SENSITIVITY ANALYSIS OF OPTIMAL LINEAR RANDOM PARAMETER SYSTEMS

by

Prashant Parikh

S.B., Massachusetts Institute of Technology
(1979)

SUBMITTED IN PARTIAL FULFILMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May, 1979

Signature of Author ....
Department of Electrical Engineering and Computer Science, May 1979.

Certified by ....
Thesis Supervisor

Accepted by ....
Chairman, Departmental Committee on Graduate Students

SENSITIVITY ANALYSIS OF OPTIMAL LINEAR RANDOM PARAMETER SYSTEMS

by

Prashant Parikh

Submitted to the Department of Electrical Engineering and Computer Science on May 15, 1979 in partial fulfilment of the requirements for the Degree of Master of Science.

ABSTRACT

This report involves the application of ideas in adaptive stochastic control to economics.

We investigate the control problem for a linear, multivariable, dynamic system with purely random (i.e. white) parameters. The quadratic cost criterion is formulated to make the problem a tracking problem. Since the parameters are modelled as white stochastic processes, there is no posterior learning and no dual effect. The certainty-equivalence principle does not hold. We find that the extension of the "Uncertainty Threshold Principle" from scalar systems to multidimensional ones turns out to be analytically intractable.

Next, we derive sensitivity equations for the above optimal system to study the effects of small variations in parameter uncertainties on the optimal performance of the system. These equations enable us to rank parameters in order of the sensitivity of the performance to variations in their variances. This makes it possible to locate the "pressure" points in a model, if any exist.

We then convert an economic policy problem into a stochastic optimal control tracking problem and analyse it with the equations we have derived. We study the different elements that enter into a tracking problem and then discuss the empirical results obtained from the sensitivity equations. The model we choose for the analysis turns out to be insensitive to variations in parameter variances which makes it reasonably reliable. We also analyse in detail the structure of the model and the inter-dependences of the state and control variables.

General purpose computer programs are included in one of the appendices.

THESIS SUPERVISOR : Michael Athans

TITLE : Professor of Electrical Engineering and Computer Science

ACKNOWLEDGEMENTS

I would like to thank my thesis supervisor Professor Michael Athans for his continued help and guidance in my research. Professor Athans always had an inspiring combination of insight and perspective in his supervision of my work. He has made a significant contribution to my graduate educational experience.

I would also like to thank Dr. David Castanon and Dr. Richard Ku for their useful comments and suggestions. I am deeply grateful to my Graduate Counsellor Professor Sanjoy Mitter for his understanding and concern for my overall graduate program.

The computer work was done at the MIT Information Processing Services on an IBM 370/168. Many thanks are due to Mr. Jim Carrig for his help in writing the computer programs. Thanks are also due to Mr. Abhyankar for an excellent job of typing the thesis and to Mr. Mevlankar for the art work.

This research was carried out at the MIT Laboratory for Information and Decision Sciences. The theoretical part of the research was supported in part by grant AF/AFOSR 77-3321 B; the economics related research was supported in part by grant NSF/SOC 76-05837.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Adaptive Stochastic Control :

Though research in stochastic control has progressed in the last decade, there does not exist at present a general, computationally viable theory of optimal stochastic control. Richard Ku, in his doctoral thesis [1], gives a survey of this area. Bellman [2] first introduced the concepts of 'information pattern' and 'learning'. Feldbaum [3] expanded on this in his celebrated four part paper on the theory of dual control, in which he identified the two distinct roles an optimal controller must play to be truly optimal. The controller must actively try to identify the unknown parameters of the system and simultaneously try to control the system. He showed that in such dual control systems there may exist an inherent conflict between applying the inputs for learning and for effective control purposes. This introduced the concepts of caution and probing and the possible trade-off between them. For some insight, the reader might want to refer to a paper by Sternby [4], in which he solves a simple dual control problem analytically and compares the optimal solution with other suboptimal strategies.

Bar-Shalom and Tse have further clarified the concept of dual control and various related concepts like separation, certainty-equivalence, neutrality and have also made precise the subtle differences between closed-loop optimal policies and feedback optimal policies arising from different information patterns. These can be found in [5] - [9]. On the last point there is an excellent paper by Dreyfus [10].

Since the permissible controls are causal, the only information about future observations that can be used by the controller is the probability distribution of these future observations. This knowledge is what makes the difference between a feedback control policy and a closed-loop control policy. It is only the latter policy that uses this information to advantage. The feedback law at time t uses information only upto time t. And it is this difference that makes the dual effect possible. A control is said to have a dual effect when, in addition to its effect on the state of the system, it is able to affect the uncertainty of the state of the system. If the control cannot affect this uncertainty, then the system is called neutral. If the dual effect is present, then the control can help to improve the future estimation and in so doing facilitate the task of the control. In this case the control is said to be actively adaptive. Precise definitions of these terms can be found in the references cited above.

It turns out, however, that we cannot solve the adaptive control problem except for special cases. In fact, the decision problem in linear systems with unknown parameters is actually a nonlinear stochastic control problem [7], [47]. There are two ways in which we can make approximations to make the original problem mathematically tractable. One is to approximate the optimal law. The second is to approximate the linear system as having random parameters that are uncorrelated in time, or white, in engineering jargon, and to obtain the optimal control for this approximate system which may now be possible analytically. This is the route we shall take in this report. We shall find that our assumption of white parameters makes identification impossible which means there is no probing action thereby making the problem solvable.

Before we turn to a mathematical description of the problem, let us first survey the interactions of control theory and economics, as we shall be applying our techniques to an economic policy problem.

1.2 Control Theory and Economics :

In recent years, several workers have begun to find the techniques of optimal control theory to be useful to the analysis of economic problems. Some of the basic concepts of system theory and, in particular, of stochastic optimal control theory may be able to provide a more unified and comprehensive analytical framework for posing and solving economic problems. Kendrick [12], Athans and Kendrick [13], and Aoki [14] have written good survey articles with extensive bibliographies on the different areas of interaction between economists and control theorists. The earliest instances of such intercourse began to appear in the 1950's with the work of Tustin [15], Phillips [16], Theil [17] and Simon [18]. After this, there seems to have been a total absence of dialogue until the 1970's. This decade has seen, however, an encouragingly large number of interactions. Aoki, Chow, Kendrick and Pindyck, amongst others, seem to have been the more prominent contributors, [19] - [38]. Though there is still a debate about the degree and kind of applicability of control theoretic ideas and methods, it is significant that the debate does not question any more the fact of the basic usefulness of control theory to economics. One cannot emphasize enough, however, the need for control theorists to thoroughly understand the economics they wish to apply themselves to. Also, economists would do well to appreciate the different tools developed in control theory together with the limitations of these tools.

The applications of control theory have been in different areas of economics : various microeconomic problems and macroeconomic stabilization and regulation problems. Examples of microeconomic applications are profit maximization in a firm, optimal advertizing levels, analysis of commodity markets, optimal price setting in the face of uncertain consumer response, and others, all in a more general dynamic setting. The reader can find references in the survey articles cited above and in [38].

A natural area for control applications is the analysis of macroeconomic policy planning problems. Economic policymakers are interested in controlling the national economy with the various instruments they have at their disposal. The economy is, firstly, a dynamic entity, in which present policy action affects not only the present but also the future course of events. Secondly, it is essentially a stochastic entity as well, so that some way of incorporating uncertainty at a basic level is needed. This makes the regulation of the economy a natural stochastic control problem.

A number of questions arise in the evaluation of the performance of the economy under different specifications of the policy instruments. First of all, we need to specify goals in terms of which this performance can be evaluated. Once we have succeeded in formulating clearly our objectives, how do we look for good policies? In general, one might expect a good policy to coordinate all the available instruments in some suitable way. How do we compare different "good" policies? Is there an unique optimal policy? Many other related questions can be asked. Optimal control seems to offer a natural, precise framework for addressing such questions.

Another point, in a slightly different vein, needs to be made here. System theory can make a far more basic contribution as well. Much conventional economics is done in a sociopolitical vacuum from which all traces of conflict, compromise, imbalances of power, human factors in policymaking and other so-called imperfections have been conveniently removed. If one is to adopt a realistic approach to real problems, then a more comprehensive viewpoint at a fundamental level is needed, and to the extent that science can illuminate our understanding of human "systems", system theory has the potential to incorporate a larger view. (This, of course, is not to ratify the argot in the pseudosciences of "General Systems Theory" [39] or "System Dynamics" [40].)

Economists and control theorists approach their models with different attitudes and this has, to some degree, influenced the tools they use. In economics, many aspects of the models are rather arbitrary since the sheer complexity of real economic phenomena force model builders to adopt many simplifying and often unrealistic assumptions for reasons not entirely justifiable on economic considerations alone. This is in addition to the fact that economic theory today does not as yet have a really fundamental grasp of economic phenomena. Conscious of this arbitrariness to some extent, economists do not take their models literally and are generally content with establishing qualitative properties of their models such as existence of optimal decision rules and properties of classes of optimal decision rules such as stationarity and stability. Time has played a relatively minor role in these models, though recent economics has considered it more adequately.

Engineers, on the other hand, do have a better and deeper understanding of the engineering systems they model, relatively speaking, and so tend to trust their models to a far greater degree. They generally analyse their systems in detailed quantitative terms, and construct and implement algorithms for optimal decision rules, in addition to studying the qualitative features of their systems. Most models do take into account the dynamics of the system.

The focal point of the interaction here has been the traditional macroeconometric model which, after suitable transformation, can be recast into the state-space representation familiar to engineers. Economists usually assume that the main state variables can be measured exactly. Also, they emphasize the estimation of unknown parameters. Engineers, on the other hand, usually take parameters as given and deal with observation errors instead. In [31], Kendrick observes that the data used by policy analysts to determine monetary and fiscal policies are known to contain errors. Such data are being constantly revised as more information becomes available. The magnitude of these revisions gives us a measure of the relative quality of different macroeconomic time series. However, economists do not at present use this new information in determining policies. Fair [11] points out that the accuracy of the model is generally improved when the actual values of the exogenous variables are used and when more recent coefficient estimates are used. From the engineering side, adaptive control algorithms that look impossible in an aerospace context may be perfectly practical when decision rules have to be computed only once a month or once every quarter.

Differences of this kind in attitude and approach help to underscore, in fact, the common thread that binds both fields : the making of decisions with imperfect information in an uncertain environment. Adaptive stochastic control seeks to tackle this basic question. Let us turn now to a mathematical formulation of the problem.

1.3 The Problem :

We shall study the following linear, multivariable, discrete-time system :

$$x_{t+1} = A_t x_t + B_t u_t + c_t \tag{1.3.1}$$

where A_t, B_t are white, Gaussian matrices and c_t is a white, Gaussian vector. Note that the noise in this system enters both additively, through c_t, and multiplicatively through A_t and B_t. Note also that all the random quantities are white. This is a crucial assumption in that it makes active learning impossible since, at each time instant, the values of A, B and C are all uncorrelated with the past. However, this assumption does enable us to deal analytically with uncertain parameters, representing in some sense a worst case situation. The assumption of a Gaussian distribution is actually superfluous. All we need to know are the first and second order statistics. The actual probability distribution does not matter.

This formulation holds a double interest. Firstly, its solution is of basic theoretical interest. An analysis of this problem can be found in [1], [41], [42], [43]. This system forms the basis of the result embodied in the "Uncertainty Threshold Principle" expounded in [1], [44], [45], [46]. The second point of this formulation is that its assumptions fit the framework of linear econometric models reasonably well. The estimated parameters of econometric models are actually random variables. The use of white processes, of course, may not be quite realistic, though this assumption makes the problem amenable to mathematical solution, and in addition represents a worst case situation which may yield useful information for further analysis.

The central result of Ku's thesis [1] that is of relevance to us is embodied in what is called the "Uncertainty Threshold Principle". It arises from an analysis of the following scalar stochastic control problem :

$$x_{t+1} = a_t x_t + b_t u_t + \xi_t ; \qquad x_o \text{ given} \tag{1.3.2}$$

where x_t is the scalar state of the first order system. We assume that the driving term ξ_t is a zero-mean Gaussian white noise with known variance Ξ. We also assume that the random parameters a_t and b_t are Gaussian and white with known means $\bar{a}$, $\bar{b}$, known variances Σ_{aa}, Σ_{bb}, and known cross-covariance Σ_{ab}. We also have perfect state information.

The optimal control problem is to find a feedback control law $u_t = \gamma(x_t,t)$, $t = 0,1,2, \ldots , N-1$, such that the expected value of the following quadratic cost functional is minimized.

$$J = E\{Qx_N^2 + \sum_{t=0}^{N-1} (Qx_t^2 + Ru_t^2)\}, \quad F,Q \geq 0, \quad R > 0 \tag{1.3.3}$$

The expectation is taken with respect to the probability distribution of the underlying random variables a_t, b_t, ξ_t.

The solution to this problem is readily obtained by applying the standard stochastic dynamic programming algorithm. We get the following equations :

$$u_t^* = -G_t x_t \tag{1.3.4}$$

$$G_t = \frac{K_{t+1}(\Sigma_{ab} + \bar{a}\bar{b})}{R + (\Sigma_{bb} + \bar{b}^2) K_{t+1}} \tag{1.3.5}$$

$$K_t = Q + (\Sigma_{aa} + \bar{a}^2)K_{t+1} - G_t^2[R + K_{t+1}(\Sigma_{bb} + \bar{b}^2)] \tag{1.3.6}$$

$$K_N = Q \tag{1.3.7}$$

The optimal cost is given by :

$$J^* = K_o x_o^2 + \sum_{\tau=0}^{N-1} K_{\tau+1} \Xi_\tau \tag{1.3.8}$$

We note, in passing, that the control law is linear in the state and the Riccati-like equation satisfied by K_t has a unique solution under the given conditions.

An inspection of the infinite horizon case $(N \to \infty)$ yields an interesting result. Assume that K_{t+1} is "large" in the following equation :

$$K_t = Q + (\Sigma_{aa} + \bar{a}^2) K_{t+1} - \frac{K_{t+1}^2 (\Sigma_{ab} + \bar{a}\bar{b})^2}{R + (\Sigma_{bb} + \bar{b}^2) K_{t+1}}$$

Then the backward in time evolution of K_t is given approximately

$$K_t \simeq K_{t+1} \cdot M$$

$$\text{where } M = \Sigma_{aa} + \bar{a}^2 - \frac{(\Sigma_{ab} + \bar{a}\bar{b})^2}{(\Sigma_{bb} + \bar{b}^2)} \qquad (1.3.9)$$

Clearly, if the threshold parameter $M > 1$, then K_t blows up. In fact, it is possible to prove that the unique positive solution to the above equation exists if and only if $M < 1$. This result, which imposes a fundamental limitation on the infinite horizon problem, is called the Uncertainty Threshold Principle. If $M > 1$, then K_t blows up and therefore the optimal cost J* also blows up. In physical terms, this principle makes the eminently reasonable statement that if one's knowledge about the present and future structure of the system is "very" uncertain, then there is no optimal action that will keep the cost finite for the infinite horizon problem. Though the result has been proved for linear-quadratic systems,it seems reasonable to assume the same qualitative result for general systems too.

1.4 Structure of Report :

In this report we shall pursue two different routes that arise from the random parameter formulation. The first is to extend the above described result to multivariable systems. This turns out to be far more difficult than what it may seem to be on first sight. The equations, though similar in structure, are far more complicated because of the appearance of matrices in all the formulas. The first difficulty one faces is the question of suitably representing the covariance of a matrix and then establishing formulas and equations that are expressed in terms of the means and covariances of the various matrices. We find that it is very difficult, if not impossible, to derive an analytical formula for the threshold in analogy with the scalar case. This part of the work is described in Chapter 2.

The second route is more practically oriented. We know that it is difficult to control large econometric models with many random parameters. If we formulate the policy problem in an optimal control framework, then it would be very useful if we could develop some method by which to rank these parameters in terms of their influence on the performance of the system. This would tell us which, if any, parameters are sensitive and give a clue as to whether better information is needed if we are to trust the model we are using. This kind of study falls under the general rubric of sensitivity analysis. A fair amount of work has already been done in this area, [48] - [63], and this methodology can be readily applied to derive equations for our case. We first derive sensitivity equations for optimal random parameter systems. Next we choose a small econometric model by Abel [47] and apply these equations to the model. We then analyse the results and comment on possible uses for this approach. This is the content of Chapters 3 and 4.

1.5 Contributions of the Report :

1. Derivation and analysis of the solution to the optimal linear - quadratic tracking problem with purely random parameters and additive noise.

2. Sensitivity analysis : development of sensitivity equations for the above system to rank parameters in terms of their influence on the performance of the system.

3. Application of above equations to a simple macroeconomic model of the U.S. economy.

4. Development of general purpose computer programs for the optimal stochastic control of multivariable linear systems with white parameters with respect to quadratic performance criteria, for both regulator and tracking applications.

CHAPTER 2

OPTIMAL LINEAR RANDOM PARAMETER SYSTEMS

2.1 Introduction :

In this chapter, we shall develop and discuss the optimal control problem for linear systems with purely random parameters. We treat the most general case of this formulation : the problem is multivariable and includes additive noise, and is stated as a tracking problem. We also state the 'Uncertainty Threshold Principle' for one-dimensional systems and consider some of the difficulties involved in trying to extend it to multivariable systems. Here we present one way of representing algebraically the solution to the multivariable control problem. Some empirical results are presented to demonstrate the behaviour of such systems. This chapter will try to lay the groundwork and motivation for the next chapter.

In the next section, we state the problem as a multivariable linear - quadratic random parameter tracking problem. In section 3, we present the solution of the problem. Since the actual derivation is slightly long and complicated we choose to present it in Appendix A. In section 4, we discuss the solution of the problem. Next, in section 5, we demonstrate the Uncertainty Threshold Principle developed by Ku [1] for further insight into the problem.

2.2 Problem Statement :

Let us begin by stating the problem. Consider a multivariable stochastic linear dynamical system with state x_t and control u_t described by the following difference equation :

$$\underline{x}_{t+1} = \underline{A}_t \underline{x}_t + \underline{B}_t \underline{u}_t + \underline{c}_t \tag{2.2.1}$$

$\underline{x}_o$ given; $t = 0,1,2, \ldots , N-1$

$\underline{x}_t \in R^n, \ \underline{u}_t \in R^m, \ \underline{A}_t \in R^{n\times n}, \ \underline{B}_t \in R^{n\times m}, \ \underline{c}_t \in R^n$

Henceforth we shall not underscore vectors or matrices for greater clarity of notation. We assume that the additive term c_t driving the system is a vector random process which is white and whose mean vector and covariance matrix are given. That is, we assume that

$$E\{c_t\} = \bar{c} \qquad \forall t$$

$$E\{(c_t - c)(c_\tau - \bar{c})'\} = \Sigma_c \delta_{t\tau} \qquad \delta_{t\tau} = \begin{cases} 1 & \text{if } t = \tau \\ 0 & \text{if } t \neq \tau \end{cases}$$

where Σ_c is an $n \times n$ matrix.

Assume that A_t and B_t are random matrices which are also white with given first and second order statistics. We assume that

$$E\{A_t\} = \bar{A}$$

$$E\{B_t\} = \bar{B}$$

Here we face the issue of how to represent the covariance of a matrix. Just as the covariance of a vector is a matrix, so the covariance of a matrix is a fourth-order tensor. We can, however, express this tensor as a higher dimensional matrix. There are many ways of doing this, an obvious one that comes to mind immediately being the Kronecker product. The manner of representation should evidently be dictated by how we wish to use the covariance. We shall find that, for our purposes, the most suitable representation is obtained by using the simple notion of a stacking operator, that is, an operator that stacks the columns of a matrix into a single vector. Mathematically, if we have a $p \times q$ matrix A whose columns are denoted by a_i i.e.

if $\quad A = (a_1\ a_2\ a_3\ \ldots\ldots\ a_q)$

then $\quad S(A) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_q \end{bmatrix}$

stacks the columns of A into a single vector of length pq.

The definition of covariance now follows quite readily :

$$\text{Cov}\ (A) = E\ \{\ [\ S(A_t) - S(\bar{A})]\ [\ S(A_t) - S(\bar{A})]'\ \}$$

An immediate advantage of this representation vis-à-vis the Kronecker product is that it is symmetric.

To return to our problem, we assume that

$$E\ \{\ [\ S(A_t) - S(\bar{A})\ ]\ [\ S(A_\tau) - S(\bar{A})\ ]'\} = \Sigma_A\ \delta_{t\tau}$$

$$E\ \{\ [\ S(B_t) - S(\bar{B})\ ]\ [\ S(B_\tau) - S(\bar{B})\ ]'\} = \Sigma_B\ \delta_{t\tau}$$

$$E\ \{\ [\ S(B_t) - S(\bar{B})\ ]\ [\ S(A_\tau) - S(\bar{A})\ ]'\} = \Sigma_{BA}\ \delta_{t\tau}$$

We also assume that the following cross-covariances are given :

$$E \{ [S(A_t) - S(\bar{A})] [c_\tau - \bar{c}]' \} = \Sigma_{Ac} \delta_{t\tau}$$

$$E \{ [S(B_t) - S(\bar{B})] [c_\tau - \bar{c}]' \} = \Sigma_{Bc} \delta_{t\tau}$$

All the covariance matrices must, of course, be positive semi-definite. In addition to this, they must also satisfy the constraint that the correlation coefficient for each pair of parameters must lie between -1 and +1. Note that all the given statistics are time-invariant - this is not really a restriction. The generalization to the nonstationary case is immediate. Note also that we have made no assumptions about the actual distributions of the various random parameters.

For any optimal control problem, it is essential to specify the information available for control, that is, the information pattern. Generally, in stochastic control problems, utilizing observations improves the performance over the open loop controls because using measurements on the system allows one to reduce the uncertainty. A causal or non-anticipative control cannot obviously use future observations, but it can, however, use the given a priori information about the future probabilistic behaviour of the system and measurement dynamics, or, in equivalent terms, it can use a probabilistic description of future observations.

