated with the extimate $\hat{X}_1$. Because of (4.92), we may make use of relationship (4.65) in a special way: for the one-sided case we may discard the variables with "2" subscripts; therefore, remembering that $P_{11} - P_{01} = 0$, (4.65) may be written

$$(P_1 - P_0)\lambda_1 = \begin{bmatrix} P_{10} - P_{00} & P_{11} - P_{01} \end{bmatrix} \begin{bmatrix} \lambda_{01} \\ \lambda_{11} \end{bmatrix} = (P_{10} - P_{00})\lambda_1 = 0 \tag{4.93}$$

Then, because of (4.93), we may write

$$P_{10}\,\lambda_{01} = P_{10}(P_{10} - P_{00})^{-1}\,(P_{10} - P_{00})\lambda_{01} = 0 \tag{4.94}$$

and thus (4.90) becomes

$$B_1 = P_{11}\,H_1^{*}\,R_1^{-1} \tag{4.95}$$

i.e., the expression for the filter gain becomes explicitly independent of the λ matrix.

Similar things happen to equation (4.91) when (4.92) is satisfied. First, (4.46) becomes

$$\begin{bmatrix} P_{10} - P_{00} & 0 \end{bmatrix} \left(\begin{bmatrix} \lambda_{00} \\ \lambda_{10} \end{bmatrix} G_1 - \begin{bmatrix} -I \\ I \end{bmatrix} K_1^{*} \right) = 0 \tag{4.96}$$

This may be written as

$$(P_{10} - P_{00})\begin{bmatrix} \lambda_{00}G_1 + K_1^{*} \end{bmatrix} = 0 \tag{4.97}$$

which will be satisfied when

$$K_1 = -G_1^* \lambda_{00} \tag{4.98}$$

This is the deterministically optimal feedback gain, as can be seen from the fact that the matrix λ_{00} satisfies the differential equation

$$\dot{\lambda}_{00} = -(\lambda_{00}\Gamma_{00} + \lambda_{01}\Gamma_{10}) - (\lambda_{00}\Gamma_{00} + \lambda_{01}\Gamma_{10})^* + K_1^* K_1 \tag{4.99}$$

Since we have chosen $A = F$, $\Gamma_{10} = 0$, so (4.99) reduces to

$$\dot{\lambda}_{00} = -\lambda_{00}\Gamma_{00} - (\lambda_{00}\Gamma_{00})^* + K_1^* K_1 \tag{4.100}$$

Substituting (4.98) into (4.100) and remembering that $\Gamma_{00} = F - G_1 K_1$, we have

$$\dot{\lambda}_{00} = -\lambda_{00} F - F^* \lambda_{00} - \lambda_{00} G_1 G_1^* \lambda_{00} \tag{4.101}$$

$$\lambda_{00}(T) = I$$

This matrix Ricatti equation is the same as that satisfied by

$$\phi^*(T,t)\left[I + T_1 T_1^*\right]^{-1} \phi(T,t)$$

showing that (4.98) is identical to (1.52) and is thus the deterministically optimal feedback gain for the one-sided case.

4.8 The Matrices B_1, B_2, K_1, and K_2

The Hamiltonian is quadratic in B_1, B_2, K_1, and K_2, and it is thus possible to obtain explicit expressions for these matrices in terms of the elements of the P and λ matrices. This has been done in (4.90) and

(4.91) for B_1 and K_1, respectively; similar expressions may be obtained for B_2 and K_2. When these expressions are substituted into the Γ matrix in equations (4.31) and (4.34), these two equations constitute a non-linear two-point boundary value problem. Since both P and λ are symmetric and 3n x 3n, the total number of variables is 3n(3n+1). For the simplest non-trivial example, n = 1; this implies that the nonlinear problem has twelve variables.