For our formulation of the problem, the information pattern is especially simple. The whiteness of each component of noise, multiplicative as well as additive, in the system, makes any learning impossible, and so renders the control law incapable of affecting future uncertainty. The law does, of course, take present uncertainty into account.

We assume perfect state measurements. We also assume that the admissible controls are real-valued and of state feedback type, '$u_t = \gamma(x_t, t)$', such that they depend only on the given a priori information and measurements upto time t.

The optimal control problem, then, is to determine the control sequence '$u_t = \gamma(x_t, t)$, $t = 0,1,2, \ldots, N-1$', that minimizes the following quadratic cost criterion :

$$J = \frac{1}{2} E\left\{ \sum_{t=0}^{N-1} [(x_t - \tilde{x}_t)'Q(x_t - \tilde{x}_t) + (u_t - \tilde{u}_t)'R(u_t - \tilde{u}_t)] + (x_N - \tilde{x}_N)'Q(x_N - \tilde{x}_N) \right\} \tag{2.2.2}$$

where $\{\tilde{x}_t\}$, $\{\tilde{u}_t\}$ are the target state and control sequences respectively. These are, of course, also specified at the beginning of the problem. Thus, the problem is what is called a 'tracking' problem in the literature. Note that the weighting matrices are taken to be constant for simplicity but the generalization to time-varying matrices is quite direct.

We now proceed to solve the problem.

2.3 Problem Solution :

The solution to the optimal control problem stated above can be obtained by applying the method of stochastic dynamic programming. Since the complete derivation is somewhat lengthy, we shall relegate it to

Appendix A and merely state the solution here.

The control law turns out to be a linear state feedback law, as one would expect. The equations are :

$$u_t^* = L_t x_t + m_t \tag{2.3.1}$$

where the gain L_t is given by :

$$L_t = -[R + \overline{B'K_{t+1}B}]^{-1} [\overline{B'K_{t+1}A}] \tag{2.3.2}$$

(We use the notation $\overline{B'K_{t+1}B}$ to denote $E\{B_t'K_{t+1}B_t\}$, etc. See Appendix A)

and where

$$m_t = -[R + \overline{B'K_{t+1}B}]^{-1} [\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t] \tag{2.3.3}$$

The matrix, K_t, in the above equations, satisfies the following Riccati-like difference equation :

$$K_t = Q + \overline{A'K_{t+1}A} + [\overline{B'K_{t+1}A}]' \cdot L_t \tag{2.3.4}$$

with the terminal condition:

$$K_N = Q \tag{2.3.5}$$

The vector, p_t, satisfies the following equation :

$$p_t = -Q\tilde{x}_t + \overline{A'K_{t+1}c} + \bar{A}'p_{t+1} + [\overline{B'K_{t+1}A}]' \cdot m_t \tag{2.3.6}$$

$$p_N = -Q\tilde{x}_N$$

The optimal cost can also be evaluated and turns out to be :

$$J^* = \frac{1}{2} x_o' K_o x_o + p_o' x_o + g_o \tag{2.3.8}$$

The scalar g_o comes from the following difference equation :

$$g_t = \frac{1}{2} \tilde{x}_t' Q \tilde{x}_t + \frac{1}{2} \tilde{u}_t' R \tilde{u}_t + \frac{1}{2} \overline{c' K_{t+1} c} + \bar{c}' p_{t+1}$$

$$+ \frac{1}{2} [\overline{B' K_{t+1} c} + \bar{B}' p_{t+1} - R\tilde{u}_t]' \; m_t + g_{t+1} \tag{2.3.9}$$

$$g_N = \frac{1}{2} \tilde{x}_N' \; Q \; \tilde{x}_N \tag{2.3.10}$$

The state of the optimal system is now given by :

$$x_{t+1} = (A_t + B_t L_t) \; x_t + B_t m_t + c_t \tag{2.3.11}$$

Since x_t is a random variable, so is the control u_t, though the gain L_t and the driving term m_t are deterministic.

Note, however, that our a priori information is in terms of means and covariances of A_t, B_t and c_t, whereas the solution is expressed in terms of certain expectations of A_t, B_t, c_t. We should like, therefore, to represent the solution in terms of the various means and covariances. As these equations are a bit complicated, let us first look to the scalar case for some insight. Let's consider the scalar system :

$$x_{t+1} = a_t x_t + b_t u_t + c_t \tag{2.3.12}$$

where a_t, b_t, c_t are now scalar random processes. The Riccati-like equation for the scalar K_t is :

$$K_t = Q + \overline{a^2 K_{t+1}} + \overline{(abK_{t+1})} \; L_t$$

$$L_t = -(R + \overline{b^2 K_{t+1}})^{-1} \overline{(abK_{t+1})}$$

$$= - \frac{\overline{ab} \cdot K_{t+1}}{R + \overline{b^2} \cdot K_{t+1}} \qquad (2.3.13)$$

Therefore,

$$K_t = Q + \overline{a^2} \cdot K_{t+1} - \frac{(\overline{ab})^2 \; K_{t+1}^2}{R + \overline{b^2} \cdot K_{t+1}}$$

But

$$E\{a^2\} = \Sigma_a + \bar{a}^2$$

$$E\{b^2\} = \Sigma_b + \bar{b}^2$$

$$E\{ab\} = \Sigma_{ba} + \bar{a}\bar{b}$$

Hence

$$K_t = Q + (\Sigma_a + \bar{a}^2)K_{t+1} - \frac{(\Sigma_{ba} + \bar{a}\bar{b})^2 K_{t+1}^2}{R + (\Sigma_b + \bar{b}^2)K_{t+1}} \qquad (2.3.14)$$

So now we see how the covariances and means of the various random parameters directly influence the evolution of K_t. In order to represent the solution to the multivariable case in a similar way we need to make a few definitions.

(a) $e_i \sim$ a vector of appropriate dimensions with all zeroes except for a one in the i-th place.

$$e_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

(b) E_{ij} ∿ a matrix of appropriate dimensions with all zeroes except for a one in the i,j-th place.

$$E_{ij} = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix}$$

(c) P_k ∿ a block matrix with n columns and an appropriate number of rows (usually either n^2 or mn) with blocks of n×n such that the k-th block is the identity I_n, and the rest are zeroes. Here 'n' refers to the number of states and 'm' to the number of controls. This is a generalization of e_i.

$$P_k = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \cdots & \cdots \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(d) $\Sigma_A^{k\ell}$ ∿ the (k,ℓ)-th block of size n×n in covariance matrix Σ_A. A similar definition holds for cross-covariance matrices too.

(e) $\Sigma_{A_{ij}}^{k\ell}$ ∿ the (i,j)-th element of the (k,ℓ)-th block of Σ_A i.e.

$$\Sigma_{A_{ij}}^{k\ell} = E\,[\,(a_{ik} - \bar{a}_{ik})\,(a_{j\ell} - \bar{a}_{j\ell})\,]$$

Note that, from the above definitions, we have,

$$\Sigma_A^{k\ell} = P_k'\,\Sigma_A P_\ell$$

We now have the following representation :

$$\overline{A'KA} = E\{ A'KA \}$$

$$= \sum_{k=1}^{n} \sum_{\ell=1}^{n} \mathrm{tr}(KP_k'\Sigma_A P_\ell)\, E_{k\ell} + \bar{A}'K\bar{A} \qquad (2.3.15)$$

Proof :

$$\overline{A'KA} = E\{ A'KA \}$$

$$= E\{ \sum_{k\ell} (A'KA)_{k\ell} E_{k\ell} \}$$ where $(A'KA)_{k\ell}$ is the (k,ℓ)-th element of $(A'KA)$

$$= \sum_{k\ell} E[(A'KA)_{k\ell}]\, E_{k\ell}$$

But

$$E[(A'KA)_{k\ell}] = E[a_k' Ka_\ell]$$

where a_k, a_ℓ are the k-th, ℓ-th columns of A respectively.

$$E[a_k' Ka_\ell] = E[\sum_{i,j} a_{ik} K_{ij} a_{j\ell}]$$

$$= \sum_{i,j} K_{ij}\, E(a_{ik}a_{j\ell})$$

$$= \sum_{i,j} K_{ij} (\Sigma_{A_{ij}}^{k\ell} + \bar{a}_{ik}\bar{a}_{j\ell})$$

$$= \sum_{i,j} K_{ij} \Sigma_{A_{ij}}^{k\ell} + \sum_{i,j} \bar{a}_{ik}K_{ij}\bar{a}_{j\ell}$$

$$= \mathrm{tr}\; K\, \Sigma_A^{k\ell} + \bar{a}_k' K\bar{a}_\ell \qquad \text{(since K is symmetric)}$$

$$= \mathrm{tr}\,(K\, P_k'\, \Sigma_A\, P_\ell) + \bar{a}_k' K\bar{a}_\ell$$

Therefore,

$$\overline{A'KA} = \sum_{k,\ell} tr\,(KP_k'\Sigma_A P_\ell)\; E_{k\ell} + \sum_{k,\ell} (\bar{A}'K\bar{A})_{k\ell}E_{k\ell}$$

$$= \sum_{k=1}^{n}\sum_{\ell=1}^{n} tr\,(KP_k'\Sigma_A P_\ell)\; E_\ell + \bar{A}'K\bar{A} \qquad \text{as required}$$

The same expansion holds obviously for the other cases as well.

Thus, we can rewrite the solution to our optimal control problem in the following way :

$$u_t^* = L_t x_t + m_t \tag{2.3.16}$$

$$L_t = -\left[R + \sum_{k=1}^{m}\sum_{\ell=1}^{m} tr\,(K_{t+1}P_k'\Sigma_B P_\ell)\; E_{k\ell} + \bar{B}'K_{t+1}\bar{B} \right]^{-1}$$

$$\left[\sum_{k=1}^{m}\sum_{\ell=1}^{m} tr\,(K_{t+1}P_k'\Sigma_{BA}P_\ell)\; E_{k\ell} + \bar{B}'K_{t+1}\bar{A} \right] \tag{2.3.17}$$

$$m_t = -\left[R + \sum_{k=1}^{m}\sum_{\ell=1}^{m} tr\,(K_{t+1}P_k'\,\Sigma_B P_\ell)\; E_{k\ell} + \bar{B}'K_{t+1}\bar{B}\right]^{-1}$$

$$\left[\sum_{k=1}^{m} tr\,(K_{t+1}P_k'\Sigma_{Bc})e_k + \bar{B}'K_{t+1}\bar{c} + \bar{B}'p_{t+1} - R\tilde{u}_t \right] \tag{2.3.18}$$

$$K_t = Q + \left[\sum_{k=1}^{n}\sum_{\ell=1}^{n} tr\,(K_{t+1}P_k'\Sigma_A P_\ell)\; E_{k\ell} + \bar{A}'K_{t+1}\bar{A} \right] +$$

$$\left[\sum_{k=1}^{m}\sum_{\ell=1}^{m} tr(K_{t+1}P_k'\Sigma_{BA}P_\ell)\; E_{k\ell} + \bar{B}'K_{t+1}\bar{A}\right]'.\, L_t \tag{2.3.19}$$

$$p_t = -Q\tilde{x}_t + \sum_{k=1}^{n} tr\,(K_{t+1}P_k'\Sigma_{Ac})e_k + \bar{A}'K_{t+1}\bar{c} + \bar{A}'p_{t+1} +$$

$$\left[\sum_{k=1}^{m}\sum_{\ell=1}^{m} tr\,(K_{t+1}P_k'\Sigma_{BA}P_\ell)\; E_{k\ell} + \bar{B}'K_{t+1}\bar{A}\right]'.\, m_t \tag{2.3.20}$$

$$g_t = \frac{1}{2}\tilde{x}_t' Q \tilde{x}_t + \frac{1}{2}\tilde{u}_t' R \tilde{u}_t + \frac{1}{2}[\, tr\,(K_{t+1}\Sigma_c) + \bar{c}' K_{t+1}\bar{c}\,] + \bar{c}' p_{t+1} + \frac{1}{2}[\sum_{k=1}^{m} tr\,(K_{t+1}P_k'\,\Sigma_{Bc})\, e_k + \bar{B}' K_{t+1}\bar{c}]'\, m_t + \frac{1}{2}[\,\bar{B}' p_{t+1} - R\tilde{u}_t\,]'\, m_t + g_{t+1} \quad (2.3.21)$$

$$K_N = Q \quad (2.3.22)$$

$$p_N = -\,Q\tilde{x}_N \quad (2.3.23)$$

$$g_N = \frac{1}{2}\tilde{x}_N' Q \tilde{x}_N \quad (2.3.24)$$

$$J^* = \frac{1}{2}x_o' K_o x_o + p_o' x_o + g_o \quad (2.3.25)$$

2.4 Comments :

Let us briefly note some of the salient features of the solution. Figure 2.1 shows the overall structure of the optimal feedback system. Since $u_t^* = L_t x_t + m_t$, the optimal controller is a linear and time-varying transformation of the state. This is so even if the system is stationary and the cost-functional is time-invariant.

The driving term 'm_t' in the control performs the function of neutralizing the mean of additive noise term c_t, whereas the gain L_t does the actual steering of the system, as can be seen by the fact that L_t is independent of c_t. Looking at L_t more closely, we see that when B_t is more uncertain, the controller is more cautious, as it should be since the control u_t affects the state x_t through B_t. If there is, on the other hand, a high correlation between A_t and B_t, then the control is more active since it can better regulate the system. This is

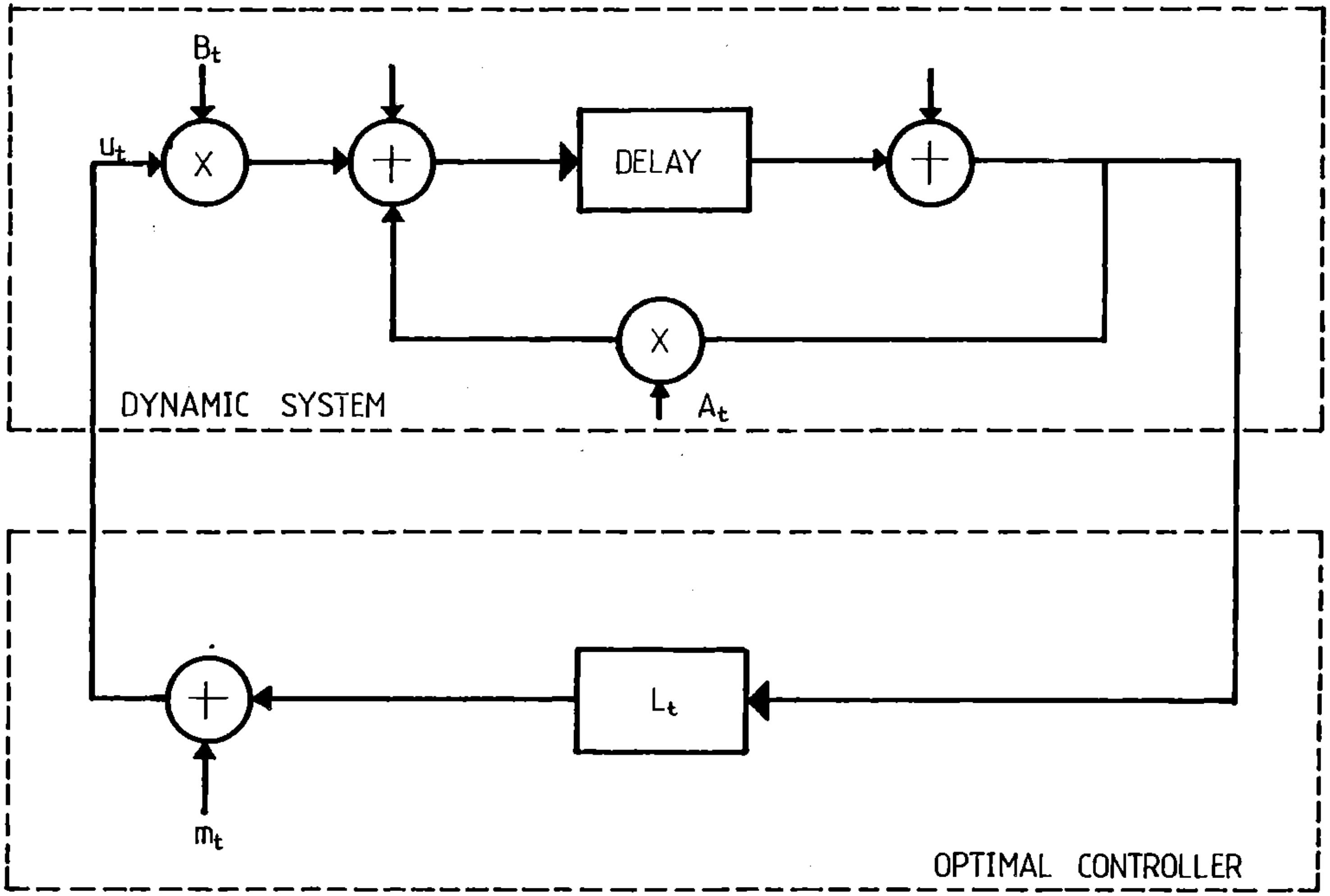


Fig. 2.1 Structure of the optimal feedback controller

so even in the extreme case where $\bar{B} = 0$, that is, when the system is 'most' uncontrollable on the average, since the controller can use the information about the high correlation in a useful way. When the matrix A_t is uncertain, then, of course, the controller will be more active, though the degree to which it will be so will depend on the other terms in the expression, since K_t appears in both the numerator and the denominator. Similar observations can be made for the various covariances in the equation for 'm_t'. For example, if B_t and c_t are strongly correlated then the magnitude of m_t is greater, as it can more effectively cancel the exogenous driving term c_t.

We note also that the certainty-equivalent control law is different from the optimal control law. It can be obtained from the optimal law by setting all covariances to zero. Basically, the optimal control takes into account the uncertainty in the parameters.

The optimal control is without any posterior learning. This, in fact, we had already anticipated when we defined our information pattern. The random matrices in the system equation are white and therefore unidentifiable. It is as if at each new time instant, the system restructures itself anew according to some unknown (and not necessarily constant) probability distribution, whose first and second moments, however, are known to us. The control system must adapt itself to this visceral change in order to minimize the cost-to-go. The whiteness of the noise does not permit us to reduce future uncertainty by present control action, which is to say that the control does not perform a dual role. Note however that the optimal decision certainly uses a priori knowledge of future randomness. That is, we know and make use of the a priori knowledge of the various future means and covariances. The problem and its solution are changed if we exclude knowledge of future statistics from the information pattern.

Physically, of course, this is quite unrealistic, and we ought to mention some ways in which this choice of modelling a stochastic system can be useful. In reality some learning is always possible and systems are never so insistently white. If we assume that the parameters are unknown but constant, we know that leads to the well-known dual problem, which does not admit of an exact analytical solution. With our assumption of whiteness we face a problem that is analytically tractable and that leads to a control that can be easily implemented. Moreover, economists have argued that in economic systems, it may be desirable to treat unknown parameters as purely random to obtain a consequent caution in the control, especially when B_t is not known accurately. Athans and Varaiya [44] have argued that the control of white parameter systems represents a worst-case situation in which the ratio (for scalar systems)

$$\frac{K\ (0 \mid \Sigma_a \neq 0,\ \Sigma_b \neq 0,\ \Sigma_{ba} \neq 0)}{K\ (0 \mid \Sigma_a = \Sigma_b = \Sigma_{ba} = 0}\ \geqslant\ 1$$

provides a measure of the deterioration in performance due to the unknown parameters, which can provide a guide as to whether sophisticated parameter estimation and adaptive control algorithms are warranted.

2.5 The "Uncertainty Threshold Principle" :

In this section we examine the asymptotic behaviour of linear random parameter systems. We assume here that all means and covariances and the weighting matrices in the cost functional are constant.

Let us first consider the simplest situation of scalar systems in a regulator problem type setting without additive noise. We have :

$$x_{t+1} = a_t x_t + b_t u_t \qquad x_o \text{ given} \qquad t = 0,1,2, \ldots , N \tag{2.5.1}$$

Here, a_t and b_t are white with given means, variances and covariance, all of which are constant. Note that the term c_t is absent.

$$J = \frac{1}{2} E \left\{ \sum_{k=o}^{N-1} [Qx_k^2 + Ru_k^2] + Qx_N^2 \right\} \tag{2.5.2}$$

Note that we have no non-zero trajectories to track.

The solution to this is obtained from our earlier general solution and is given by :

$$u_t^* = L_t x_t \tag{2.5.3}$$

$$L_t = - \frac{K_{t+1} (\Sigma_{ab} + \bar{a}\bar{b})}{R + (\Sigma_b + \bar{b}^2) K_{t+1}} \tag{2.5.4}$$

$$K_t = Q + K_{t+1}(\bar{a}^2 + \Sigma_a) - \frac{K_{t+1}^2 (\Sigma_{ab} + \bar{a}\bar{b})^2}{R + (\Sigma_b + \bar{b}^2) K_{t+1}} \tag{2.5.5}$$

$$K_N = Q \tag{2.5.6}$$

$$J^* = \frac{1}{2} x_o^2 K_o \tag{2.5.7}$$

This set of equations has been investigated by Ku [1] and gives rise to what is called the Uncertainty Threshold Principle. This is basically a result regarding the stability of the nonlinear difference equation for K_t. Its implications are discussed fully in Ku [1]. Here we shall merely give an informal expositional argument and then see what can be said for the general multivariable case.

In Eq. 2.5.5 assume that K_{t+1} is "large". Then we have the approximate relation :

$$K_t \approx m \cdot K_{t+1}$$

where 'm', the threshold parameter, is given by :

$$m = \Sigma_a + \bar{a}^2 - \frac{(\Sigma_{ab} + \bar{a}\bar{b})^2}{\Sigma_b + \bar{b}^2} \qquad (2.5.8)$$

If $m > 1$, then obviously K_t blows up as $N \to \infty$, so that a steady-state solution does not exist in this case. In fact, the uncertainty threshold principle states that for the infinite horizon problem, a necessary and sufficient condition for a solution to exist is $m < 1$.

If K_t blows up for the infinite horizon problem, then so does the cost J^* which means the optimal control problem has no solution. This makes good intuitive sense too, because if there is too much uncertainty in a system then there is little one can do to control its evolution over a long period of time.

We would expect a similar result to hold for multivariable systems as well. However, it seems that a neat mathematical expression for the threshold is not possible owing to the complexity of the equations involved. A special case of multivariable systems has been explored by Ku [1] in which the eigenvalues of the A matrix have to satisfy a threshold. The general case, in which we consider the multivariable tracking problem with additive noise is, as one would imagine, hopelessly complicated. Here we must consider the stability of three equations, for K_t, p_t and g_t, to determine whether the infinite-horizon cost remains finite or not.

2.6 Conclusion :

In this chapter, we have stated and solved the optimal tracking problem for a linear-quadratic system with purely random parameters. We briefly noted the salient characteristics of the 'Uncertainty Threshold Principle' and found that the multivariable case presents formidable analytical problems which may make it impossible to derive a mathematical expression for the threshold.

Now that we have the complete solution, we can explore, in the next chapter, the derivation of the sensitivity equations for this problem and then apply them to a macroeconometric model of the U.S. economy.

CHAPTER 3

SENSITIVITY EQUATIONS

3.1 Introduction :

In this chapter, our main objective will be to develop equations to analyse the sensitivity of linear systems with random parameters to variations in parameter uncertainties.

The concept of sensitivity is a very general one and 'sensitivity analysis' is a fairly well-developed tool. In any real system, there is always some uncertainty associated with the exact values of its parameters, either because of imperfect information or because of approximations made in the modelling process or possibly because of some inherent randomness in the behaviour of its parameters. This obviously affects the efficacy of any control law, whether open or closed loop, as well as the accuracy of any simulation of the system. If the behaviour of a system is dramatically different as a result of variations in parameter values, then we say the system is very sensitive to such variations. This gives us some useful information in assessing the reliability of our efforts. An excellent example of such a situation is provided by the now infamous 'Limits to Growth' report by the Club of Rome [48]. Sharply different qualitative results, such as lack of evidence on which to base a prediction of the collapse of world population, can be obtained by appropriate combinations of small changes. This illustrates the caution that is necessary in basing policy judgments on sensitive models.

There are many different questions one can ask in this area of sensitivity analysis. One basic question is how perturbations in the parameters affect the optimal performance of the system. If the optimal cost or optimal welfare are significantly altered as a result of small variations in the parameters, then obviously our analysis and policy recommendations are not very reliable. This kind of study is probably most useful in dealing with large economic and socio-economic systems, in which little is known about the actual structure of the system, and in which there is almost always a great deal of uncertainty about parameter values.

For systems with parameters that are modelled as being deterministic, the standard procedure is to derive sensitivity equations with respect to variations in the parameter values themselves. This has already been done and is readily available in the literature.

For systems whose parameters are modelled as random processes, however, it makes sense to look instead at the effects of variations in the parameter uncertainties, that is, the variances and covariances of these parameters. This leads to a slightly modified set of equations, though the basic approach remains the same. Sensitivities may either be absolute, or relative to the parameter and optimal cost values, and it may be useful, in general, to look at both sets of numbers. We can even rank parameters in order of their sensitivities which may help to identify the 'pressure points' of a system.

We shall first derive general sensitivity equations from the optimal control solution developed in the previous chapter. Next, we briefly describe a small econometric model of the U.S. economy and do a sensitivity analysis of the model. We end with a discussion of the results and possible uses for a sensitivity analysis and ranking of parameters.

3.2 Problem Statement :

We are given the following linear multivariable system :

$$x_{t+1} = A_t x_t + B_t u_t + c_t \qquad x_o = \bar{x}_o \qquad t = 0,1,2, \ldots , N-1 \tag{3.2.1}$$

We have perfect measurements of the state. The elements of the matrices A_t, B_t and the vector c_t are all random variables. Each element constitutes a white stochastic process with given mean and variance. That is, we are given the covariance matrices Σ_A, Σ_B, Σ_c, Σ_{BA}, Σ_{Bc}, Σ_{Ac}, where each covariance matrix is defined by the convention described in chapter 2, and we are given the mean matrices A and B and the mean vector c. We choose to minimize the standard quadratic cost functional:

$$J = \frac{1}{2} E \left\{ \sum_{t=o}^{N-1} [(x_t - \tilde{x}_t)'Q (x_t - \tilde{x}_t) + (u_t - \tilde{u}_t)'R(u_t - \tilde{u}_t)] + (x_N - \tilde{x}_N)'Q(x_N - \tilde{x}_N) \right\} \tag{3.2.2}$$

The sequences $\{\tilde{x}_t\}$, $\{\tilde{u}_t\}$ are, of course, given.

This is so far only a restatement of the optimal control problem considered in the previous chapter. Its solution has also been given there.

Now we would like to pose the following question. Let σ denote any element of any one of the six covariance matrices. The question is : how sensitive is the optimal cost to small variations in σ ?

If J^* denotes the optimal cost, then the answer is given by the number $\frac{\partial J^*}{\partial \sigma}\Big|_0$. Here the symbol $\Big|_0$ is used to mean 'evaluated at the given values of the various means and covariances'. This number is an absolute measure of sensitivity. If there is a small absolute change $\delta\sigma$ in σ, it induces a corresponding absolute change δJ^* in J^*, whose magnitude is given by the relation :

$$\delta J^* = \frac{\partial J^*}{\partial \sigma}\Big|_0 \delta\sigma \tag{3.2.3}$$

If $\frac{\partial J^*}{\partial \sigma}\Big|_0$ is large, then the induced change δJ^* is also proportionally large. It is in this sense that $\frac{\partial J^*}{\partial \sigma}\Big|_0$ is an absolute measure of sensitivity.

We can also obtain a relative measure of sensitivity by noting that :

$$\frac{\delta J^*}{J^*} = \left[\frac{\partial J^*}{\partial \sigma}\Big|_0 \frac{\sigma}{J^*}\right] \frac{\partial \sigma}{\sigma} \tag{3.2.4}$$

This number, $\frac{\partial J^*}{\partial \sigma}\Big|_0 \frac{\sigma}{J^*}$, tells us how a percentage or relative change in σ is transformed into a percentage or relative change in J^*. In general, the appropriate measure will depend upon the application at hand, and in some cases both measures may provide useful information.

For now, let us turn to deriving equations that will enable us to evaluate the quantity $\left.\frac{\partial J^*}{\partial \sigma}\right|_0$.

3.3 Derivation of Sensitivity Equations :

The derivation of sensitivity equations for a linear random parameter system is quite straightforward though the final equations are somewhat cumbersome to use. We first restate the solution to the optimal control problem (see Chapter 2).

$$u_t^* = L_t x_t + m_t \tag{3.3.1}$$

$$L_t = - [R + \overline{B'K_{t+1}B}]^{-1} [\overline{B'K_{t+1}A}] \tag{3.3.2}$$

$$m_t = - [R + \overline{B'K_{t+1}B}]^{-1} [\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t] \tag{3.3.3}$$

$$K_t = Q + [\overline{A'K_{t+1}A}] + [\overline{B'K_{t+1}A}]' L_t \tag{3.3.4}$$

$$p_t = - Q\tilde{x}_t + [\overline{A'K_{t+1}c}] + \bar{A}'p_{t+1} + [\overline{B'K_{t+1}A}]' m_t \tag{3.3.5}$$

$$g_t = \frac{1}{2} \tilde{x}_t' Q \tilde{x}_t + \frac{1}{2} \tilde{u}_t' R \tilde{u}_t + \frac{1}{2} [\overline{c'K_{t+1}c}] + \bar{c}'p_{t+1} + \frac{1}{2} [\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t]' m_t + g_{t+1} \tag{3.3.6}$$

$$K_N = Q \tag{3.3.7}$$

$$p_N = - Q\tilde{x}_N \tag{3.3.8}$$

$$g_N = \frac{1}{2} \tilde{x}_N' Q \tilde{x}_N \tag{3.3.9}$$

$$J^* = \frac{1}{2} x_o'K_o x_o + p_o x_o + g_o \tag{3.3.10}$$

The evolution of the state is now given by :

$$x_{t+1} = (A_t + B_t L_t)\, x_t + B_t m_t + c_t \quad ; \quad x_o = \bar{x}_o \tag{3.3.11}$$

In order to calculate $\frac{\partial J^*}{\partial \sigma}$, we need to calculate $\frac{\partial K_o}{\partial \sigma}, \frac{\partial p_o}{\partial \sigma}, \frac{\partial g_o}{\partial \sigma}$, which in turn require us to calculate $\frac{\partial L_o}{\partial \sigma}, \frac{\partial m_o}{\partial \sigma}$. Let us, therefore, differentiate the appropriate equations.

Preliminaries :

Before we actually carry out the differentiation let us state a few simple algebraic results in order to make the derivation a little clearer.

(b) $$\frac{\partial}{\partial \sigma} \operatorname{tr} A = \operatorname{tr} \frac{\partial A}{\partial \sigma} \tag{3.3.12}$$

(b) Let G be a random matrix with mean G and covariance Σ_G and let H be a deterministic matrix and some function of σ, where σ may be an element of Σ_G. Then,

$$\begin{aligned}
\frac{\partial}{\partial \sigma} [\overline{G'HG}\,] &= \frac{\partial}{\partial \sigma} [\, \sum_{k,\ell} \operatorname{tr} (HP_k' \Sigma_G P_\ell)\, E_{k\ell} + \bar{G}'H\bar{G}\,] \\
&= \sum_{k,\ell} \frac{\partial}{\partial \sigma} \operatorname{tr} (HP_k' \Sigma_G P_\ell)\, E_{k\ell} + \frac{\partial}{\partial \sigma} (\bar{G}'H\bar{G}) \\
&= \sum_{k,\ell} \operatorname{tr} \left(\frac{\partial H}{\partial \sigma} P_k' \Sigma_G P_\ell \right) E_{k\ell} + \bar{G}' \frac{\partial H}{\partial \sigma} \bar{G} \\
&\quad + \sum_{k,\ell} \operatorname{tr} \left(HP_k' \frac{\partial \Sigma_G}{\partial \sigma} P_\ell\right) E_{k\ell}
\end{aligned}$$

Let $$f(\overline{G'HG}) \triangleq \sum_{k,\ell} \operatorname{tr} \left(\frac{\partial H}{\partial \sigma} P_k' \Sigma_G P_\ell \right) E_{k\ell} + \bar{G}' \frac{\partial H}{\partial \sigma} \bar{G} \tag{3.3.13}$$

We make this definition only to save us some repetitious writing.

(c) Let

$$r = 1 + \text{quotient } \left[\frac{i-1}{n}\right]$$

$$s = 1 + \text{quotient } \left[\frac{j-1}{n}\right]$$

$$u = 1 + \text{remainder } \left[\frac{i-1}{n}\right]$$

$$v = 1 + \text{remainder } \left[\frac{j-1}{n}\right]$$

where $i = 1,2, \ldots, n^2$; $j = 1,2, \ldots, n^2$

Let σ_{ij} be the (i,j)-th element of Σ_G

Then,

$$\frac{\partial \Sigma_G}{\partial \sigma_{ij}} = E_{ij} + E_{ji} - E_{ij}\delta_{ij} \qquad \text{(because } \Sigma_G \text{ is symmetric)}$$

Therefore,

$$P_k' \frac{\partial \Sigma_G}{\partial \sigma_{ij}} P_\ell = P_k' E_{ij} P_\ell + P_k' E_{ji} P_\ell - P_k' E_{ij} \delta_{ij} P_\ell$$

$$= E_{uv}\delta_{kr}\delta_{\ell s} + E_{vu}\delta_{ks}\delta_{\ell r} - E_{uv}\delta_{kr}\delta_{\ell s}\delta_{ij}$$

which follows from the fact that (i,j) must belong to the (k,ℓ)-th block of E_{ij} for a non-zero product. Hence

$$\sum_{k,\ell} \text{tr}\left(HP_k' \frac{\partial \Sigma_G}{\partial \sigma_{ij}} P_\ell\right) E_{k\ell}$$

$$= \text{tr}(HE_{uv})E_{rs} + \text{tr}(HE_{vu})E_{sr} - \text{tr}(HE_{uv})E_{rs}\delta_{ij}$$

$$= h^{vu}E_{rs} + h^{uv}E_{sr} - h^{vu}E_{rs}\delta_{ij} \qquad \text{where } h^{vu} \text{ is the (v,u)-th element of H, etc.} \qquad (3.3.14)$$

For i = j, this simplifies to :

$$h^{vu}E_{rs}$$

(d) $$\frac{\partial}{\partial\sigma}(AA^{-1}) = \frac{\partial}{\partial\sigma}(I) = 0$$

$$\frac{\partial A}{\partial\sigma} . A^{-1} + A . \frac{\partial A^{-1}}{\partial\sigma} = 0$$

$$\frac{\partial A^{-1}}{\partial\sigma} = - A^{-1} . \frac{\partial A}{\partial\sigma} . A^{-1} \qquad (3.3.15)$$

Derivation :

We shall now differentiate the optimal equations stated above. There are six separate cases to be considered : σ_{ij} can be the (i,j)-th element of any one of Σ_A, Σ_B, Σ_c, Σ_{BA}, Σ_{Bc}, Σ_{Ac}. We shall only look at Σ_A, Σ_B, Σ_{BA}.

Let $$S_t = [R + \overline{B'K_{t+1}B}]$$

$$P_t = \frac{\partial K_t}{\partial\sigma_{ij}}$$

1. $$\frac{\partial L_t}{\partial\sigma_{ij}} = -\frac{\partial}{\partial\sigma_{ij}}[R + \overline{B'K_{t+1}B}]^{-1} . (\overline{B'K_{t+1}A})$$

$$- (R + \overline{B'K_{t+1}B})^{-1} \frac{\partial}{\partial\sigma_{ij}}(\overline{B'K_{t+1}A})$$

$$= S_t^{-1} \frac{\partial S_t}{\partial\sigma_{ij}} S_t^{-1} . \overline{B'K_{t+1}A} - S_t^{-1}\frac{\partial}{\partial\sigma_{ij}}(\overline{B'K_{t+1}A})$$

$$= S_t^{-1} \frac{\partial}{\partial\sigma_{ij}}(\overline{B'K_{t+1}B}) S_t^{-1} . \overline{B'K_{t+1}A} - S_t^{-1}\frac{\partial}{\partial\sigma_{ij}}(\overline{B'K_{t+1}A})$$

$$= S_t^{-1} [f(\overline{B'K_{t+1}B}) + \sum_{k,\ell} tr (K_{t+1} P_k' \frac{\partial\Sigma_B}{\partial\sigma_{ij}} P_\ell) E_k] S_t^{-1} . \overline{B'K_{t+1}A}$$

$$- S_t^{-1}[f(\overline{B'K_{t+1}A}) + \sum_{k,\ell} tr(K_{t+1}P_k' \frac{\partial\Sigma_{BA}}{\partial\sigma_{ij}} P_\ell) E_{k\ell}] \qquad (3.3.16)$$

(a) $\sigma_{ij} \varepsilon \Sigma_A$:

$$\frac{\partial \Sigma_B}{\partial \sigma_{ij}} = 0 \qquad \frac{\partial \Sigma_{BA}}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial L_t}{\partial \sigma_{ij}} = S_t.^{-1} f(\overline{B'K_{t+1}B}).S_t.^{-1}(\overline{B'K_{t+1}A}) - S_t.^{-1} f(\overline{B'K_{t+1}A})$$

$$= (R + \overline{B'K_{t+1}B})^{-1}. (\sum_{k,\ell} tr (P_{t+1}P_k'\Sigma_B P_\ell) E_{k\ell} + \bar{B}'P_{t+1}\bar{B}).$$

$$(R + \overline{B'K_{t+1}B})^{-1}. (\overline{B'K_{t+1}A})$$

$$- (R + \overline{B'K_{t+1}B})^{-1} (\sum_{k,\ell} tr (P_{t+1}P_k' \Sigma_{BA}P_\ell)E_{k\ell} + \bar{B}'P_{t+1}\bar{A})$$

(3.3.17)

(b) $\sigma_{ij} \varepsilon \Sigma_B$: $\qquad \frac{\partial \Sigma_{BA}}{\partial \sigma_{ij}} = 0$

Therefore,

$$\frac{\partial L_t}{\partial \sigma_{ij}} = S_t^{-1}[f(\overline{B'K_{t+1}B}) + k_{t+1}^{vu} E_{rs}]S_t^{-1}.(\overline{B'K_{t+1}A}) - S_t^{-1}.f(\overline{B'K_{t+1}A})$$

$$= (R + \overline{B'K_{t+1}B})^{-1}(\sum_{k,\ell} tr (P_{t+1}P_k'\Sigma_B P_\ell)E_{k\ell} + \bar{B}'P_{t+1}\bar{B} +$$

$$k_{t+1}^{vu} E_{rs}). (R + \overline{B'K_{t+1}B})^{-1}.(\overline{B'K_{t+1}A})$$

$$- (R+\overline{B'K_{t+1}B})^{-1}(\sum_{K,\ell} tr(P_{t+1}P_k'\Sigma_{BA}P_\ell) + \bar{B}'P_{t+1}\bar{A})$$

(3.3.18)

(c) $\sigma_{ij} \varepsilon \Sigma_{BA}$:

$$\frac{\partial \Sigma_B}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial L_t}{\partial \sigma_{ij}} = S_t^{-1}.f([\overline{B'K_{t+1}B}).S_t^{-1}.(\overline{B'K_{t+1}A}) - S_t^{-1}[f(\overline{B'K_{t+1}A}) + \overset{vu}{k_{t+1}} E_{rs}]$$

$$= (R+\overline{B'K_{t+1}B})^{-1}(\sum_{k,\ell} tr(P_{t+1}P_k'\Sigma_B P_\ell)E_{k\ell} + \bar{B}'P_{t+1}\bar{B}).$$

$$(R + \overline{B'K_{t+1}B})^{-1}.(\overline{B'K_{t+1}A})$$

$$- (R+\overline{B'K_{t+1}B})^{-1}[\sum_{k,\ell} tr(P_{t+1}P_k'\Sigma_{BA}P_\ell)E_{k\ell} + \bar{B}'P_{t+1}\bar{A} + \overset{vu}{k_{t+1}} E_{rs}] \qquad (3.3.19)$$

2. $$\frac{\partial K_t}{\partial \sigma_{ij}} = \frac{\partial}{\partial \sigma_{ij}}(\overline{A'K_{t+1}A}) + \frac{\partial}{\partial \sigma_{ij}}(\overline{B'K_{t+1}A}).L_t + (\overline{B'K_{t+1}A})\frac{\partial L_t}{\partial \sigma_{ij}}$$

$$= f(\overline{A'K_{t+1}A}) + \sum_{k,\ell} tr\,(K_{t+1}P_k'\frac{\partial \Sigma_A}{\partial \sigma_{ij}}P_\ell)\,E_{k\ell}$$

$$+ [f(\overline{B'K_{t+1}A}) + \sum_{k,\ell} tr(K_{t+1}P_k'\frac{\partial \Sigma_{BA}}{\partial \sigma_{ij}}P_\ell)\,E_{k\ell}]'\,.\,L_t$$

$$+ \overline{(B'K_{t+1}A)}'\,.\,\frac{\partial L_t}{\partial \sigma_{ij}} \qquad (3.3.20)$$

(a) $\sigma_{ij} \in \Sigma_A$:

$$\frac{\partial \Sigma_{BA}}{\partial \sigma_{ij}} = 0 \qquad \frac{\partial \Sigma_B}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial K_t}{\partial \sigma_{ij}} = f(\overline{A'K_{t+1}A}) + k^{vu}_{t+1}E_{rs} + [f(\overline{B'K_{t+1}A})]' \cdot L_t + (\overline{B'K_{t+1}A})' \cdot \frac{\partial L_t}{\partial \sigma_{ij}}$$

Therefore,

$$P_t = \sum_{k,\ell} tr(P_{t+1}P_k'\Sigma_A P_\ell)E_{k\ell} + \bar{A}'P_{t+1}\bar{A} + k^{vu}_{t+1} E_{rs}$$

$$+ [\sum_{k,\ell} tr(P_{t+1}P_k' \Sigma_{BA}P_\ell)E_{k\ell} + \bar{B}'P_{t+1}\bar{A}]' \cdot L_t + (\overline{B'K_{t+1}A})' \cdot \frac{\partial L_t}{\partial \sigma_{ij}}$$

(3.3.21)

(b) $\sigma_{ij} \in \Sigma_B$:

$$\frac{\partial \Sigma_A}{\partial \sigma_{ij}} = 0 \qquad \frac{\partial \Sigma_{BA}}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial K_t}{\partial \sigma_{ij}} = f(\overline{A'K_{t+1}A}) + [f(\overline{B'K_{t+1}A})]' \cdot L_t + (\overline{B'K_{t+1}A})' \cdot \frac{\partial L_t}{\partial \sigma_{ij}}$$

Therefore,

$$P_t = \sum_{k,\ell} tr(P_{t+1}P_k' \Sigma_A P_\ell)E_{k\ell} + \bar{A}'P_{t+1}\bar{A}$$

$$+ [\sum_{k,\ell} tr(P_{t+1}P_k' \Sigma_{BA}P_\ell)E_{k\ell} + \bar{B}'P_{t+1}\bar{A}]' \cdot L_t$$

$$+ [\overline{B'K_{t+1}A}]' \cdot \frac{\partial L_t}{\partial \sigma_{ij}}$$

(3.3.22)