Solution of nonlinear two-point boundary value problems by iterative computational methods is a subject covered fairly well in the literature [2,10,16,20] and will not be discussed here in any detail. However, when such problems arise out of differential games, two important aspects must be considered. The first of these is the number of variables involved, large even by optimal control standards. Whereas a one-sided stochastic optimal control problem with n = 1 involves solution for two variables, the two-player case of the same dimension involves solution for twelve variables. The second aspect is the particular nature of the nonlinear equations: specifically, if the elements of the Γ, G, and Q matrices in (4.31) and (4.34) were known, these equations would be linear differential equations with one-sided boundary conditions. This fact suggests a fairly simple iterative computational scheme:

i) Choose an initial set of values for P(t) and λ(t).

ii) On the basis of (i), compute the values of the elements of Γ(t), G(t), and Q(t).

iii) Using the values computed in (ii), solve (4.31) and (4.34) as linear equations with one-sided boundary values.

iv) Using the solution obtained in (iii), update the calculations done in (ii).

v) Repeat until solution converges.

Convergence in step (v) is not guaranteed, of course, and depends on an intelligent choice of initial values in step (i) as well as fortuitous conditioning of the equations by the physical parameters of the system and by a proper choice of the A matrices.

As an alternative to solving the nonlinear problem, we may consider a direct approach to optimization by some gradient technique; however, it would seem that the convergence difficulties inherent in gradient computational solutions of one-sided optimal control problems would be increased enormously when two sets of variables are involved, one set minimizing and the other maximizing. Thus, it appears that the indirect approach to differential game problems described in this chapter is, at least in some situations, the most promising method.

Chapter 5

OBTAINING PAYOFF BOUNDS FOR CONSTRAINED STRATEGIES

5.1 Removing Constraints on One Controller

In Chapter 4 it was indicated that the optimal coefficient matrices could be obtained by solving a nonlinear differential equation with split boundary conditions whose order is $3n(3n+1)$. It was also pointed out that the computational difficulties of doing so are potentially great. It is thus the natural question to ask what is obtained in return for the effort required to solve the nonlinear problem, particularly in view of the fact that the solutions obtained give only control functionals which are optimal within a certain, somewhat artificial constraint.

Fortunately, this question is easier to answer than is that which inquires as to the optimal control itself. Once the constrained problem of Chapter 4 is solved, the solution so obtained may be evaluated by either player by comparing the payoff under the constrained solution to the payoff which would result should his opponent be unconstrained. This comparison is easily made, since the separation principle tells us that if one controller uses a set n-dimensional control-generating system, his opponent's optimal opposing strategy is generated by a 2nth order differential equation.

This fact allows either controller, once he has established the form of his control-generating system and its parameters, to obtain a worst-case bound on the payoff when he employs that strategy. He is not able to obtain a best-case bound, because the best-case payoff depends upon how poorly the opponent chooses his strategy and may be unbounded. He is able to solve, of course, for the payoff when his opponent uses an optimal constrained control.

5.2 Obtaining Worst-Case Bounds on Payoff

When one player, say number 1, specifies the parameters of his n-dimensional control, the system from player number 2's viewpoint may be described by the 2n-dimensional system of equations

$$\begin{bmatrix} \dot{X} \\ \dot{\epsilon}_1 \end{bmatrix} = \begin{bmatrix} F-G_1K_1 & G_1K_1 \\ F-A_1 & A-B_1H_1 \end{bmatrix} \begin{bmatrix} X \\ \epsilon_1 \end{bmatrix} - \begin{bmatrix} 0 \\ B_1 \end{bmatrix} \eta_1 + \begin{bmatrix} G_2 \\ G_2 \end{bmatrix} U_2 \tag{5.1}$$

Player number 2's observation equation remains

$$Z_2 = H_2X_2 + \eta_2 \tag{5.2}$$

which may be rewritten

$$Z_2 = \left[H_2 \;\vdots\; 0 \right] \begin{bmatrix} X \\ \epsilon_1 \end{bmatrix} + \eta_1 \tag{5.3}$$

Thus, the payoff functional may be rewritten as

$$J(U_2) = E\left\{ \left[X^*(T) \;\; \epsilon_1^*(T) \right] \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X(T) \\ \epsilon_1(T) \end{bmatrix} + \int_0^T \left(\left[X^*(\tau), \epsilon_1^*(\tau) \right] \right.\right.$$

$$\left.\left. \begin{bmatrix} K_1^*K_1 & -K_1^*K_1 \\ -K_1^*K_1 & K_1^*K_1 \end{bmatrix} \begin{bmatrix} X(\tau) \\ \epsilon_1(\tau) \end{bmatrix} - U_2^*(\tau)U_2(\tau) \right) d\tau \right\} \tag{5.4}$$

Equations (5.1) through (5.4) constitute a standard one-sided stochastic optimal control problem, the solution to which is given by

$$U_2^o = -G_o^*(t)S(t)\hat{X}(t) \tag{5.5}$$

where the matrix S(t) satisfies the differential equation

$$\dot{S} = SF_o - F_o^*S + SG_oG_o^*S - A \tag{5.6}$$

$$S(T) = Q_T$$

and where $\hat{X}$ satisfies the differential equation

$$\dot{\hat{X}} = \left[F_o - G_oG_o^*S\right]\hat{X} + K\left[z_2 - H_o\hat{X}\right] \tag{5.7}$$

where

$$K = P_{\epsilon\epsilon}H_o^*R_2^{-1} \tag{5.8}$$

where $P_{\epsilon\epsilon}$ satisfies

$$P_{\epsilon\epsilon} = F_oP_{\epsilon\epsilon} + P_{\epsilon\epsilon}F_o + P_{\epsilon\epsilon}H_o^*R_2^{-1}H_oP_{\epsilon\epsilon} + B_oR_1B_o^* \tag{5.9}$$

$$P_{\epsilon\epsilon}(0) = \begin{bmatrix} \psi_{X_o} & 0 \\ 0 & \psi_{X_o} \end{bmatrix}$$

and where we make the identification

$$
\begin{aligned}
F_o &= \begin{bmatrix} F-G_1K_1 & G_1K_1 \\ G-A_1 & A-B_1H_1 \end{bmatrix} & B_o &= \begin{bmatrix} 0 \\ B_1 \end{bmatrix} \quad X = \begin{bmatrix} X \\ \epsilon_1 \end{bmatrix} \\
G_o &= \begin{bmatrix} G_2 \\ G_2 \end{bmatrix} & H_o &= \begin{bmatrix} H_2 & 0 \end{bmatrix} \\
A &= \begin{bmatrix} K_1^*K_1 & -K_1^*K_1 \\ -K_1^*K_1 & K_1^*K_1 \end{bmatrix} & Q_T &= \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}
\end{aligned}
\tag{5.10}
$$

and R_1 and R_2 are noise covariances as defined previously.

The worst case bound is then obtained by inserting the optimal opposing strategy given by (5.5) through (5.10) into the functional (5.4) and evaluating it.

An interesting parallel to the development of Chapter 4 is the problem of choosing an optimum control strategy of the form

$$U_1 = K_1\hat{X}_1 \tag{5.11}$$

where $\hat{X}_1$ satisfies

$$\dot{\hat{X}}_1 = (A_1-G_1K_1)\hat{X}_1 + B_1(z_1-H_1\hat{X}_1) \tag{5.12}$$

and where controller number 2 is unconstrained and, therefore, uses a strategy of the form (5.5) through (5.10).

Chapter 6

CONCLUSION

6.1 Summary

In Chapters 1 and 2 a general stochastic differential game characterized by linear differential equations, a quadratic cost functional, and additive white Gaussian observation noise was presented. It was shown that the certainty-equivalence principle, valid for one-player game situations, was not correct for two-player problems. Specifically, if one assumed a control form consisting of a matrix transformation of the conditional mean plus a linear operation on the residuals, the matrix transformation was the deterministic optimal feedback gain; however, the linear operation on the residuals was not a zero operation, as was true in the one-sided case.

In Chapter 3 it was shown that in order to generate the conditional mean of the state vector, each player was required to store all past observations. However, since this was considered to be an impractical requirement for many practical systems, a state estimation scheme was developed which generated the estimate as a solution to a differential equation forced by the observations. The order of this differential equation was that of the controlled system.