(c) $\sigma_{ij} \ \varepsilon \ \Sigma_{BA}$:

$$\frac{\partial \Sigma_A}{\partial \sigma_{ij}} = 0 \qquad \frac{\partial \Sigma_B}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial K_t}{\partial \sigma_{ij}} = f(\overline{A'K_{t+1}A}) + [f(\overline{B'K_{t+1}A}) + k^{vu}_{t+1}E_{rs}]' \cdot L_t + (\overline{B'K_{t+1}A})' \cdot \frac{\partial L_t}{\partial \sigma_{ij}}$$

Therefore,

$$P_t = \sum_{k,\ell} tr(P_{t+1}P_k' \Sigma_A P_\ell)E_{k\ell} + \bar{A}'P_{t+1}\bar{A} + [\sum_{k,\ell} tr(P_{t+1}P_k' \Sigma_{BA}P_\ell)E_{k\ell} + \bar{B}'P_{t+1}\bar{A} + k^{vu}_{t+1}E_{rs}]' \cdot L_t + (\overline{B'K_{t+1}A})' \cdot \frac{\partial L_t}{\partial \sigma_{ij}} \tag{3.3.23}$$

3. $$\frac{\partial m_t}{\partial \sigma_{ij}} = - \frac{\partial}{\partial \sigma_{ij}} [R + \overline{B'K_{t+1}B}]^{-1}(\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t) - (R + \overline{B'K_{t+1}B})^{-1} \frac{\partial}{\partial \sigma_{ij}} [\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t]$$

$$= S_t^{-1} \frac{\partial S_t}{\partial \sigma_{ij}} S_t^{-1}(\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t) - (S_t^{-1}) \left(\frac{\partial}{\partial \sigma_{ij}} (\overline{B'K_{t+1}c}) + \bar{B}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}} \right)$$

$$= S_t^{-1} \frac{\partial S_t}{\partial \sigma_{ij}} St^{-1} (\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - Ru_t)$$

$$- S_t^{-1}[f(\overline{B'K_{t+1}c}) + \sum_k tr(K_{t+1}P_k' \frac{\partial \Sigma_{Bc}}{\partial \sigma_{ij}}) e_k + \bar{B}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}}]$$

$$= S_t^{-1} [f(\overline{B'K_{t+1}B}) + \sum_{k,\ell} tr(K_{t+1}P_k' \Sigma_B P_\ell) E_{k\ell}] St^{-1}.$$

$$(\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t)$$

$$- S_t^{-1}[f(\overline{B'K_{t+1}c}) + \sum_{k,\ell} tr(K_{t+1}P_k' \frac{\partial \Sigma_{Bc}}{\partial \sigma_{ij}} P_\ell) E_{k\ell} + \bar{B}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}}] \tag{3.3.24}$$

(a) $\sigma_{ij} \in \Sigma_A$:

$$\frac{\partial \Sigma_B}{\partial \sigma_{ij}} = 0 \qquad \frac{\partial \Sigma_{Bc}}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial m_t}{\partial \sigma_{ij}} = S_t^{-1} [f(\overline{B'K_{t+1}B})] S_t^{-1} . (\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t)$$

$$- St^{-1} [f(\overline{B'K_{t+1}c}) + \bar{B}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}}]$$

$$= (R + \overline{B'K_{t+1}B})^{-1}(\sum_{k,\ell} tr(P_{t+1}P_k' \Sigma_B P_\ell) E_{k\ell} + \bar{B}'P_{t+1}\bar{B}).$$

$$(R + \overline{B'K_{t+1}B})^{-1} (\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t)$$

$$- (R + \overline{B'K_{t+1}B})^{-1} [\sum_k tr(P_{t+1} P_k' \Sigma_{Bc})e_k + \bar{B}'P_{t+1}\bar{c}$$

$$+ \bar{B}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}}] \tag{3.3.25}$$

(b) $\sigma_{ij} \; \varepsilon \; \Sigma_B$:

$$\frac{\partial \Sigma_{Bc}}{\partial \sigma_{ij}} = 0$$

$$\frac{\partial m_t}{\partial \sigma_{ij}} = S_t^{-1} \left[f(\overline{B'K_{t+1}B}) + k_{t+1}^{vu} E_{rs} \right] S_t^{-1} . (\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t)$$

$$- S_t^{-1} \left[f(\overline{B'K_{t+1}c}) + \bar{B}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}} \right]$$

$$= (R + \overline{B'K_{t+1}B})^{-1} \left[\sum_{k,\ell} tr(P_{t+1}P_k'\Sigma_B P_\ell) E_{k\ell} + \bar{B}'P_{t+1}\bar{B} + k_{t+1}^{vu} E_{rs} \right) . (R+\overline{B'K_{t+1}B})^{-1} (\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t)$$

$$- (R + \overline{B'K_{t+1}B})^{-1} \left[\sum_k tr(P_{t+1}P_k'\Sigma_{Bc}) e_k + \bar{B}'P_{t+1}\bar{c} + \bar{B}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}} \right] \qquad (3.3.26)$$

(c) $\sigma_{ij} \; \varepsilon \; \Sigma_{BA}$:

$$\frac{\partial \Sigma_B}{\partial \sigma_{ij}} = 0 \qquad \frac{\partial \Sigma_{Bc}}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial m_t}{\partial \sigma_{ij}} = S_t^{-1} \left[f(\overline{B'K_{t+1}B}) \right] S_t^{-1} . (\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t)$$

$$- S_t^{-1} \left[f(\overline{B'K_{t+1}c}) + \bar{B}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}} \right]$$

$$= \quad (R + \overline{B'K_{t+1}B})^{-1} \; [\sum_{k,\ell} tr(P_{t+1}P_k'\Sigma_B P_\ell) \; E_{k\ell} + \bar{B}'P_{t+1}\bar{B} \;]$$

$$(R + \overline{B'K_{t+1}B})^{-1} \; (\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t)$$

$$- \; (R + \overline{B'K_{t+1}B})^{-1} \; [\sum_k tr(P_{t+1}P_k' \; \Sigma_{Bc})e_k + \bar{B}'P_{t+1}c$$

$$+ \; \bar{B}' \; . \; \frac{\partial p_{t+1}}{\partial \sigma_{ij}} \;] \tag{3.3.27}$$

4. $$\frac{\partial p_t}{\partial \sigma_{ij}} = \frac{\partial}{\partial \sigma_{ij}} (\overline{A'K_{t+1}c}) + \bar{A}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}} + \frac{\partial}{\partial \sigma_{ij}} (\overline{B'K_{t+1}A})' \; . \; m_t$$

$$+ \; (\overline{B'K_{t+1}A}) \; . \; \frac{\partial m_t}{\partial \sigma_{ij}}$$

$$= \quad [\; f(\overline{A'K_{t+1}c}) + \sum_k tr(K_{t+1}P_k' \frac{\partial \Sigma_{Ac}}{\partial \sigma_{ij}}) \; e_k \;] + \bar{A}' \; \frac{\partial p_{t+1}}{\partial \sigma_{ij}}$$

$$+ \; [\; f(\overline{B'K_{t+1}A}) + \sum_{k,\ell} tr(K_{t+1}P_k' \; \frac{\partial \Sigma_{BA}}{\partial \sigma_{ij}} \; P_\ell)E_{k\ell}]' . \; m_t$$

$$+ \; (\overline{B'K_{t+1}A})' \; . \; \frac{\partial m_t}{\partial \sigma_{ij}} \tag{3.3.28}$$

(a) $\sigma_{ij} \; \epsilon \; \Sigma_A$:

$$\frac{\partial \Sigma_{Ac}}{\partial \sigma_{ij}} = 0 \qquad \frac{\partial \Sigma_{BA}}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial p_t}{\partial \sigma_{ij}} = \quad [\sum_k tr(P_{t+1}P_k' \; \Sigma_{A_c}) \; e_k + \bar{A}'P_{t+1}\bar{c} \;] + \bar{A}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}}$$

$$+ \; [\sum_{k,\ell} tr(P_{t+1}P_k' \; \Sigma_{BA}P_\ell)E_{k\ell} + \bar{B}'P_{t+1}\bar{A} \;]' . \; m_t$$

$$+ \; (\overline{B'K_{t+1}A})' \; . \; \frac{\partial m_t}{\partial \sigma_{ij}} \tag{3.3.29}$$

(b) $\sigma_{ij} \in \Sigma_B$:

$$\frac{\partial \Sigma_{Ac}}{\partial \sigma_{ij}} = 0 \qquad \frac{\partial \Sigma_{BA}}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial p_t}{\partial \sigma_{ij}} = [\sum_k tr\,(P_{t+1}P_k'\Sigma_{Ac})e_k + \bar{A}'P_{t+1}\bar{c}\,] + \bar{A}'\frac{\partial p_{t+1}}{\partial \sigma_{ij}} + [\sum_{k,\ell} tr(P_{t+1}P_k'\,\Sigma_{BA}P_\ell)E_{k\ell} + \bar{B}'P_{t+1}\bar{A}\,]' \cdot m_t + (\overline{B'K_{t+1}A}\,)' \cdot \frac{\partial m_t}{\partial \sigma_{ij}} \tag{3.3.30}$$

(c) $\sigma_{ij} \in \Sigma_{BA}$:

$$\frac{\partial \Sigma_{Ac}}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial p_t}{\partial \sigma_{ij}} = [\sum_k tr(P_{t+1}P_k'\,\Sigma_{Ac})e_k + \bar{A}'P_{t+1}\bar{c}\,] + \bar{A}'\frac{\partial p_{t+1}}{\partial \sigma_{ij}} + [\sum_{k,\ell} tr(P_{t+1}P_k'\,\Sigma_{BA}P_\ell)E_{k\ell} + \bar{B}'P_{t+1}\bar{A} + k_{t+1}^{vu}\,E_{rs}\,]'\,m_t + (\overline{B'K_{t+1}A})' \cdot \frac{\partial m_t}{\partial \sigma_{ij}} \tag{3.3.31}$$

5. $$\frac{\partial g_t}{\partial \sigma_{ij}} = \frac{1}{2}\ \frac{\partial}{\partial \sigma_{ij}}(\overline{c' K_{t+1} c}) + \bar{c}'\ \frac{\partial p_{t+1}}{\partial \sigma_{ij}}$$

$$+ \frac{1}{2}\ \frac{\partial}{\partial \sigma_{ij}}(\overline{B' K_{t+1} c} + \bar{B}' p_{t+1} - R\tilde{u}_t)'\ m_t$$

$$+ \frac{1}{2}\ \frac{\partial}{\partial \sigma_{ij}}(\overline{B' K_{t+1} c} + \bar{B}' p_{t+1} - R\tilde{u}_t)' .\ \frac{\partial m_t}{\partial \sigma_{ij}} + \frac{\partial g_{t+1}}{\partial \sigma_{ij}}$$

$$= \frac{1}{2} [\ f(\overline{c' K_{t+1} c}) + tr(K_{t+1} \frac{\partial \Sigma_c}{\partial \sigma_{ij}})\ ] + \bar{c}'\ \frac{\partial p_{t+1}}{\partial \sigma_{ij}}$$

$$+ \frac{1}{2} [\ f(\overline{B' K_{t+1} c}) + \sum_k tr(K_{t+1} P_k' \frac{\partial \Sigma_{Bc}}{\partial \sigma_{ij}})\ e_k + \bar{B}'\ \frac{\partial p_{t+1}}{\partial \sigma_{ij}}\ ]' . m_t$$

$$+ \frac{1}{2}\ (\overline{B' K_{t+1} c} + \bar{B}' p_{t+1} - R\tilde{u}_t)' .\ \frac{\partial\ m_t}{\partial \sigma_{ij}} + \frac{\partial g_{t+1}}{\partial \sigma_{ij}} \qquad (3.3.32)$$

(a) $\sigma_{ij}\ \varepsilon\ \Sigma_A$:

$$\frac{\partial \Sigma_c}{\partial \sigma_{ij}} = 0 \qquad \frac{\partial \Sigma_{Bc}}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial g_t}{\partial \sigma} = \frac{1}{2} [\ tr\ (P_{t+1} \Sigma_c) + \bar{c}' P_{t+1} \bar{c}\ ] + \bar{c}'\ \frac{\partial p_{t+1}}{\partial \sigma_{ij}}$$

$$+ \frac{1}{2} [\ \sum_k tr(P_{t+1} P_k' \Sigma_{Bc}) e_k + \bar{B}' P_{t+1} \bar{c} + \bar{B}'\ \frac{\partial p_{t+1}}{\partial \sigma_{ij}}\ ]' .\ m_t$$

$$+ \frac{1}{2} [\ \overline{B' K_{t+1} c} + \bar{B}' p_{t+1} - R\tilde{u}_t\ ]' .\ \frac{\partial m_t}{\partial \sigma_{ij}} + \frac{\partial g_{t+1}}{\partial \sigma_{ij}}$$

(3.3.33)

(b) $\sigma_{ij}\ \varepsilon\ \Sigma_B$:

$$\frac{\partial \Sigma_c}{\partial \sigma_{ij}} = 0 \qquad \frac{\partial \Sigma_{Bc}}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial g_t}{\partial \sigma_{ij}} = \frac{1}{2} [\, tr(P_{t+1}\Sigma_c) + \bar{c}'P_{t+1}\bar{c} \,] + \bar{c}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}}$$

$$+ \frac{1}{2} [\, \sum_k tr(P_{t+1}P_k'\Sigma_{Bc})e_k + \bar{B}'P_{t+1}\bar{c} + \bar{B}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}} \,]' . \; m_t$$

$$+ \frac{1}{2} [\, \overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t \,]' . \; \frac{\partial m_t}{\partial \sigma_{ij}} + \frac{\partial g_{t+1}}{\partial \sigma_{ij}} \qquad (3.3.34)$$

(c) $\sigma_{ij} \in \Sigma_{BA}$:

$$\frac{\partial \Sigma_c}{\partial \sigma_{ij}} = 0 \qquad \frac{\partial \Sigma_{Bc}}{\partial \sigma_{ij}} = 0$$

Therefore,

$$\frac{\partial g_t}{\partial \sigma_{ij}} = \frac{1}{2} [\, tr(P_{t+1} \, \Sigma_c) + \bar{c}'P_{t+1}\bar{c} \,] + \bar{c}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}}$$

$$+ \frac{1}{2} [\, \sum_k tr(P_{t+1}P_k' \, \Sigma_{Bc})e_k + \bar{B}'P_{t+1}\bar{c} + \bar{B}' \frac{\partial p_{t+1}}{\partial \sigma_{ij}} \,]' . m_t$$

$$+ \frac{1}{2} [\, \overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t]' . \frac{\partial m_t}{\partial \sigma_{ij}} + \frac{\partial g_{t+1}}{\partial \sigma_{ij}} \qquad (3.3.35)$$

6. $$\frac{\partial K_N}{\partial \sigma_{ij}} = 0 \qquad (3.3.36)$$

$$\frac{\partial p_N}{\partial \sigma_{ij}} = 0 \qquad (3.3.37)$$

$$\frac{\partial g_N}{\partial \sigma_{ij}} = 0 \qquad (3.3.38)$$

7. $$\frac{\partial J^*}{\partial \sigma_{ij}} = \frac{1}{2} x_o' \frac{\partial K_o}{\partial \sigma_{ij}} x_o + \frac{\partial p_o'}{\partial \sigma_{ij}} x_o + \frac{\partial g_o}{\partial \sigma_{ij}} \qquad (3.3.39)$$

Evaluating this number finally gives us an absolute measure of the sensitivity of the optimal cost to variations in parameter uncertainties. As we mentioned before, we can also calculate from this a dimensionless number, a relative sensitivity, for each parameter, viz.

$$\left[\frac{\partial J^*}{\partial \sigma_{ij}} \cdot \frac{\sigma_{ij}}{J^*}\right]$$

We have, at this point, completed our derivation of the cost sensitivity equations. It is also frequently useful to look at the sensitivity of the optimal control law to parameter variations. Though the transformation itself in the optimal law is deterministic, the control is random because the state is random. Here again, therefore, it is more meaningful to calculate the sensitivity of the covariance matrix of the optimal control to parameter uncertainties. Mathematically, we would like to calculate $\frac{\partial \Sigma u_t}{\partial \sigma}$ where Σu_t is the covariance matrix of the optimal control u_t^*. We have :

$$u_t^* = L_t x_t + m_t$$

Therefore,

$$\Sigma_{u_t} = L_t \Sigma_t L_t' \qquad \text{where } \Sigma_t = \text{cov } \{x_t\} \qquad (3.3.40)$$

We need, therefore, to calculate Σ_t. This turns out to be a gargantuan mess, so we shall not bother to reproduce it here, and merely indicate the source of the complexity.

$$x_{t+1} = (A_t + B_t L_t)\, x_t + B_t m_t + c_t \qquad (3.3.41)$$

The point is that A_t, B_t and c_t are themselves random, so that calculation of variances becomes doubly complicated. Some relief is afforded by the fact that, at each time instant, x_t is independent of A_t, B_t and c_t, but even so, the complexity is too great to warrant a derivation here.

3.4 Computer Code :

In Appendix B, we code the solution to our stochastic control problem and the sensitivity equations we have derived in this chapter. More precisely, we code Equations (3.3.1) - (3.3.11) and (3.3.16) - (3.3.39). Though all the quantities represented in these equations are not printed out, they are all used in various intermediate calculations, and so can easily be made available by minor alterations in the program if the user needs them. The program does not contain sensitivity equations for $\sigma \in \Sigma_c$, Σ_{Bc}, Σ_{Ac}. Since this program was used for a specific application it also has a particular specification for the target sequence $\{ \tilde{x}_t \}$ which can again be altered by the user. No sequence $\{ \tilde{u}_t \}$ was needed for this application because we used $R = 0$. The user must provide both target sequences, the values for the Q and R matrices, the values of the means and covariances of A, B and c, and the time horizon N.

3.5 Conclusion :

Now that we have derived the relevant equations let us see how we can use them in analysing a specific model. For this we choose a small econometric model of the U.S. economy and analyse it in the next chapter.

CHAPTER 4

SENSITIVITY ANALYSIS

4.1 Introduction :

In this chapter, we shall use the equations we derived in the previous chapter to analyse the sensitivity of a small macroeconomic model of the U.S. economy. We first describe the model, then recast the equations into the appropriate optimal control framework, and finally present some simulation results with a discussion of their interpretation. Let us begin in the next section with the model.

4.2 A Simple Macroeconomic Model :

We shall describe, in this section, an especially simple macroeconomic model of the U.S. economy. This model was developed and estimated by Andrew Abel [47] to analyse the relative effectiveness of monetary and fiscal policies in an optimal control framework.

It is based on real quarterly data covering the period from 1954/I to 1963/IV, which corresponds roughly to the period between the end of the Korean War and the beginning of heavy U.S. involvement in Vietnam. It is an extremely small model, consisting of only two endogenous target variables, consumption C_t and investment I_t, and two instruments, government expenditures E_t and the money supply M_t. We assume that, in the short run, government authorities can control E_t and M_t in real terms since prices do not change rapidly enough to seriously neutralize their actions. Over the time period covered by our data, the rate of inflation was low enough to make this assumption plausible.

This model is based on a closed economy. Desired consumption is a linear function of GNP, and the realized period-to-period adjustment in consumption is subject to a partial adjustment factor :

$$C_t = aC_{t-1} + bI_t + bE_t + d \qquad (4.2.1)$$

The structural equation for investment is based upon a modification of Samuelson's private consumption accelerator. We posit that the desired level of the capital stock is a linear function of consumption and that the realized adjustment of the capital stock is subject to a partial adjustment factor. Since gross investment, I_t, is defined as $K_t - (1 - D)\,K_{t-1}$, where D is the depreciation rate of the capital stock, we have

$$I_t = eC_t - (1 - D)eC_{t-1} + f\,I_{t-1} + g$$

In addition, we assume that the level of gross investment is linearly related to the money supply in order to capture some of the effects of interest rates upon investment :

$$I_t = e'C_t - (1 - D)e'C_{t-1} + f'I_t + hM_t + g' \qquad (4.2.2)$$

The estimated reduced form equations corresponding to the structural equations are :

$$C_t = \underset{(0.0534)}{0.9266}\,C_{t-1} - \underset{(0.0916)}{0.0203}\,I_{t-1} + \underset{(0.1389)}{0.3190}\,E_t - \underset{(0.1863)}{0.4206}\,M_t - \underset{(25.7718)}{63.2386}$$

$$R^2 = 0.9958$$

$$D - W = 1.7084 \qquad (4.2.3)$$

$$I_t = \underset{(0.0781)}{0.1527}\, C_{t-1} + \underset{(0.1339)}{0.3806}\, I_{t-1} - \underset{(0.2031)}{0.0735}\, E_t + \underset{(0.2724)}{1.5389}\, M_t - \underset{(37.6899)}{210.8994}$$

$$R^2 = 0.8749$$

$$D\text{-}W = 1.7582 \qquad (4.2.4)$$

Note that each of these estimated equations has a high value of R^2. In addition, the Durbin-Watson statistic, although biased towards 2.0 because of the lagged endogenous variable, does not suggest significant serial correlation in either equation. The figures in parentheses are the corresponding standard errors.

4.3 Conversion into Optimal Control Framework :

Let us recast the reduced form equations in the previous section into state variable form. We shall write the model as a first-order linear vector difference equation with random coefficients :

$$x_{t+1} = A_t\, x_t + B_t\, u_t + c_t \qquad (4.3.1)$$

where

$$x_t = \begin{bmatrix} C_t \\ I_t \end{bmatrix}$$

$$u_t = \begin{bmatrix} E_{t+1} \\ M_{t+1} \end{bmatrix}$$

Note that $u_t = \begin{bmatrix} E_{t+1} \\ M_{t+1} \end{bmatrix}$ and not $\begin{bmatrix} E_t \\ M_t \end{bmatrix}$.