Chapter 4 generalized this concept to that of optimal control strategies within the class of strategies generated as solutions to differential equations forced by the observations. The order of these differential equations was taken to be that of the controlled system. This approach resulted in expressions for the control strategies given in terms of functions which are known only as solutions to a set of nonlinear differential equations with split boundary conditions. A computational approach to solving these equations is suggested.

In Chapter 5 it was pointed out that, once a set of dimensionally constrained strategies is calculated, either player may compute a worst-case bound on the payoff by assuming his opponent uses an unconstrained, and therefore higher-dimensional, strategy. Formulas are given for computing this bound.

6.2 Results of Research

Optimal dimensionally-constrained control strategies are of interest in practical problems where computational capacity is limited. A great deal of importance in choosing a control strategy is bound up in the question of what one is willing to assume about his opponent's strategy. Computation of an optimal unconstrained but linear strategy is quite complicated, and so it is reasonable to assume that one's opponent will impose some complexity constraint upon himself. As we have seen in Chapter 4, there are various ways in which such constraints may be imposed, e.g., by specifying the order of the control-generating differential equation. The specific form of the self-imposed constraint of one player is unknown to the other player and may not reasonably be treated as a random variable in most cases. For this reason it is of interest to compute worst-case bounds on the payoff under varying assumptions about the player's strategies. These bounds may then be used as a guide to making engineering decisions about the utility of a particular strategy.

6.3 Suggestions for Future Investigations

In this work we have analyzed a linear-quadratic-Gaussian problem of a rather uncomplicated type. The natural extensions of this work should follow the patterns established by investigators of one-sided stochastic control problems: examinations of cases with plant noise, colored noise, or no noise and cases where the payoff is described in terms of non-negative definite rather than positive definite matrices.

Investigation should also be continued into the computational aspects of the problem. The indirect approach described in Chapter 4 results in a set of non-linear equations with split boundary conditions. These equations are of such a nature that when the control gains are fixed the equations may be separated into sets of linear differential equations with one-sided boundary conditions. It may be possible to exploit this property to simplify the computational problem.

It would also be of interest to investigate the problem of direct optimization by some type of gradient method or local optimization scheme and to determine how the two-sided nature of the problem affects convergence.

BIBLIOGRAPHY

1. Aoki, Masanao. Optimization of stochastic systems. New York, Academic Press, 1967.

2. Bellman, Richard E., and Robert E. Kalaba. Quasilinearization and nonlinear boundary-value problems. New York, American Elsevier, 1965.

3. Behn, R.D., and Y.C. Ho. "On a class of linear stochastic differential games." IEEE Transactions on Automatic Control, Vol. AC-13, No. 3, pp. 227-239, June, 1968.

4. Berkowitz, L.D. "A variational approach to differential games." Annals of Math. Studies, No. 52, Princeton, N.J., Princeton University Press, 1964, pp. 127-173.

5. Berkowitz, L.D., and W.H. Fleming. "On differential games with integral payoff." Annals of Math. Studies, No. 39, Princeton, N.J., Princeton University Press, 1957, pp. 413-435.

6. Bley, Kenneth Boas. Randomized and stochastic linear, discrete differential games with quadratic payoff functions. Ph.D. Department of Engineering, University of California at Los Angeles, Los Angeles, 1969.

7. Bryson, Arthur E., Jr., and Yu-Chi Ho. Applied optimal control. Waltham, Mass., Blaisdell, 1969.

8. Deutsch, Ralph. Estimation theory. Englewood Cliffs, N.J., Prentice-Hall, 1965.

9. Gunckel, T.F., and G.F. Franklin. "A general solution for linear, sampled-data control." Transactions of the ASME, Journal of Basic Engineering, Vol. 85D, No. 2, pp. 197-203, 1963.

10. Hestenes, M.R. "Numerical methods of obtaining solutions of fixed end point problems in the calculus of variations." RM-102, The Rand Corporation, August, 1949.

11. Ho, Y.C., A.E. Bryson, Jr., and S. Baron. "Differential games and optimal pursuit-evasion strategies." IEEE Transactions on Automatic Control, Vol. AC-10, No. 4, pp. 385-389, October, 1965.