This is a small difference in the approach of control theorists and econometricians and is merely a matter of definition. Both refer to the policy variable that must be used to directly influence the state at time (t+1).

The coefficients of the various variables in the reduced form equations give us the respective means of the random matrices A_t, B_t and the random vector c_t. We have :

$$A_t = A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 0.9266 & -0.0203 \\ 0.1527 & 0.3806 \end{bmatrix}$$

$$B_t = B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 0.3190 & 0.4206 \\ -0.0735 & 1.5389 \end{bmatrix}$$

$$c_t = c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -63.2386 \\ -210.8994 \end{bmatrix}$$

The covariance matrices are defined by the convention in Chapter 2. These are obtained from the standard errors of the various random variables. The square of each standard error, that is the number in parentheses under each coefficient in Eqs. (4.2.3) - (4.2.4) gives the variance of the corresponding variable. Thus the diagonal entries of Σ_A are the variances of a_{11}, a_{21}, a_{12} and a_{22} in that order. The off-diagonal entries, the covariances, we somewhat arbitrarily set to zero. (Ignoring the covariances will usually tend to overestimate the size of the model's forecast errors. The majority of the estimated covariances are usually negative and cancel part of the variance in each coefficient.

Ignoring the covariances thus tends to overemphasize the degree of fluctuation in the coefficients.) All the covariance matrices are constant.

$$\Sigma_A = \text{diag}\ [\ \text{var}(a_{11}),\ \text{var}(a_{21}),\ \text{var}(a_{12}),\ \text{var}(a_{22})\ ]$$

$$= \begin{bmatrix} .0029 & 0 & 0 & 0 \\ 0 & .0070 & 0 & 0 \\ 0 & 0 & .0084 & 0 \\ 0 & 0 & 0 & .0179 \end{bmatrix}$$

$$\Sigma_B = \text{diag}\ [\ \text{var}(b_{11}),\ \text{var}(b_{21}),\ \text{var}(b_{12}),\ \text{var}(b_{22})\ ]$$

$$= \begin{bmatrix} .0193 & 0 & 0 & 0 \\ 0 & .0412 & 0 & 0 \\ 0 & 0 & .0347 & 0 \\ 0 & 0 & 0 & .0742 \end{bmatrix}$$

$$\Sigma_c = \text{diag}\ [\ \text{var}(c_1),\ \text{var}(c_2)\ ]$$

$$= \begin{bmatrix} 664.1908 & 0 \\ 0 & 1420.5286 \end{bmatrix}$$

We also need to define the values of the cross-covariance matrices Σ_{BA}, Σ_{Bc}, Σ_{Ac}. The estimation procedure used in Abel's paper does not provide us with estimates of these covariances, so here again we shall arbitrarily set them all equal to zero. This will also help a little in reducing the complexity of the various equations we have derived. We have, therefore :

$$\Sigma_{BA} = 0$$

$$\Sigma_{Ac} = 0$$

$$\Sigma_{Bc} = 0$$

At this point, we have completely specified the linear, random coefficient structure of the economic system in state variable form. To analyse the system in an optimal control framework, we need to specify a cost criterion.

$$J = \frac{1}{2} E \{ \sum_{t=0}^{N-1} [(x_t - \tilde{x}_t)' Q(x_t - \tilde{x}_t) + (u_t - \tilde{u}_t) R(u_t - \tilde{u}_t)] + (x_N - \tilde{x}_N)' Q (x_N - \tilde{x}_N) \}$$

We need to choose suitable values for the targets $\{\tilde{x}_t\}$, $\{\tilde{u}_t\}$ $t = 0,1,\ldots,N$ and specify the weighting matrices Q, R and the time horizon N. Following Abel, we examine the historical growth rates for consumption and investment over the period of estimation, 1954/I to 1963/IV, which turn out to be 0.91 % and 1.14 % per quarter respectively. With these in mind, we select target growth rates of 1.25 % per quarter for both C_t and I_t. Mathematically,

$$\tilde{x}_t = (1.0125)^t x_0 \qquad t = 0,1,2, \ldots , N$$

We shall restrict our choices for Q to diagonal matrices for the purpose of the analysis. We shall use the following five values for the Q matrix to compare different solutions.

$$Q = \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix} \triangleq \text{diag}\ (10,1) \triangleq (10,1) \quad \text{for simplicity}$$

$$Q = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \triangleq \text{diag}\ (2,1) \triangleq (2,1)$$

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \triangleq \text{diag}\ (1,1) \triangleq (1,1)$$

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \triangleq \text{diag}\ (1,2) \triangleq (1,2)$$

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix} \triangleq \text{diag}\ (1,10) \triangleq (1,10)$$

Henceforth the notation (10,1), (2,1) etc. will be used to denote the diagonal entries of diagonal Q matrices. We shall use this simplified notation especially when we present the simulation results.

We choose the R matrix to be zero throughout to simplify the analysis.

$$R = 0$$

Since R is chosen to be zero, we do not need to specify the targets $\{u_t\}$. The cost criterion is reduced to :

$$J = \frac{1}{2} E \left\{ \sum_{t=0}^{N} (x_t - \tilde{x}_t)' \, Q \, (x_t - \tilde{x}_t) \right\}$$

After doing a few simulations, it was decided that N = 15 would be large enough for the analysis without incurring too great a cost for the simulations.

The last item that needs to be specified is the initial conditions. From the historical record we find that

$$x_o = \begin{bmatrix} C_o \\ I_o \end{bmatrix} = \begin{bmatrix} 362 \\ 89 \end{bmatrix}$$

The units used are billions of dollars. E_t and M_t, the instruments, also have the same units. Note that $x_o = \tilde{x}_o$ by definition.

This completes the statement of the problem. In the next section, we present some simulation results.

4.4 Interpretation and Discussion of Results :

We shall now present, in the form of graphs and tables, some simulation results describing the behaviour of our econometric model in an optimal control framework. In this section we shall analyse some of these results and leave others for future research.

First, some general observations. As with other tracking problems, this problem can be split into one part that helps to regulate the state and another that helps it to track the desired trajectory and cancel any additive driving terms. We see that, in the event that all the covariance matrices are zero, the optimal control tracks perfectly. This is seen from the uppermost curves in Figs. 1 and 2. This is to be expected since R = 0 and there is no constraint on the control energy expended in the process. Also, in our problem, $x_o = \tilde{x}_o$, so there is no initial error. This deterministic solution is also the certainty-equivalence solution [ ], and we observe that the certainty-equivalence principle does not hold.

In the stochastic case, with Σ_A, Σ_B and Σ_c nonzero, we must first understand what it means for the state to track the desired trajectory. Since A,B and c are all random so are x_t and u_t (though the gain L_t and the correction cum tracking term m_t are deterministic). The control attempts to minimize the mean square error of the state trajectory which means it tries to keep the mean of the error plus the variance of the error small. In other words, there is a trade-off between keeping the average state close to the desired trajectory and keeping the variance of the error low. In general, therefore, we shall find that the average state evolution does not track perfectly. This is so even though R = 0. In Figs. 1 and 2, we have plotted the means of the state trajectories for the different values of Q. We see here that these mean trajectories fall short of the perfect certainty-equivalent trajectory. Of course, the actual trajectory we would get from any stochastic simulation would be different each time since we would have different realizations of A_t, B_t and c_t - this is true for both the state and control variables.

The certainty-equivalent solution for R = 0 simplifies to :

$$L_t = -\bar{B}^{-1}\bar{A} \tag{4.4.1}$$

$$K_t = Q \tag{4.4.2}$$

$$m_t = -\bar{B}^{-1}(\bar{c} + \bar{Q}^{-1} p_{t+1}) = -\bar{B}^{-1}(\bar{c} - \tilde{x}_{t+1}) \tag{4.4.3}$$

$$p_t = -Q\,\tilde{x}_t \tag{4.4.4}$$

$$g_t = \frac{1}{2}\,\tilde{x}_t'\,Q\,\tilde{x}_t \tag{4.4.5}$$

$$J^* = 0 \tag{4.4.6}$$

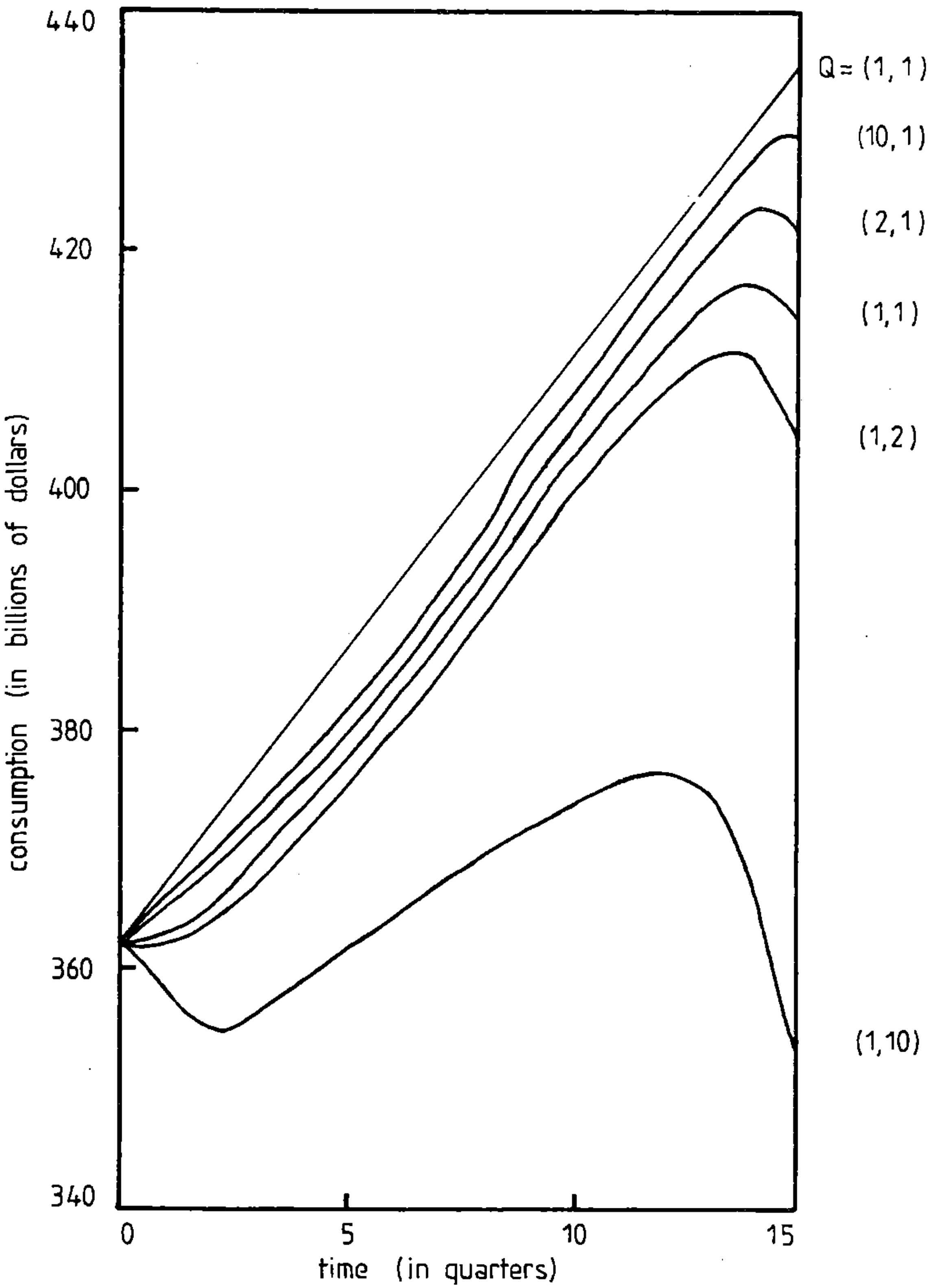


Fig.1. Consumption vs. time, Eq. (3.3.11), for N = 15. For the C.E. case, all covariance matrices are set equal to zero. The C.E. curve is identical with the desired trajectory.

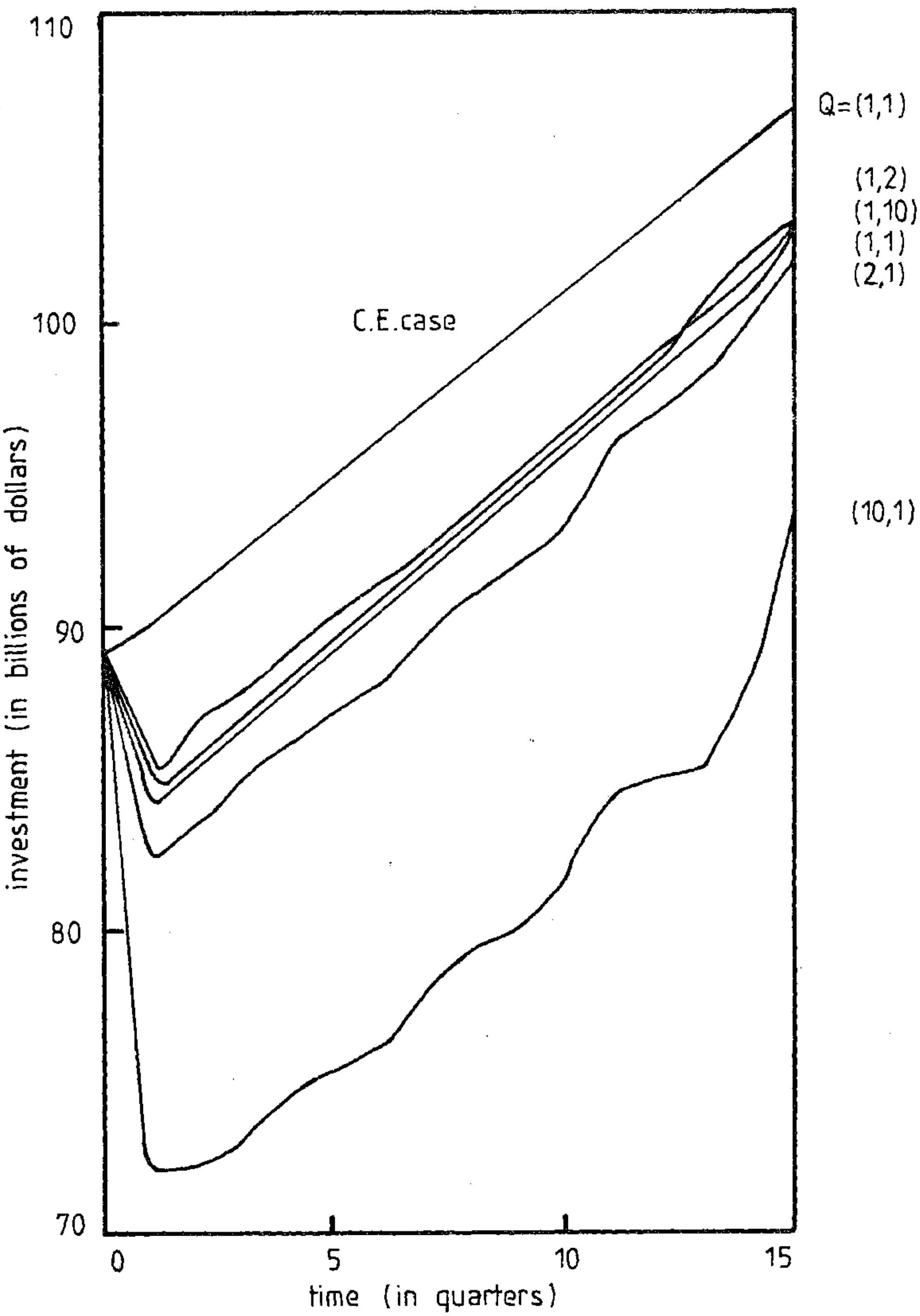


Fig.2. Investment vs. time, Eq. (3.3.11), for N = 15. For the C.E. case, all covariance matrices are set equal to zero. The C.E. curve is identical with the desired trajectory.

Substituting these equations into the mean of the state equation we get,

$$
\begin{aligned}
x_{t+1} &= \bar{A}\bar{x}_t + \bar{B}L_t\bar{x}_t + \bar{B}m_t + \bar{c} \\
&= (\bar{A} - \bar{B}.\bar{B}^{-1}\bar{A})\ \bar{x}_t - \bar{B}.\bar{B}^{-1}\bar{c} + \bar{B}.\bar{B}^{-1}\ \tilde{x}_{t+1} + \bar{c} \\
&= \tilde{x}_{t+1}, \text{ as expected.} \qquad (4.4.7)
\end{aligned}
$$

Note that the gain L_t, the additive term m_t and the average state $\bar{x}_t$ and the average control law $\bar{u}_t$ are all independent of the choice of Q. This is why we need not specify the value of Q for the certainty-equivalence curves in Figs. 1 and 2. The different curves for the stochastic case are identified by the corresponding values of Q.

The gain L_t in Eq. 4.4.1 serves to cancel the coefficient matrix A which it does exactly in the mean case when $A = \bar{A}$, whereas the term m_t cancels the additive exogenous term c as well as forces the state to track the target, both of which again are done exactly in the mean case. Note that the optimal cost J^* is zero (Eq. 4.4.6), the absolute minimum of J, because R = 0 and because the state tracks perfectly. J^* is also independent of Q.

Let us now examine the stochastic case more closely. Our first observation of the simulation results is that the regulator part of the problem viz. L_t and K_t, is well behaved. We have plotted in Fig. 3 the certainty-equivalent and the stochastic K_t for Q = (1,1). There are four graphs, one for each element of K_t. Since K_t is symmetric two of the graphs representing the off-diagonal terms are identical. We plot, in a

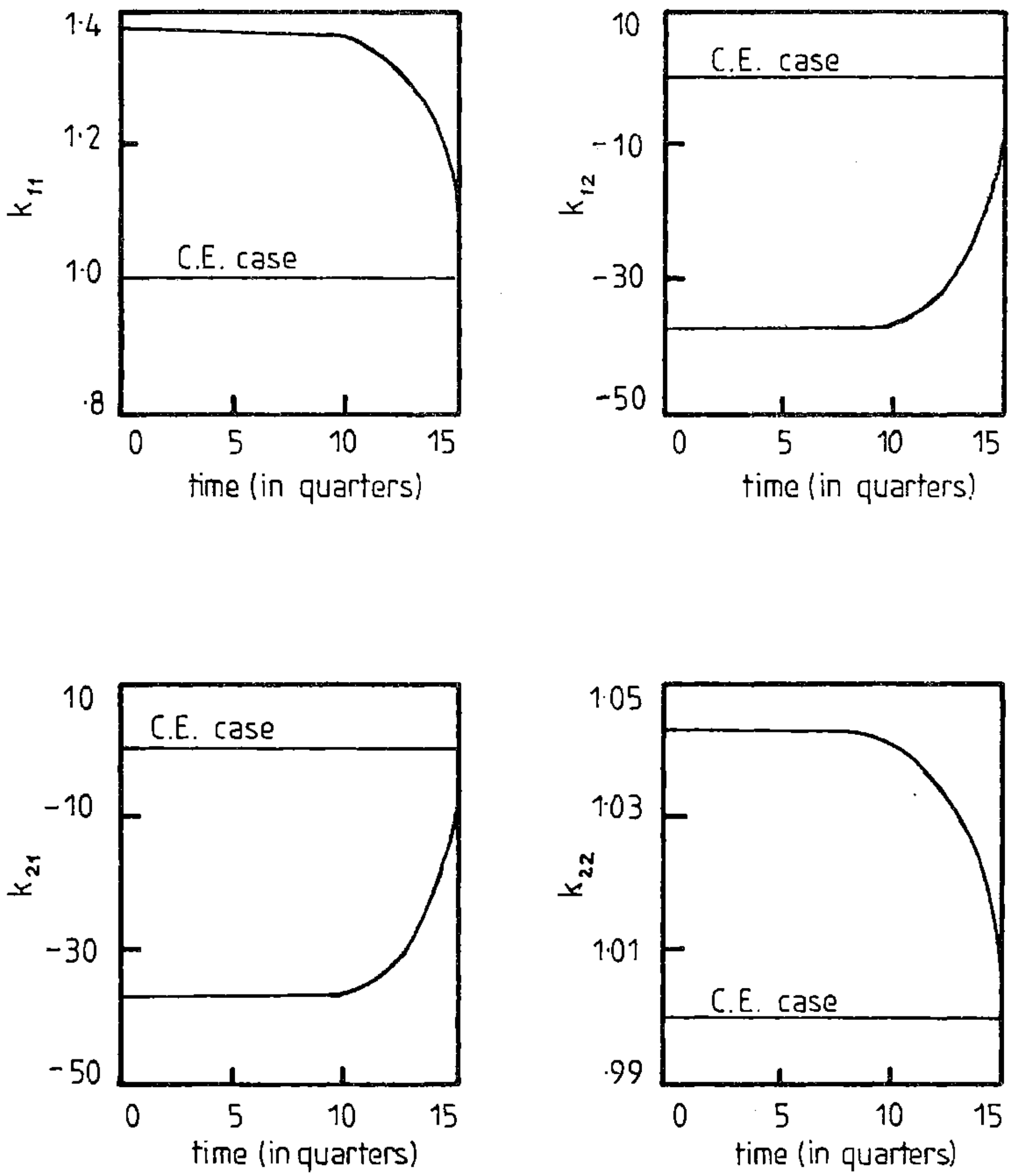


Fig.3. Solution of Riccati-like equation, Eq.(A.8). K_t is a symmetric 2 x 2 matrix. Q = (1,1), N = 15.