12. Ho, Y.C. "Optimal terminal maneuver and evasion strategy." J. SIAM Control, Vol. 4, No. 3, pp. 421-428, August 1966.

13. Isaacs, Rufus. Differential Games. New York, Wiley, 1965.

14. Johansen, D.E. "Optimal control of linear stochastic systems with complexity constraints." Advances in Control Systems, C.T. Leondes (ed.), Vol. 4, New York, 1966, pp. 181-278.

15. Joseph, P.D., and J.T. Tou. "On linear control theory." Transactions of the AIEE, Part II, pp. 193-196, September, 1961.

16. Kalaba, R.E. "On nonlinear differential equations, the maximum operation, and monotone convergence." Journal of Mathematics and Mechanics, Vol. 8, No. 4, pp. 519-574, July, 1959.

17. Kalman, R.E., and R.S. Bucy. "New results in linear filtering and prediction." Transactions of the ASME, Journal of Basic Engineering, Vol. 83D, pp. 95-108, 1961.

18. Kalman, R.E. "A new approach to linear filtering and prediction problems." Transactions of the ASME, Journal of Basic Engineering, Vol. 82D, pp. 35-45, 1960.

19. Kelley, H.J. "A second variation test for singular extremals." AIAA Journal, Vol. 2, No. 8, pp. 1380-1382, August, 1964.

20. Kenneth, P. and R. McGill. "Two-point boundary-value-problem techniques." Advances in Control Systems, C.T. Leondes (ed.), Vol. 3, New York, 1966, pp. 69-109.

21. Porter, William A. Modern foundations of systems engineering. New York, Macmillan, 1966.

22. Porter, W.A. "On function space pursuit-evasion games." J. SIAM Control, Vol. 5, No. 4, pp. 555-574, November, 1967.

23. Rhodes, I.B., and D.G. Luenberger. "Differential games with imperfect state information." IEEE Transactions on Automatic Control, Vol. AC-14, No. 1, pp. 29-38, February, 1969.

24. Rhodes, I.B., and D.G. Luenberger. "Stochastic differential games with constrained state estimators." IEEE Transactions on Automatic Control, Vol. AC-14, No. 5, pp. 476-481, October 1969.

25. Speyer, J.L. "A stochastic differential game with controllable statistical parameters." IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-3, No. 1, pp. 17-20, June, 1967.

26. Striebel, C. "Sufficient statistics in the optimum control of stochastic systems." J. Math. Analysis and Applications, Vol. 12, pp. 576-592, 1965.

27. Willman, W.W. "On a class of stochastic pursuit-evasion games." Harvard University Technical Report, TR 585, April, 1969.

28. Willman, W.W. "Formal solutions for a class of stochastic pursuit-evasion games." IEEE Transactions on Automatic Control, Vol. AC-14, No. 5, pp. 504-509, October, 1969.

29. Witsenhausen, H.S. "A counterexample in stochastic optimum control." J. SIAM Control, Vol. 6, No. 1, pp. 131-147, February, 1968.

APPENDIX

The general linear-quadratic-Gaussian stochastic differential game functional is given in terms of L_1, L_2, K_1, and K_2 by (2.78) which is

$$J(L_1,L_2,K_1,K_2) = \frac{1}{2} E\Big\{\Big\langle \phi X - T_1K_1\big[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\big] + T_2K_2\big[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\big],\ \phi X - T_1K_1\cdot \big[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\big] + T_2K_2\big[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\big]\Big\rangle + \Big\langle K_1\big[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\big],\ K_1\big[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\big]\Big\rangle - \Big\langle K_2\big[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\big],\ K_2\big[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\big]\Big\rangle\Big\} \tag{A.1}$$

Differentiating the payoff functional with respect to L_1 and L_2 and setting the results equal to zero, we have

$$0 = \frac{\partial J}{\partial L_1} = -K_1^*T_1^*\Big[\phi X - T_1K_1\big[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\big] + T_2K_2\big[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\big]\Big] + K_1^*K_1\big[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\big] \tag{A.2}$$

$$0 = \frac{\partial J}{\partial L_2} = K_2^*T_2^*\Big[\phi X - T_1K_1\big[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\big] + T_2K_2\big[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\big]\Big] - K_2^*K_2\big[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\big] \tag{A.3}$$