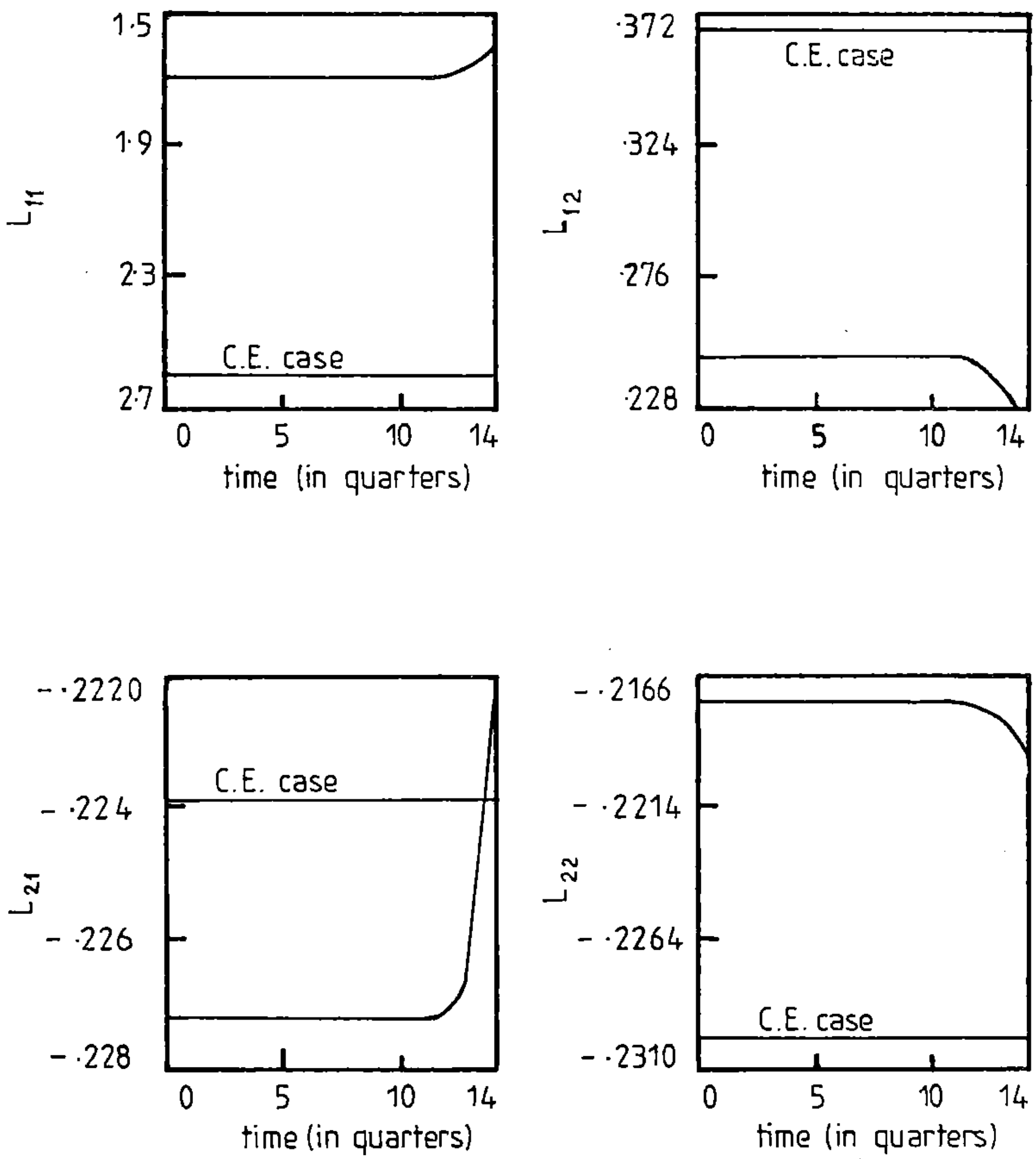
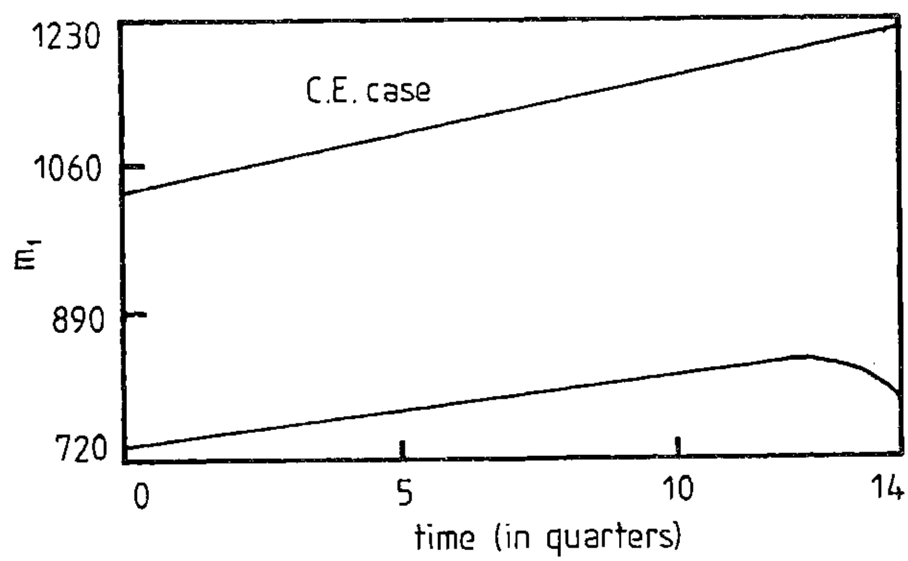


Fig.4. Graph of gain matrix L_t vs. time, Eq. (A.5), for N = 15, Q=(1,1).



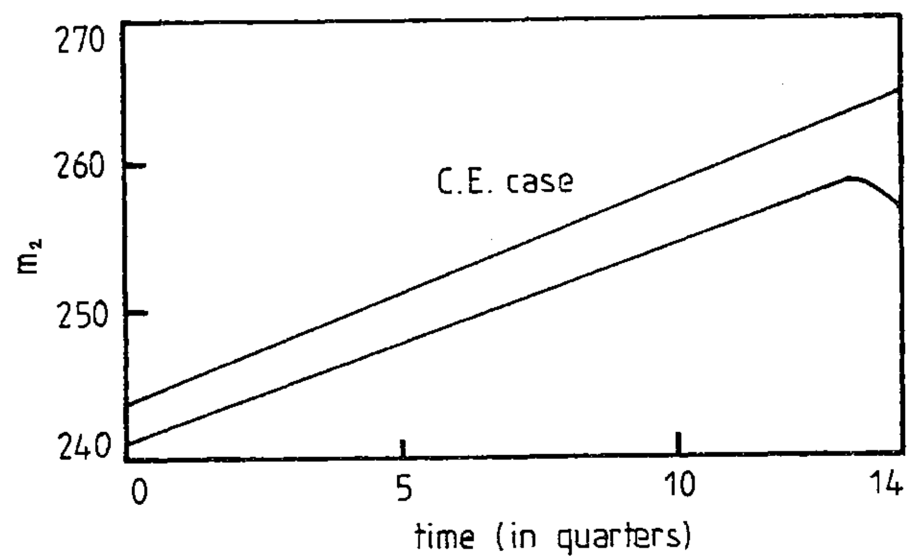


Fig.5. Additive term $\underline{m}_t$ vs. time, Eq. (A,6), for N = 15. Q = (1,1) for all curves. For the C.E. case all covariance matrices are set equal to zero.

similar way, L_t in Fig. 4, again for Q = (1,1). The certainty-equivalent value of L_t in this figure is given by Eq. (4.4.1). Both quantities soon reach a steady state, seen backward in time. The correction terms m_t in Fig. 5 keeps growing because it has to track $\tilde{x}_t$ in addition to cancelling the exogenous term c_t. The optimal cost also keeps growing. However, since K_t is steady initially, we can deduce that the regulator component of the cost, $\frac{1}{2} x_o' K_o x_o$, settles to a steady state. The tracking error naturally keeps accumulating and this makes the cost grow. The behaviour of K_t (Fig. 3) leads us to the conclusion that the uncertainties in the problem are within the uncertainty threshold (even though we do not know exactly what the threshold is). We shall find later that even if Σ_A is multiplied by a scale factor of 30, K_t does not blow up. This seems reasonable when one inspects the numerical values of A, Σ_A, B, Σ_B which are all fairly small. The elements of Σ_A, Σ_B in particular are all << 1.

$$K_t = Q + [\bar{A}'K_{t+1}\bar{A} + \sum_{k,\ell} tr(K_{t+1}\Sigma_A^{k\ell})E_{k\ell}] - [\bar{A}'K_{t+1}\bar{B}].[\bar{B}'K_{t+1}\bar{B} + \sum_{k,\ell} tr(K_{t+1}\Sigma_B^{k\ell})E_{k\ell}]^{-1}.[\bar{B}'K_{t+1}\bar{A}]$$

Note that $\Sigma_{BA} = 0$ in our problem. Since $Q \geqslant 0$, $\Sigma_A \geqslant 0$, $\Sigma_B \geqslant 0$, the structure of the equation tells us to expect $K_t > Q$ or equivalently, $\| K_t \| \geqslant \| Q \|$ where $\| M \| \equiv (\det M)^{1/2}$. This is in fact borne out by the simulation results. In Table 1, we present some norms of K_t for different Q. This demonstrates that the steady state "value" of K_t in the stochastic case is greater than that in the certainty-equivalent case. This confirms our intuition that we need more "force" when there is

Q	$\| Q \|$	$\| K_o \|$
(10,1)	3.16	3.64
(2,1)	1.41	1.62
(1,1)	1.00	1.18
(1,2)	1.41	1.75
(1,10)	3.16	4.74

Table 1. Comparison of norms of Q and corresponding K_o.

uncertainty. The end point constraint $K_N = Q$ forces K_t to come down to the C.E. value at N (Fig. 3). Physically, K_t represents a sort of cumulative weighting matrix which incorporates both the present error at time t as well as the propagation of this error as t progresses to N. When $t << N$ we would expect the slope of K_t to be relatively horizontal since the future error weighs about the same for small t far from N. However, as t gets close to N, K_t is determined more by the present error since the propagation error gets smaller, so that it begins to fall to Q, till at $t = N$, there is no future and K_N exactly equals the present error weighting matrix Q. We have ignored here the effects of non-zero R. The steady-state value is greater in the uncertain case because we are minimizing the mean square error, as opposed to just the mean error so that there is greater propagation of the present error and $K_t > Q$. This description can quite easily be extended to the case of time-varying Q's. Note also that if $\Sigma_{BA} \neq 0$, then the propagation of the uncertainty in the error is somewhat reduced, since B and A are now correlated and the control can make use of this additional information. However, because of the restrictions placed by the various correlation coefficients, the effects of uncertainty cannot be completely nullified. This is also supported by the mathematics.

The gain L_t, Fig. 4, follows the behaviour of K_t in a mathematical sense. It is steady initially and, as t approaches N, it moves away from the steady-state value just as K_t does. Again, it basically attempts to minimize the mean square error instead of just the mean error. Note that L_t represents only the regulator part of the control and is totally independent of the targets and the driving term c. The scalar case provides some insight into its behaviour.

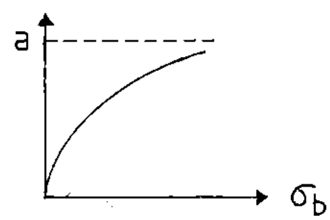
$$\ell_t = -\frac{\bar{a}\bar{b}}{\bar{b} + \sigma_b^2} = -\frac{\bar{a}}{\bar{b}}\left(\frac{1}{1 + \sigma_b^{\,2}/\bar{b}^2}\right)$$

Note that in the scalar case ℓ_t is constant even in the stochastic case. Also, note that ℓ_t decreases in absolute value as σ_b increases, other things remaining the same.

$$x_{t+1} = (a + b\ell_t)\, x_t + bm_t + c$$

$$= a \cdot \frac{\sigma_b^{\,2}}{b^2 + \sigma_b^{\,2}} \cdot x_t + b\, m_t + c$$

The coefficient of x_t has the following approximate behaviour :



when $\sigma_b = 0$, the coefficient vanishes, thereby keeping x_{t+1} close to zero, as required by the regulator. The optimal gain is chosen so as to minimize $E(a + b\ell_t)^2$.

i.e. $\frac{d}{d\ell_t} E(a + b\ell_t)^2 = 0$

Therefore $2\,\bar{a}\bar{b} + (\bar{b}^2 + \sigma_b^{\,2})\ell_t = 0$

Therefore $\ell_t = -\frac{\bar{a}\bar{b}}{\bar{b}^2 + \sigma_b^{\,2}}$ as required

This short derivation merely shows, from a different perspective, that ℓ_t does the stochastically optimal thing. The vector case behaves essentially in the same way though the mathematics is a trifle opaque because the appropriate quantity to minimize for the one-step optimal gain is $E[(A + BL_t)'\ K_{t+1}(A + BL_t)]$, because K_{t+1} embodies the correct cumulative weighting at time t.

The term p_t is again essentially a mathematical entity like K_t. The equation for p_t is :

$$\begin{aligned} p_t &= -Q\tilde{x}_t + \overline{A'K_{t+1}c} + \bar{A}'p_{t+1} + (\overline{A'K_{t+1}B}).m_t \\ &= -Q\tilde{x}_t + \overline{A'K_{t+1}c} + \bar{A}'p_{t+1} \\ &\qquad - (\overline{A'K_{t+1}B}).(\overline{B'K_{t+1}B})^{-1}(\overline{B'K_{t+1}c} + \bar{B}'p_{t+1}) \end{aligned}$$

$$p_N = -Q\,\tilde{x}_N$$

Its behaviour can be understood in analogy with that of K_t. It has two basic functions. The first is its role in providing a correction term to cancel the exogenous term c and the second to provide a cumulative weighted measure of the desired trajectory. To understand these roles more clearly let us look at them separately. First let us assume that the desired trajectory is zero i.e. $\tilde{x}_t = 0$ for all t. Then,

$$p_t = \bar{A}'K_{t+1}\bar{c} + \bar{A}'p_{t+1} - (\bar{A}'K_{t+1}\bar{B})(\overline{B'K_{t+1}B})^{-1}(\bar{B}'K_{t+1}\bar{c} + \bar{B}'p_{t+1})$$

$$p_N = 0$$

We note here that the behaviour of p_t is directed towards c. At t=N, $p_N = 0$ because c_N cannot affect the optimal cost. Now let us assume that c=0, we get,

$$p_t \quad = \quad - Q\tilde{x}_t + \bar{A}'p_{t+1} - (\bar{A}'K_{t+1}\bar{B}) \cdot (\overline{B'K_{t+1}B})^{-1} \cdot (\bar{B}'p_{t+1})$$

$$p_N \quad = \quad - Q\,\tilde{x}_N$$

This shows how at t=N, p_N represents a weighted target and for earlier t, how it incorporates both the present target in the term $- Q\tilde{x}_t$ and the propagation of this in the future as well as future targets in the rest of the equation. In the general case when $R \neq 0$, p_t also includes the weighted control targets in the term $- R\tilde{u}_t$.

Just as K_{t+1} gives us the gain L_t so p_{t+1} (in combination with K_{t+1}) gives us the additive term m_t, which embodies the two roles of p_t explicitly in the control. The first role is to act as a correction term to offset the exogenous vector c. This function is independent of the regulator and tracking parts of the problem or, in other words, it is needed in both. The second function is tracking. It is responsible for making the state track the desired trajectory. These two objectives are clearly observable in the equation for m_t.

$$m_t \quad = \quad - [\ \bar{B}'K_{t+1}\bar{B} + \sum_{k,\ell} tr(K_{t+1}\Sigma_B^{k\ell})E_{k\ell}]^{-1}[\bar{B}'K_{t+1}\bar{c} + \bar{B}'p_{t+1}]$$

We see from Fig. 5 that the behaviour of m_t shows an approximately steady growth. Though the corrective component does reach a steady state the tracking component does not since the target itself grows with time. Its behaviour could also be understood in terms of the minimization of a suitable expression as we did for L_t. However, this is complicated by the fact that both K_{t+1} and p_{t+1} enter into it.

Now that we have some description of the behaviour of the various components of the problem we can better appreciate the behaviour of the control u_t and the state x_t.

The certainty-equivalent control u_t is given by :

$$u_t = L_t x_t + m_t = - \bar{B}^{-1}\bar{A} x_t - \bar{B}^{-1}(\bar{c} - \tilde{x}_{t+1})$$

and the certainty-equivalent x_t is :

$$x_t = \tilde{x}_t$$

This shows that u_t and x_t in the certainty-equivalent case must be approximately linear (since $\tilde{x}_t \simeq [1 + 0.0125_t]\ x_o$). This is borne out by Figs. 1-2 and Figs. 6-7. In the stochastic case we find that $\bar{u}_t$ tries to approach u_t^{CE} in the "middle", as we would expect. At this point it is useful to look at the mean values of the A and B matrices :

$$\bar{A} \simeq \begin{bmatrix} .93 & -.02 \\ .15 & .38 \end{bmatrix} \qquad \bar{B} \simeq \begin{bmatrix} .32 & .42 \\ -.07 & 1.53 \end{bmatrix} \qquad \bar{c} \simeq \begin{bmatrix} -63.24 \\ -210.90 \end{bmatrix}$$

and

$$\begin{bmatrix} \bar{C}_{t+1} \\ \bar{I}_{t+1} \end{bmatrix} = \begin{bmatrix} .93 & -.02 \\ .15 & .38 \end{bmatrix} \begin{bmatrix} \bar{C}_t \\ \bar{I}_t \end{bmatrix} + \begin{bmatrix} .32 & .42 \\ -.07 & 1.53 \end{bmatrix} \begin{bmatrix} \bar{E}_t \\ \bar{M}_t \end{bmatrix} - \begin{bmatrix} 63.24 \\ 210.90 \end{bmatrix}$$

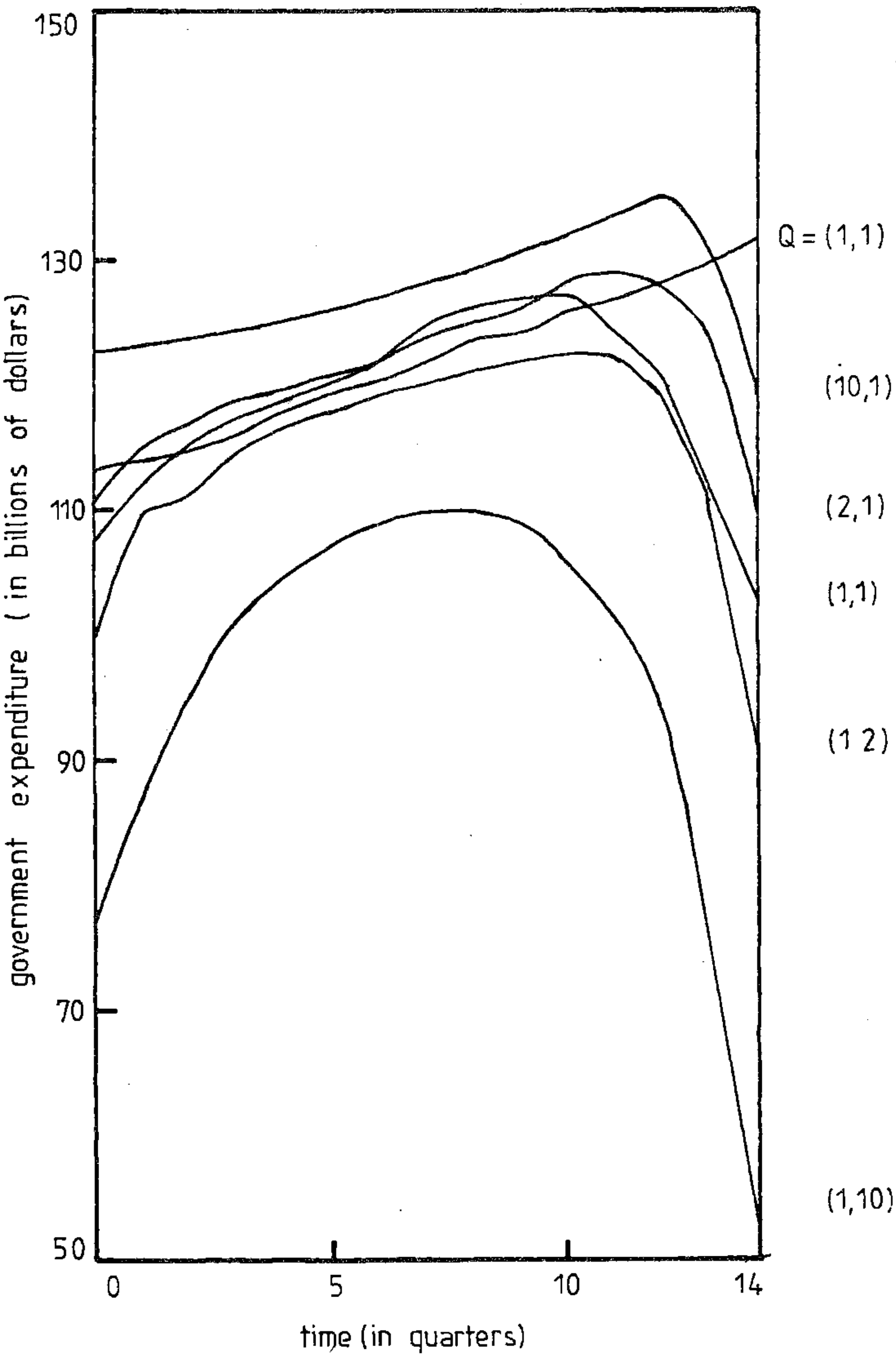


Fig.6. Government expenditure vs. time, Eq. (A.4), for N = 15. For the C.E. case, all covariance matrices are set equal to zero.

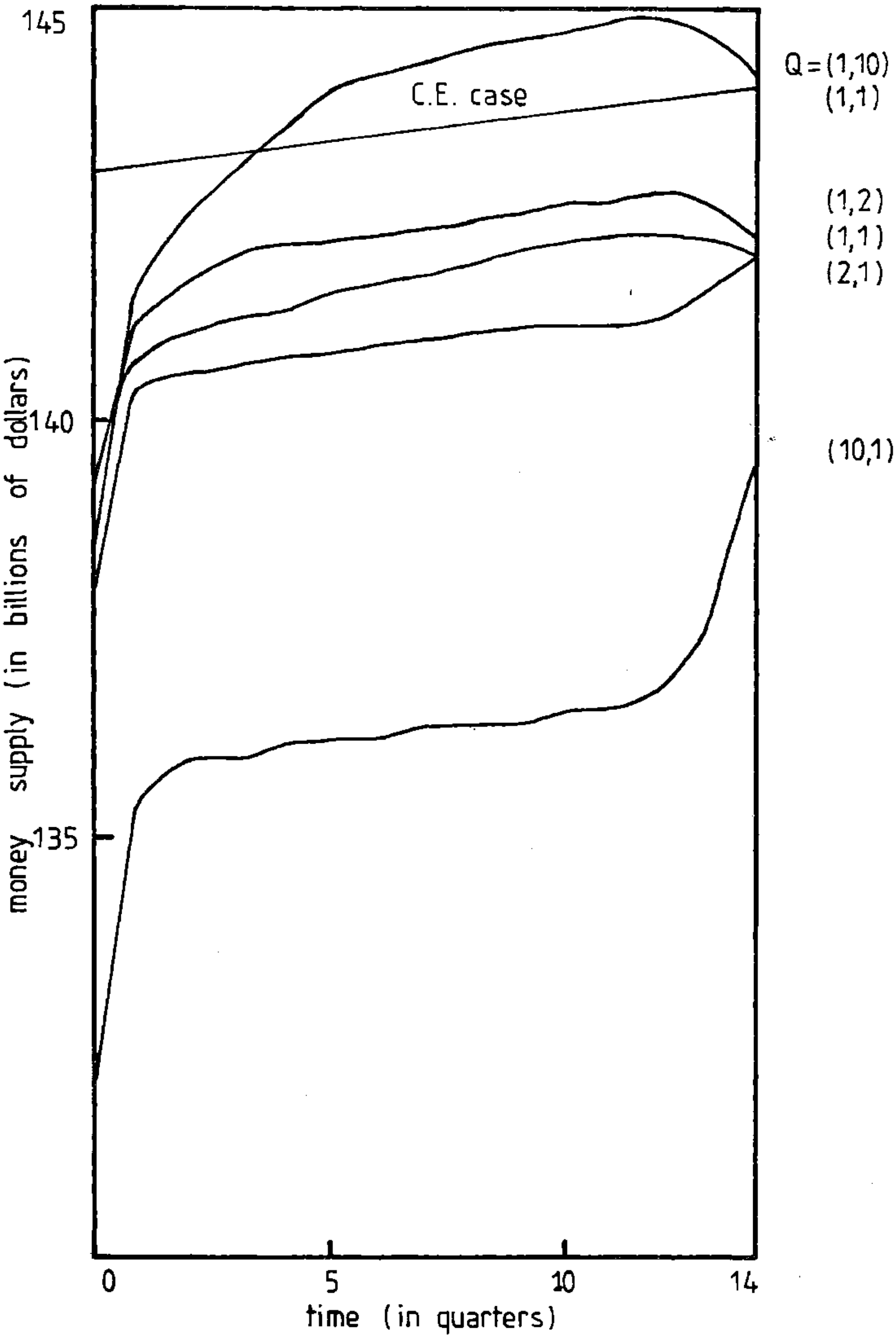


Fig.7. Money supply vs. time, Eq. (A.4), for N = 15. For the C.E. case, all covariance matrices are set equal to zero.

Looking at the relative values of the elements of $\bar{A}$, we see that average consumption $\bar{C}$ is essentially independent of average investment $\bar{I}$, though $\bar{I}$ does depend on $\bar{C}$. Also, owing to the relative values of the elements of B we see that the average government expenditure $\bar{E}_t$ does not really affect investment $\bar{I}_{t+1}$. However, $\bar{E}_t$ influences $\bar{C}_{t+1}$ directly which in turn influences $\bar{I}_{t+2}$, so that the effect of average government expenditure on average investment is experienced two periods later. We note also that both the instruments can influence consumption.

In the stochastic case we see that as the relative weighting of consumption and investment in the weighting matrix Q changes in favour of one or the other, the corresponding state approaches the target more closely, as one would expect. In Fig. 1, the perfect C.E. case is at the top. Below this comes the curve corresponding to Q = diag (10,1). As the relative weighting of consumption decreases to Q = (2,1) the mean consumption trajectory drops even further down. This trend continues till Q = (1,10). In Fig. 2, we observe exactly the opposite. Q = (1,10) represents the case for which investment tracks most closely since the relative weight of I is highest here and it gets progressively worse as we go to Q = (10,1).

Finally, the optimal cost J^* needs to be considered. We find that it can also be divided into two parts : the regulator part and the tracking part. The regulator part comes from the term $\frac{1}{2} x_o' K_o x_o$, which is the same as the cost for the corresponding regulator problem. The additional terms $p_o' x_o$ and g_o explain the tracking part of the cost.

The term 'g_o' represents a residual type cost (the dynamic counterpart of the constant term 'c' in the minimization of a quadratic function $ax^2 + bx + c$). We note also that J^* increases as Σ_A increases, since the control becomes less and less capable of controlling the system effectively, (Fig. 8).

Let us now look at the sensitivities of some of the parameters. To keep things simple we shall only look at the sensitivities of the diagonal elements of Σ_A and Σ_B. Note that σ_{11} = var (a_{11}), σ_{22} = var (a_{21}), σ_{33} = var (a_{12}), σ_{44} = var (a_{22}) when $\sigma_{ij} \varepsilon \Sigma_A$. Similarly, when $\sigma_{ij} \varepsilon \Sigma_B$, σ_{11} = var (b_{11}), σ_{22} = var (b_{21}), σ_{33} = var (b_{12}) and σ_{44} = var (b_{22}). For convenience we shall denote var (a_{ij}) by $\sigma(a_{ij})$ and var (b_{ij}) by $\sigma(b_{ij})$. The relative sensitivities corresponding to different Q matrices are given in Table 2 and are then ranked in Table 3. We do the same with the absolute sensitivities in Tables 4 and 5.

Our first observation is that none of the parameters are overly sensitive. We note that the highest relative sensitivity is only .3 or 30%. We can call a relative sensitivity of 1 or 100% high because that implies a variation of a magnitude commensurate with the actual value. Judging by this standard sensitivities of .3 or less are negligible. Thus, in a general sense, this model is quite insensitive to variations in parameter variances. In other words, at least for this model, this method of analysing sensitivity does not yield much useful information, besides the fact that the model is insensitive and therefore reasonably reliable.

Parameters			Q:(10,1)	Q:(2,1)	Q:(1,1)	Q:(1,2)	Q:(1,10)
Σ_A	σ_{11}	$\sigma(a_{11})$	.171	.107	.075	.050	.016
	σ_{22}	$\sigma(a_{21})$	.037	.099	.129	.152	.169
	σ_{33}	$\sigma(a_{12})$	.058	.041	.028	.017	.005
	σ_{44}	$\sigma(a_{22})$	.012	.038	.048	.053	.051
Σ_B	σ_{11}	$\sigma(b_{11})$	.216	.127	.082	.048	.010
	σ_{22}	$\sigma(b_{21})$	.047	.117	.140	.145	.105
	σ_{33}	$\sigma(b_{12})$	.250	.168	.121	.081	.030
	σ_{44}	$\sigma(b_{22})$	.054	.155	.206	.249	.316

Table 2. Relative sensitivities

Parameters			C.E.Q:(1,1)	Q:(10,1)	Q:(2,1)	Q:(1,1)	Q:(1,2)	Q:(1,10)
Σ_A	σ_{11}	$\sigma(a_{11})$	1176272	13405233	2860309	1515515	1646233	2098791
	σ_{22}	$\sigma(a_{21})$	1176272	1354913	1234413	1209006	2362638	10445811
	σ_{33}	$\sigma(a_{12})$	71100	1537704	371865	191135	195192	217415
	σ_{44}	$\sigma(a_{22})$	71100	155412	160615	152658	280599	1086178
Σ_B	σ_{11}	$\sigma(b_{11})$	111247	2501325	503951	244319	233663	199524
	σ_{22}	$\sigma(b_{21})$	111247	252895	217093	194130	332498	962598
	σ_{33}	$\sigma(b_{12})$	154526	1609666	370167	199918	222138	322734
	σ_{44}	$\sigma(b_{22})$	154526	162709	159488	159089	317734	1603107

Table 4. Absolute sensitivities

Q:(10,1)	Q:(2,1)	Q:(1,1)	Q:(1,2)	Q:(1,10)
$\sigma(b_{12})$	$\sigma(b_{12})$	$\sigma(b_{22})$	$\sigma(b_{22})$	$\sigma(b_{22})$
$\sigma(b_{11})$	$\sigma(b_{22})$	$\sigma(b_{21})$	$\sigma(a_{21})$	$\sigma(a_{21})$
$\sigma(a_{11})$	$\sigma(b_{11})$	$\sigma(a_{21})$	$\sigma(b_{21})$	$\sigma(b_{21})$
$\sigma(a_{12})$	$\sigma(b_{21})$	$\sigma(b_{12})$	$\sigma(b_{12})$	$\sigma(a_{22})$
$\sigma(b_{22})$	$\sigma(a_{11})$	$\sigma(b_{11})$	$\sigma(a_{22})$	$\sigma(b_{12})$
$\sigma(b_{21})$	$\sigma(a_{21})$	$\sigma(a_{11})$	$\sigma(a_{11})$	$\sigma(a_{11})$
$\sigma(a_{21})$	$\sigma(a_{12})$	$\sigma(a_{22})$	$\sigma(b_{11})$	$\sigma(b_{11})$
$\sigma(a_{22})$	$\sigma(a_{22})$	$\sigma(a_{12})$	$\sigma(a_{12})$	$\sigma(a_{12})$

Table 3. Ranking of parameters in order of decreasing relative sensitivity

Q:(10,1)	Q: (2,1)	Q: (1,1)	Q: (1,2)	Q:(1,10)
$\sigma(a_{11})$	$\sigma(a_{11})$	$\sigma(a_{11})$	$\sigma(a_{21})$	$\sigma(a_{21})$
$\sigma(b_{11})$	$\sigma(a_{21})$	$\sigma(a_{21})$	$\sigma(a_{11})$	$\sigma(a_{11})$
$\sigma(b_{12})$	$\sigma(b_{11})$	$\sigma(b_{11})$	$\sigma(b_{21})$	$\sigma(b_{22})$
$\sigma(a_{12})$	$\sigma(a_{12})$	$\sigma(b_{12})$	$\sigma(b_{22})$	$\sigma(a_{22})$
$\sigma(a_{21})$	$\sigma(b_{12})$	$\sigma(b_{21})$	$\sigma(a_{22})$	$\sigma(b_{21})$
$\sigma(b_{21})$	$\sigma(b_{21})$	$\sigma(a_{12})$	$\sigma(b_{11})$	$\sigma(b_{12})$
$\sigma(b_{22})$	$\sigma(a_{22})$	$\sigma(b_{22})$	$\sigma(b_{12})$	$\sigma(a_{12})$
$\sigma(a_{22})$	$\sigma(b_{22})$	$\sigma(a_{22})$	$\sigma(a_{12})$	$\sigma(b_{11})$

Table 5. Ranking of parameters in order of decreasing absolute sensitivity

If we look at the variations in the sensitivity ranks as Q is changed, we find a reasonable pattern. When consumption is more heavily weighted than investment, we find that the parameters $\sigma(a_{11})$, $\sigma(a_{12})$, $\sigma(b_{11})$, $\sigma(b_{12})$ tend to be more sensitive, whereas when investment is more heavily weighted the parameters $\sigma(a_{21})$, $\sigma(a_{22})$, $\sigma(b_{21})$, $\sigma(b_{22})$ are more sensitive (Tables 3 and 5). This is as it should be as is evinced by the positions of these parameters in the covariance matrices :

$$\Sigma_A = \left[\begin{array}{cc|cc} \sigma(a_{11}) & 0 & 0 & 0 \\ 0 & \sigma(a_{21}) & 0 & 0 \\ \hline 0 & 0 & \sigma(a_{12}) & 0 \\ 0 & 0 & 0 & \sigma(a_{22}) \end{array}\right]$$

$$\Sigma_B = \left[\begin{array}{cc|cc} \sigma(b_{11}) & 0 & 0 & 0 \\ 0 & \sigma(b_{21}) & 0 & 0 \\ \hline 0 & 0 & \sigma(b_{12}) & 0 \\ 0 & 0 & 0 & \sigma(b_{22}) \end{array}\right]$$

What happens in the sensitivity equations is that the above shown 2x2 blocks enter into the mathematics directly through the terms $P_k'\Sigma_A P_\ell$, $P_k'\Sigma_B P_\ell$. Since $\sigma(a_{11})$, $\sigma(a_{12})$, $\sigma(b_{11})$, $\sigma(b_{12})$ occupy the top left positions in these blocks they contribute to the error in the propagation of consumption and as consumption assumes a greater relative importance in the cost functional, these parameter variances become more sensitive. This is shown by the column of rankings under Q = (10,1) in Table 3. Exactly the same happens in the other direction with investment. The parameters $\sigma(a_{21})$, $\sigma(a_{22})$, $\sigma(b_{21})$, $\sigma(b_{22})$ occupy the bottom right positions in these blocks and thereby contribute to the error in investment, so that they become more sensitive as the relative weighting of investment increases. As we move from the column under Q = (10,1) to the column under Q = (1,10) from left to right in Tables 3 and 5, we find that the parameters $\sigma(a_{21})$, $\sigma(a_{22})$, $\sigma(b_{21})$, $\sigma(b_{22})$ move from the bottom of the columns gradually to the top when we get to Q = (1,10). This pattern also makes sense physically. When consumption is more important, one would expect the higher sensitivities to be with the first rows of A and B which parameters affect consumption directly. More explicitly

$$C_{t+1} = a_{11} C_t + a_{12} I_t + b_{11} E_t + b_{12} M_t + c_1$$

The other parameters a_{21}, a_{22}, b_{21}, b_{22} affect C_t only indirectly. The same is true for investment.

$$I_{t+1} = a_{21} C_t + a_{22} I_t + b_{21} E_t + b_{22} M_t + c_2$$

From this one would expect $\sigma(a_{21})$, $\sigma(a_{22})$, $\sigma(b_{21})$, $\sigma(b_{22})$ to be more sensitive as is borne out by the results.

Q = (2,1) seems to represent some sort of a "break-point" that weights consumption and investment in some "equitable" manner. Firstly, we find that the relative sensitivities at this value are all evenly distributed i.e. there is no priority in ranking in either group, $[\sigma(a_{11}), \sigma(a_{12}), \sigma(b_{11}), \sigma(b_{12})]$ or $[\sigma(a_{21}), \sigma(a_{22}), \sigma(b_{21}), \sigma(b_{22})]$.

See the column under Q = (2,1) in Table 3. If we increase the relative weight of consumption towards Q = (10,1), then we find elements of $[\sigma(a_{11}), \sigma(a_{12}), \sigma(b_{11}), \sigma(b_{12})]$ becoming more sensitive whereas if we decrease it towards Q = (1,1), (1,2) and (1,10), we find $[\sigma(a_{21}), \sigma(a_{22}), \sigma(b_{21}), \sigma(b_{22})]$ becoming more sensitive. Of course, since our data comes from only five Q matrices, we cannot have the exact break-point but we can say that it lies roughly near Q = (2,1). This also seems to be the Q that gives the lowest value for the optimal cost J^* scaled by the norm of the corresponding Q, as can be seen from Table 6. In addition to this, Table 7 indicates that $\|L_o\|$ is largest in the Q = (2,1) case. Of course, the certainty equivalent J^* equals zero and is lower than the above scaled J^*, and $\|L_o\|_{CE} = .682$ is also higher than $\|L_o\|$ for Q = (2,1). The fact that J^* is lowest for this Q means that this represents the minimum of J^* taken over all Q. Similarly, the fact that $\|L_o\|$ is highest seems to imply that the control is most forceful in this case. All this points to the fact that Q = (2,1) represents a special weighting matrix. The specific value of Q depends of course in some complicated way on the values of $\bar{A}$, $\bar{B}$ and Σ_A, Σ_B. However, the important point is that it gets closest to the certainty-equivalent case in some average way. It represents, in a certain sense, an "optimal" choice for Q.

As we increase Σ_A gradually, scaling the entire matrix Σ_A by factors of 1.1, 2, 6, 15 and 30 progressively, we find first that the optimal cost J^* increases (Fig. 8). This is reasonable physically since the system becomes increasingly difficult to control with increasing uncertainty. We find the other variables behaving reasonably too. For example, the

Q	$J^*/\|Q\|$
(10,1)	7.27
(2,1)	5.66
(1,1)	6.00
(1,2)	6.72
(1,10)	12.02
C.E.	0.00

Table 6. Normalised values of the optimal cost for different weighting matrices Q.

Q	$\|L_o\|$
(10,1)	.444
(2,1)	.470
(1,1)	.439
(1,2)	.387
(1,10)	.230
C.E.	.682

Table 7. Normed values of initial gain matrices for different weighting matrices Q.

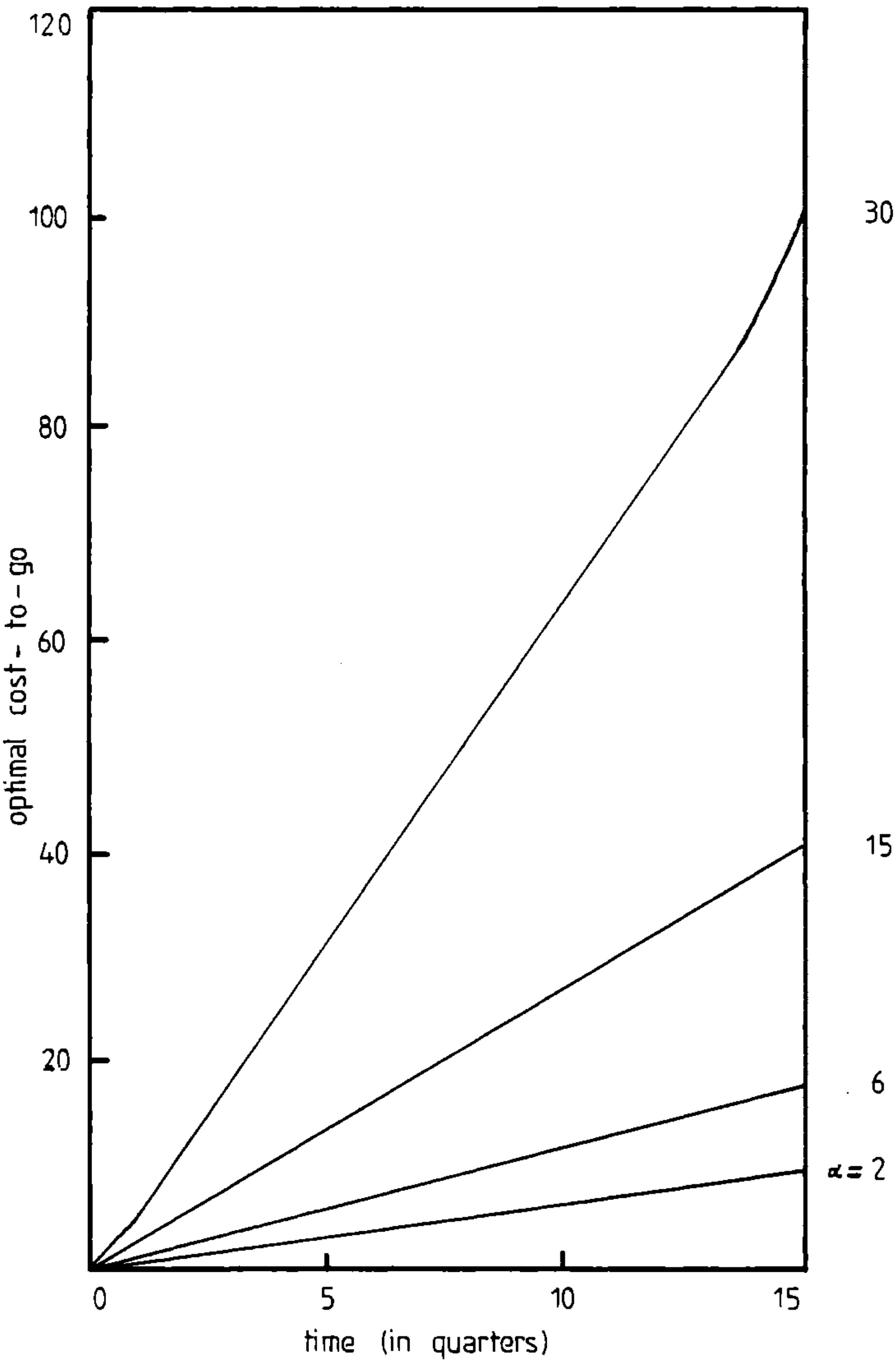


Fig.8. Optimal cost-to-go vs. time, Eq. (A.13), for N = 15.
α is the scale factor for the covariance matrix Σ_A.

quality of the state trajectory drops and we find in some sense a greater expenditure of control energy (Fig.9-10). The behaviour of the sensitivities does not show any useful regularities as can be seen by carefully studying Tables 8 and 9. Since the relative sensitivity is given by $\frac{\partial J^*}{\partial \sigma} \cdot \frac{\sigma}{J^*}$ and σ and J^* both increase, and the change in $\frac{\partial J^*}{\partial \sigma}$ itself is hard to guess, we are left without any reasonable predictions. For example, the first row of Table 8, which shows the values of $\sigma(a_{11})$ as the scale factor α of Σ_A increases, indicates that $\sigma(a_{11})$ increases as α goes from 1.1 upto 15 and then drops at $\sigma = 30$. Similarly, the third row shows that $\sigma(a_{12})$ increases till $\alpha = 6$ and then drops for $\alpha = 15$ and $\alpha = 30$. The second row keeps increasing whereas the fourth row behaves like the first. However, there is no identifiable pattern which allows us to predict the behaviour of these sensitivities. Also, since the values of Σ_A are very small, even a scale factor of 30 does not succeed in making K_t blow up. We are still within the threshold even though we do not know exactly what it is.