These equations will be satisfied if

$$0 = -T_1^*\left[\phi X - T_1K_1\left[\hat{X}_1 + L_1(Z_1 - \hat{Z}_1)\right] + T_2K_2\left[\hat{X}_2 + L_2(Z_2 - \hat{Z}_2)\right]\right]$$

$$+ K_1\left[\hat{X}_1 + L_1(Z_1 - \hat{Z}_1)\right] \qquad (A.4)$$

and

$$0 = T_2^*\left[\phi X - T_1K_1\left[\hat{X}_1 + L_1(Z_1 - \hat{Z}_1)\right] + T_2K_2\left[\hat{X}_2 + L_2(Z_2 - \hat{Z}_2)\right]\right]$$

$$- K_2\left[\hat{X}_2 + L_2(Z_2 - \hat{Z}_2)\right] \qquad (A.5)$$

We interpret these equations in the usual manner; i.e., the right side of (A.4) is orthogonal to any linear trancformation of $Z_1 - \hat{Z}_2$, and the right side of (A.5) is orthogonal to any linear transformation of $Z_2 - \hat{Z}_2$. We define the linear transformations

$$M = \phi - T_1K_1 + T_2K_2 \qquad (A.6)$$

$$A_1 = (T_1^*T_1 + I)K_1 \qquad (A.7)$$

$$A_2 = (T_2^*T_2 - I)K_2 \qquad (A.8)$$

and note that we may write $\hat{X}_1 = X - \epsilon_1$, $\hat{X}_2 = X - \epsilon_2$. Thus (A.4) may be written

$$0 = -(T_1^*M - K_1)X - A_1\left[\epsilon_1 - L_1(Z_1 - \hat{Z}_1)\right]$$

$$+ T_1^*T_2K_2\left[\epsilon_2 - L_2(Z_2 - \hat{Z}_2)\right] \qquad (A.9)$$

and (A.5) becomes

$$0 = (T_2^*M-K_2)\,X + T_2^*T_1K_1\left[\epsilon_1-L_1(Z_1-\hat{Z}_1)\right]$$

$$-\Lambda_2\left[\epsilon_2-L_2(Z_2-\hat{Z}_2)\right] \tag{A.10}$$

We may also differentiate the payoff functional with respect to K_1 and K_2. Doing this and setting the resulting expression equal to zero, we have

$$0 = \frac{\partial J}{\partial K_1} = -T_1^*\left[\phi X-T_1K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right]\right.$$

$$\left.+T_2K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right]\right] + K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right] \tag{A.11}$$

$$0 = \frac{\partial J}{\partial K_2} = T_2^*\left[\phi X-T_1K_1\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right]\right.$$

$$\left.+T_2K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right]\right] - K_2\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right] \tag{A.12}$$

Again we interpret these equations to mean that the right side of (A.11) is orthogonal to any linear transformation of $\left[\hat{X}_1+L_1(Z_1-\hat{Z}_1)\right]$ and the right side of (A.12) is orthogonal to any linear transformation of $\left[\hat{X}_2+L_2(Z_2-\hat{Z}_2)\right]$. We note that (A.11) and (A.12) have the same form as (A.4) and (A.5) and may be written as (A.9) and (A.10). Thus the right side of (A.9) is orthogonal to any linear transformation of $Z_1 - \hat{Z}_1$ or of $\left[\hat{X}_1-L_1(Z_1-\hat{Z}_1)\right]$, and these two relations imply that the right-hand side of (A.9) is orthogonal to any linear transformation of $\hat{X}_1$. Analogous statements apply to (A.10): the right side of (A.10) is orthogonal to any linear transformations of $\hat{X}_2$ or of $Z_2 - \hat{Z}_2$. Because of the fact that for normal random variables the error in the conditional mean state estimate is orthogonal to all linear transformations of the conditional mean and of the observations, we may rewrite (A.9) and (A.10) as

$$0 = (T_1^* M - K_1)\hat{X}_1 + \Lambda_1 L_1 (Z_1 - \hat{Z}_1) + T_1^* T_2 K_2 \left[Z_2 - \hat{Z}_2) \right] - \Lambda_1 \epsilon_1 \quad \text{(A.13)}$$

$$0 = (T_2^* M - K_2)\hat{X}_2 + T_2^* T_1 K_1 \left[\epsilon_1 - L_1 (Z_1 - \hat{Z}_1) \right] + \Lambda_2 L_2 (Z_2 - \hat{Z}_2) - \Lambda_2 \epsilon_2 \quad \text{(A.14)}$$

Again, we recall that the right side of (A.13) is orthogonal to all linear transformations of $\hat{X}_1$ or of $Z_1 - \hat{Z}_1$ and that the right side of (A.14) is orthogonal to all linear transformations of $\hat{X}_2$ and of $Z_2 - \hat{Z}_2$. This is true for the particular transformation of $\hat{X}_1$: $(T_1^* M + K_1)\hat{X}_1$. Thus, from (A.13)

$$\begin{aligned} 0 = & -\left\langle (T_1^* M - K_1)\hat{X}_1,\ (T_1^* M - K_1)\hat{X}_1 \right\rangle + \left\langle \Lambda_1 L_1 (Z_1 - \hat{Z}_1),\ (T_1^* M - K_1)\hat{X}_1 \right\rangle \\ & + \left\langle T_1^* T_2 K_2 \left[\epsilon_2 - L_2 (Z_2 - \hat{Z}_2) \right],\ (T_1^* M - K_1)\hat{X}_1 \right\rangle \end{aligned} \quad \text{(A.15)}$$

It is also true for the transformation of $Z_1 - \hat{Z}_1$: $\Lambda_1 L_1 (Z_1 - \hat{Z}_1)$. Therefore,

$$\begin{aligned} 0 = & -\left\langle (T_1^* M - K_1)\hat{X}_1,\ \Lambda_1 L_1 (Z_1 - \hat{Z}_1) \right\rangle + \left\langle \Lambda_1 L_1 (Z_1 - \hat{Z}_1),\ \Lambda_1 L_1 (Z_1 - \hat{Z}_1) \right\rangle \\ & + \left\langle T_1^* T_2 K_2 \left[\epsilon_2 - L_2 (Z_2 - \hat{Z}_2) \right],\ \Lambda_1 L_1 (Z_1 - \hat{Z}_1) \right\rangle \end{aligned} \quad \text{(A.16)}$$

Adding (A.15) and (A.16), we have

$$0 = -\left\langle (T_1^* M - K_1)\hat{X}_1,\ (T_1^* M - K_1)\hat{X}_1 \right\rangle + \left\langle \Lambda_1 L_1(Z_1 - \hat{Z}_1),\ \Lambda_1 L_1(Z_1 - \hat{Z}_1) \right\rangle$$
$$+ \left\langle T_1^* T_2 K_2 \left[\epsilon_2 - L_2(Z_2 - \hat{Z}_2)\right],\ \Lambda_1 L_1(Z_1 - \hat{Z}_1) + (T_1^* M - K_1)\hat{X}_1 \right\rangle \tag{A.17}$$

Now the second and third terms on the right depend only on the covariances of the noise and the initial state, while the first term on the right is also dependent on the mean of the initial state. Thus, for (A.17) to be satisfied for all values of the initial state, we must have

$$T_1^* M - K_1 = 0 \tag{A.18}$$

This being true, equation (A.13) reduces to

$$0 = \Lambda_1 L_1(Z_1 - \hat{Z}_1) + T_1^* T_2 K_2 \left[\epsilon_2 - L_2(Z_2 - \hat{Z}_2)\right] \tag{A.19}$$

Furthermore, (A.18), (A.6), and (A.7) lead to an alternate representation of Λ_1

$$\Lambda_1 = T_1^* \left[\phi + T_2 K_2\right] \tag{A.20}$$

so that (A.19) may be written

$$0 = T_1^* \left[\phi + T_2 K_2\right] L_1(Z_1 - \hat{Z}_1) + T_1^* T_2 K_2 \left[\epsilon_2 - L_2(Z_2 - \hat{Z}_2)\right] \tag{A.21}$$

which will be satisfied if

$$0 = \left[\phi+T_2K_2\right]L_1(Z_1-\hat{Z}_1) + T_2K_2\left[\epsilon_2-L_2(Z_2-\hat{Z}_2)\right] \qquad (A.22)$$

Since (A.27) is interpreted to mean that the right side of (A.27) is orthogonal to any linear transformation of $Z_1 - \hat{Z}_1$, we must have

$$\left[\phi+T_2K_2\right]L_1\psi_{Z_1Z_1} - T_2K_2L_2\psi_{Z_2Z_1} = - T_2K_2\psi_{\epsilon_2Z_1} \qquad (A.23)$$

where $\psi_{Z_1Z_1} = E\left\{(Z_1-\hat{Z}_1)(Z_1-\hat{Z}_1)^*\right\}$, etc.

A completely analogous manipulation starting with (A.14) leads us first to the conclusion that

$$T_2^*M - K_2 = 0 \qquad (A.24)$$

This then reduces (A.14) to

$$0 = T_2T_1K_1\left[\epsilon_1-L_1(Z_1-\hat{Z}_1)\right] + \Lambda_2L_2(Z_2-\hat{Z}_2) - \Lambda_2\epsilon_2 \qquad (A.25)$$

Then (A.24), (A.6), and (A.8) give an alternative expression for Λ_2

$$\Lambda_2 = - T_2^*\left[\phi-T_1K_1\right] \qquad (A.26)$$

so that (A.25) may be written

$$0 = T_2^*T_1K_1\left[\epsilon_1-L_1(Z_1-\hat{Z}_1)\right] - T_2^*\left[\phi-T_1K_1\right]L_2(Z_2-\hat{Z}_2)$$

$$+ T_2^*\left[\phi-T_1K_1\right]\epsilon_2 \qquad (A.27)$$

which will be satisfied if

$$0 = T_1K_1\left[\epsilon_1-L_1(Z_1-\hat{Z}_1)\right] - \left[\phi-T_1K_1\right]L_2(Z_2-\hat{Z}_2)$$

$$+ \left[\phi-T_1K_1\right]\epsilon_2 \tag{A.28}$$

As before, we interpret this to mean that the right side of (A.28) is orthogonal to any linear transformation of $Z_2 - \hat{Z}_2$; hence,

$$T_1K_1\psi_{Z_1Z_2} + \left[\phi-T_1K_1\right]L_2\psi_{Z_2Z_2} = T_1K_1\psi_{\epsilon_1Z_2} \tag{A.29}$$

We have in (A.18), (A.23), (A.24), and (A.29) a set of four simultaneous linear equations describing K_1, K_2, L_1, and L_2. We may solve for K_1 and K_2 quite easily from (A.18) and (A.24). Using (A.6) and (A.18), we have

$$T_1^*\phi - T_1^*T_1K_1 + T_1^*T_2K_2 = K_1 \tag{A.30}$$

Comparing these equations to (1.44) and (1.45), we see they are similar in form; thus, we have the solutions

$$K_1 = T_1^*\left[I + T_1T_1^* - T_2T_2^*\right]^{-1}\phi \tag{A.31}$$

$$K_2 = T_2^*\left[I + T_1T_1^* - T_2T_2^*\right]^{-1}\phi \tag{A.32}$$

These expressions may be substituted into (A.23) and (A.29), but for notational compactness it is better to retain the equations in their present form, which is

$$\left[\phi+T_2K_2\right]L_1\psi_{z_1z_1} - T_2K_2L_2\psi_{z_2z_2} = -T_2K_2\psi_{\epsilon_2z_2} \qquad (A.23)$$

$$T_1K_1L_1\psi_{z_1z_2} + \left[\phi-T_1K_1\right]L_2\psi_{z_2z_2} = T_1K_1\psi_{\epsilon_1z_2} \qquad (A.29)$$

The above equations are necessary conditions which must be satisfied by linear operations on noisy state observations which make up part of the strategies assumed in (2.38) and (2.39).

END