To sum up, we could say that the outcome of the analysis on this model is basically positive. There are no really sensitive parameters, so we can trust the results of the model (on the assumption, of course, that the underlying economics is accurate).

4.5 Conclusion :

In this chapter, we have presented a simple macroeconomic model of the U.S. economy and recast it into state-variable form. Next, we have applied the equations developed in Chapters 2 and 3 to this model, and

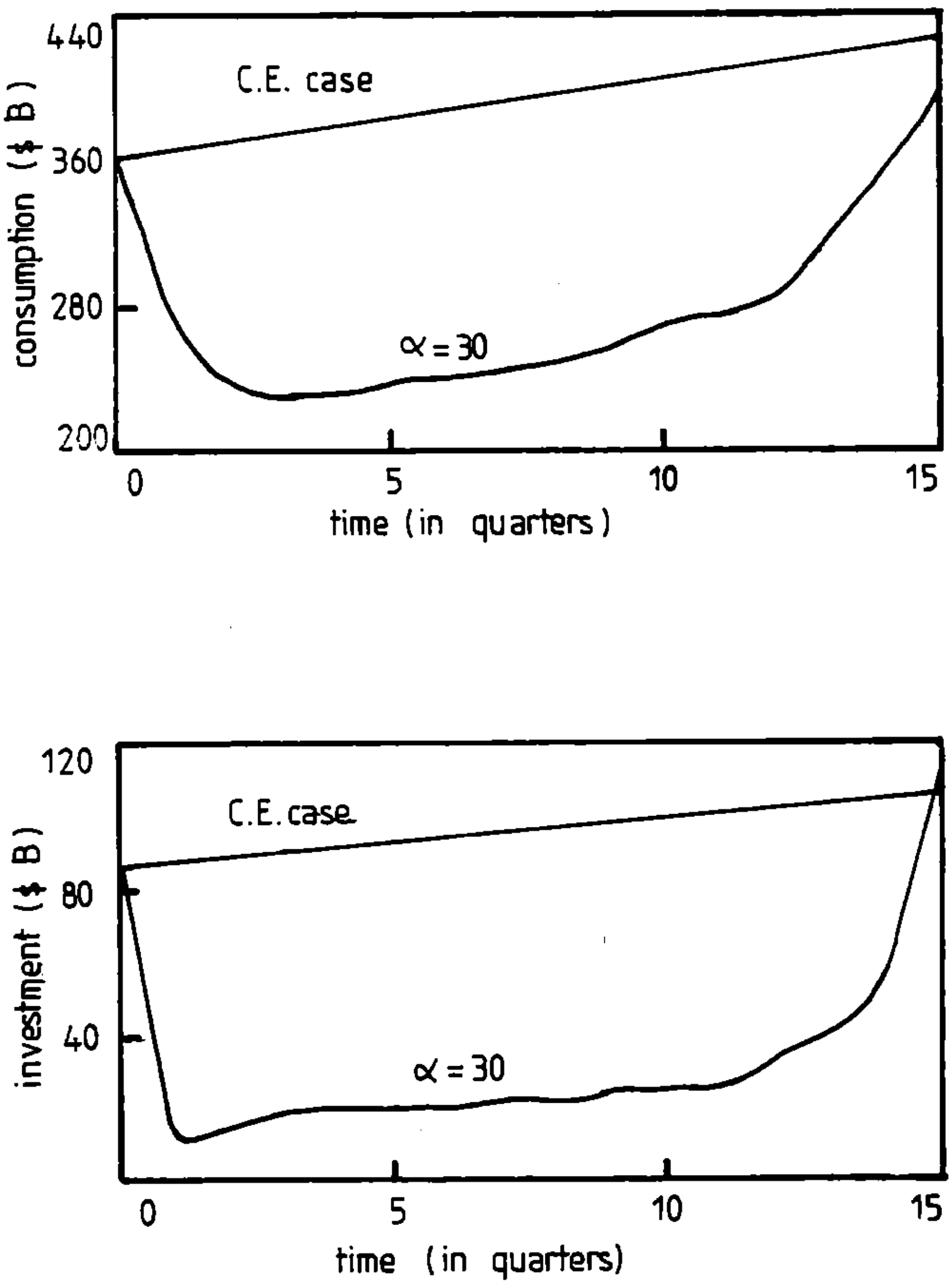
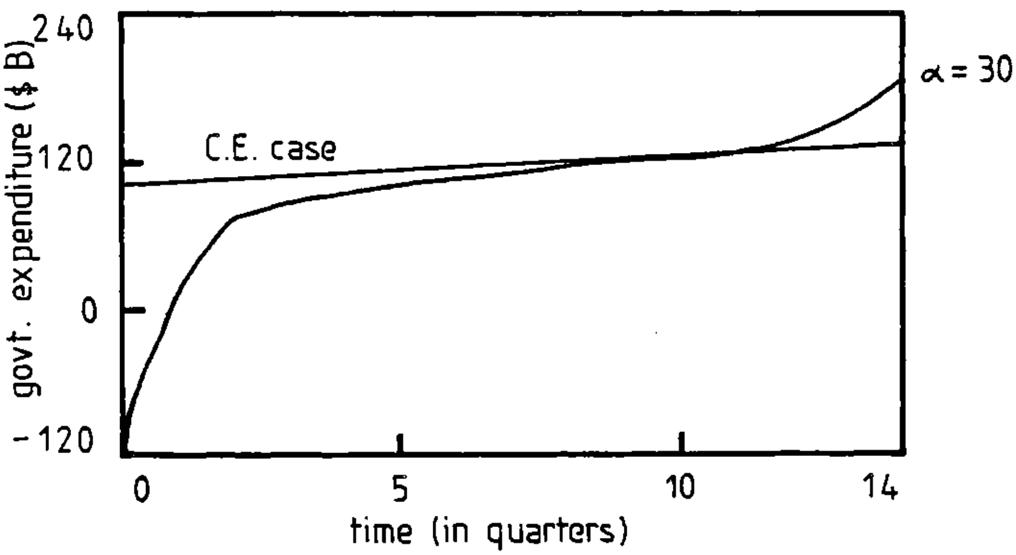


Fig. 9. State trajectory, Eq. (3.3.11). Comparison of trajectories for C.E. case with the stochastic case when Σ_A is scaled by a factor of 30. Q = (2,1) for all curves.



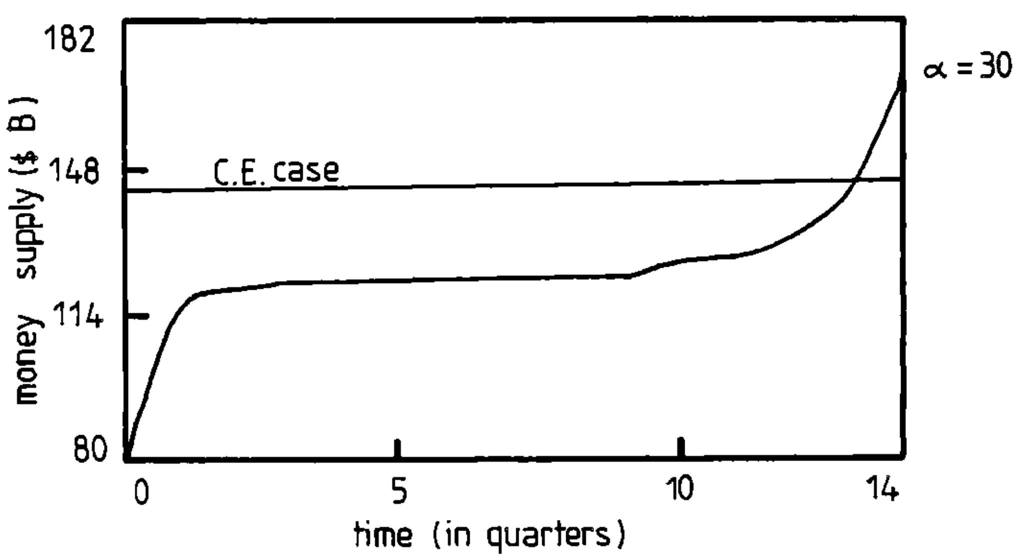


Fig.10. Control trajectory, Eq. (A.4), for N = 15. Comparison of trajectories for C.E. case with the stochastic case when Σ_A is scaled by a factor of 30. Q = (2,1) for all curves.

	α:1.1	α:2	α:6	α:15	α:30
$\sigma(a_{11})$	.115	.170	.272	.294	.215
$\sigma(a_{21})$	.106	.161	.287	.397	.444
$\sigma(a_{12})$	.044	.063	.089	.077	.049
$\sigma(a_{22})$	.040	.060	.094	.104	.101
$\sigma(b_{11})$	.127	.122	.106	.080	.048
$\sigma(b_{21})$	.117	.115	.112	.108	.098
$\sigma(b_{12})$	.164	.131	.065	.026	.012
$\sigma(b_{22})$	.151	.124	.069	.035	.026

Table 8. Relative sensitivities for Q = (2,1) and different scale factors α for Σ_A

(i.e. the actual covariance used in simulations is $\alpha\Sigma_A$ where Σ_A is given on page 63).

	α:1.1	α:2	α:6	α:15	α:30
$\sigma(a_{11})$	2860368	2860618	2855406	2801006	2547648
$\sigma(a_{21})$	1237788	1268444	1410905	1768482	2455696
$\sigma(a_{12})$	370616	359680	317017	249433	197812
$\sigma(a_{22})$	160508	159594	156699	157288	190995
$\sigma(b_{11})$	514235	605546	985255	1697967	2542816
$\sigma(b_{21})$	222153	266288	486861	1068166	2405713
$\sigma(b_{12})$	369435	363047	333829	308774	362907
$\sigma(b_{22})$	159614	160796	167413	194690	349698

Table 9. Absolute sensitivities for Q = (2,1) and different scale factors α for Σ_A
(i.e. the actual covariance used in simulations is $\alpha\Sigma_A$ where Σ_A is given on page 63).

presented some empirical results together with a discussion of these results.

Our model turns out to be fairly insensitive to parameter uncertainty variations and therefore quite reliable. Applications of this method to more models is required for a better understanding of the equations we have developed. It seems, however, that the complexity of these equations and their relative resistance to deeper insight makes this method of approaching sensitivity issues undesirable. The computation involved increases at a prohibitively untrammelled rate as the dimension of the model increases and since most useful econometric models are large, this method is not quite practical. It can, however, be useful when a small subset of the parameters in a large model needs to be analysed for its sensitivity. This, of course, is to be expected since this method is essentially a brute force way of identifying sensitive parameters.

CHAPTER 5

CONCLUSION

5.1 Summary of Results

In this report, we have investigated the structure of optimal, linear, random parameter systems. We model these parameters as white stochastic processes. Thus, the model contains both additive and multiplicative white noise. This white parameter approach to adaptive stochastic control is important for two reasons. Firstly, it makes the problem solvable analytically. The general adaptive control problem is in fact a nonlinear stochastic control problem and cannot be solved without making approximations. Secondly, it shows, in a worst case sense, the fact that the control gains of an optimal stochastic system with purely random parameters depend not only upon the mean values, but also upon the variances of the random parameters. The scalar case of this problem was investigated by Ku [1]. Here we investigate the most general multivariable version. The problem is formulated as a tracking problem and includes additive noise as well. We do this work in Chapter 2.

In the next chapter, we develop sensitivity equations to analyse the sensitivity of the system performance to small variations in the variances of the system parameters. The equations turn out to be fairly cumbersome in the general multivariable case. Deriving equations for the sensitivity of the optimal control and the optimal trajectory turns out to be hopelessly complicated.

We describe a simple macroeconomic model, recast it into an optimal control framework, and make a thorough investigation of its structure and of the optimal solution together with the sensitivities of the different parameters. We present some of the relevant simulation results for the analysis.

5.2 Conclusions :

The multivariable case for linear random parameter systems, though solvable analytically, turns out to be somewhat opaque and does not yield much further insight than the scalar case. The main result for the scalar case described in Ku [1] is the Uncertainty Threshold Principle. In the scalar case it is possible to find an analytic expression for this threshold (some function of all the means and covariances). In the multivariable case, we find that it is very difficult, if not impossible, to obtain an analytical expression for the threshold. The source of the problem is that we are dealing with matrix quantities and matrix multiplication is non-commutative and operations like the trace of a product of matrices do not decouple. However, a threshold certainly exists as can be verified by trying out different values for the various mean and covariance matrices.

The sensitivity equations, since they are derived from the above optimal solution, turn out to be even less amenable to any insight. We do not even bother to reproduce the equations for the sensitivities of the optimal control and state trajectory. The application of these equations to Abel's model also turns out to be of dubious value. Though they do supply us with some valuable information - that the model is basically insensitive and therefore reasonably reliable - it is questionable whether such a brute force approach to sensitivity analysis is worthwhile. Many currently popular econometric models are large and nonlinear and this approach would become far too involved computationally. The cpu time depends geometrically ($\sim n^2$) on the order of the system and linearly on the time horizon. However, if we restrict the set of parameters whose sensitivities we wish to examine to a small subset of all the parameters, then we can hope to extract some useful information at a reasonable cost.

5.3 Suggestions for Future Research :

1. More analysis is required to thoroughly understand the different aspects of tracking problems. Specifically, one needs to understand the end-point behaviour of various variables like x_t, u_t, L_t and m_t physically. It may help to reduce these matrix and vector quantities to scalars by using suitable norms.

2. We have calculated quantities like $\frac{\partial K_t}{\partial \sigma}$. It may be useful to consider quantities like $\frac{\partial K_{t+\theta}}{\partial \sigma}$ as well. This represents the effect of a change in the present value of σ on the future value of K_t. This may prove to be useful in adaptive control schemes where such information may be used to guide control action.

3. Though the equations turn out to be very complicated, it would be useful to look at the behaviour of $\frac{\partial \Sigma u_t}{\partial \sigma}$, $\frac{\partial \Sigma x_t}{\partial \sigma}$. Perhaps somewhat different initial assumptions might lead to a more tractable problem which might yield useful information.

4. The scheme developed in this report can be applied to assess the reliability of different models of a given system. This affords a selection criterion which can aid in choosing one out of a number of models.

5. This sensitivity analysis can also be applied to an analysis of the monetarist-fiscalist debate in Abel's paper [47].

BIBLIOGRAPHY

1. R. Ku, "Adaptive Stochastic Control of Linear Systems with Random Parameter," ESL-TH-820, MIT, May 1978

2. R. Bellman, "Adaptive Control Processes", Princeton Univ. Press, 1961

3. A. Feldbaum, "Optimal Control Systems", N.Y., Academic Press, 1965

4. J. Sternby, "A Simple Dual Control Problem with an Analytical Solution," IEEE Trans. on Automatic Control. Vol. AC-21, Dec. 1976

5. Y. Bar-Shalom and E. Tse, "Dual Effect, Certainty Equivalence and Separation in Stochastic Control," IEEE Trans. on Automatic Control, Vol. AC-19, No. 5, Oct. 1974

6. E. Tse and Y. Bar-Shalom, "Generalized Certainty Equivalence and Dual Effect in Stochastic Control," IEEE Trans. on Automatic Control, Vol. AC-20, No. 6, Dec. 1975

7. Y. Bar-Shalom and E. Tse, "Caution, Probing and the Value of Information in the Control of Uncertain Systems," Annals of Economic and Social Measurement, 5/3, 1976

8. E. Tse, Y. Bar-Shalom and L. Meier, "Wide-Sense Adaptive Dual Control for Nonlinear Stochastic Systems," IEEE Trans. on Automatic Control, Vol. AC-18, No. 2, April 1973

9. E. Tse and Y. Bar-Shalom, "An Actively Adaptive Control for Linear Systems with Random Parameters via the Dual Control Approach," IEEE Trans. on Automatic Control, Vol. AC-18, No. 2, April 1973

10. S. Dreyfus, "Some Types of Optimal Control of Stochastic Systems," J. SIAM Control, Vol. 2, No. 1, 1962

11. R. Fair, "An Evaluation of a Short-Run Forecasting Model," Int. Econ. Rev., Vol. 15, No. 2, June 1974.

12. Davind Kendrick, "Applications of Control Theory to Economics," Annals of Economic and Social Measurement, Vol. 5/2, 1976 pp.171-190

13. M. Athans and D. Kendrick, "Control Theory and Economics : A Survey, Forecast and Speculations," IEEE Trans. on Automatic Control, pp. 518-524, Vol. AC-19, Oct. 1974

14. M. Aoki, "Stochastic Control in Economic Theory and Economic Systems," IEEE Trans. on Automatic Control, Vol. AC-21, pp. 213-220, April 1976

15. A. Tustin, "The Mechanism of Economic Systems," 1953. Harvard Univ. Press

16. A.W. Phillips, "Stabilization Policy in a Closed Economy," Economic Journal, Vol. 6A, June 1954

17. H. Theil, "A Note on Certainty Equivalence in Dynamic Planning," Econometrica, Vol. 23, pp. 346-349, April 1957

18. H. Simon, "Dynamic Programming under Uncertainty with a Quadratic Criterion Function," Econometrica, Vol. 24, Jan. 1956

19. M. Aoki, "Control of Linear Discrete Time Systems with Multiplicative Stochastic Disturbances in Gain," IEEE Trans. on Automatic Control, Vol. AC-20, No. 3, pp. 388-391, 1975

20. M. Aoki, "Optimal Control and System Theory in Dynamic Economic Analysis," North Holland, 1976

21. G. Chow, "Multiplier, Accelerator and Liquidity Preference in the Determination of National Income in the U.S.," Review of Economics and Statistics, Vol. 49, pp. 1-15, Feb. 1967

22. G. Chow, "Optimal Stochastic Control of Linear Economic Systems," Journal of Money Credit and Banking, Vol. 2, pp. 411-425, 1970

23. G. Chow, "Optimal Control of Linear Econometric Systems with Finite Time Horizons," International Economic Review, Vol. 13, No. 1, pp. 16-25, Feb. 1972

24. G. Chow, "How Much Could Be Gained By Optimal Stochastic Control Policies," Annals of Economic and Social Measurement, Vol. 1, No. 4, pp. 391-406, 1972

25. G. Chow, "Effect of Uncertainty on Optimal Control Policies," International Economic Review, Vol. 14, pp. 632-645

26. G. Chow, "Problems of Economic Policy from the Viewpoint of Optimal Control," American Economic Review, Vol. LXIII, No. 5, pp. 825-837, Dec. 1973

27. G. Chow, "Analysis and Control of Dynamic Economic Systems," John Wiley & Sons, N.Y., 1975

28. G. Chow, "The Control of Nonlinear Econometric Systems with Unknown Parameters," Econometrica, Vol. 44, No. 4, pp. 685-695, 1976

29. G. Chow and S. Megdal, "The Control of Large-Scale Nonlinear Econometric Systems," IEEE Trans. on Automatic Control, Vol. AC-23, No. 2, April 1978

30. D. Kendrick and J. Majors, "Stochastic Control with Uncertain Macroeconomic Parameters," Automatica, Vol. 10, No. 2, pp. 587-594

31. D. Kendrick, "Adaptive Control of Macroeconomic Models with Measurement Error," Discussion Paper, Centre for Economic Research, Dept. of Economics, Univ. of Texas, Austin, Texas, 1977

32. R. Pindyck, "An Application of the Linear Quadratic Tracking Problem to Economic Stabilization Policy," IEEE Trans. on Automatic Control, Vol. AC-17, No. 3, June 1972

33. R. Pindyck, "Optimal Stabilization Policies via Deterministic Control," Annals of Economic and Social Measurement, Vol. I, No. 4, pp. 385-390, Oct. 1972

34. R. Pindyck, "Optimal Planning for Economic Stabilization," North Holland Publishing Co., Amsterdam, 1973

35. M. Athans, E. Kuh, L. Papademos, R. Pindyck, R. Ku, T. Ozkan, K. Wall, "Sequential Open Loop Optimal Control of a Nonlinear Macroeconomic Model," MIT, 1975

36. A. Abel, "A Comparison of Three Control Algorithms as Applied to the Monetarist-Fiscalist Debate," Annals of Economic and Social Measurement, Vol. 4, No. 2, Spring 1975, pp. 239-252

37. M. Athans, "The Importance of Kalman Filtering Methods for Economic Systems," Annals of Economic and Social Measurement, Vol. 3, No. 1, pp. 49-64, 1974

38. D. Castanon, M. Athans, E. Hnyilicza, L.F. Pau, S. Gershwin, "Applications of Advanced Control Theory in Microeconomics," ESL-R-662, MIT, June 1977

39. L. von Bertalanffy, "General Systems Theory,"

40. J. Forrester, "System Dynamics,"

41. R. Kalman, "Control of randomly varying linear dynamical systems," Proc. of Symposium in Applied Math. Vol. 13, Hydrodynamic Instability, pp. 287-298, 1962.

42. P. Joseph and J. Tou, "On Linear Control Theory," AIEE Trans. (Appl. and Ind.) Pt. II, Vol. 80, pp. 193-196, Sept. 1961

43. T.L. Gunckel and G. Franklin, "A General Solution for Linear Sampled - Data Control," Trans. of the ASME, Journal of Basic Engineering, pp. 197-203, June 1963

44. M. Athans, R. Ku and S. Gershwin, "The Uncertainty Threshold Principle : Some Fundamental Limitations of Optimal Decision Making under Dynamic Uncertainty," IEEE Trans. on Automatic Control, Vol. AC-22, pp. 491-495, June 1977

45. R. Ku and M. Athans, "Further Results on the Uncertainty Threshold Principle," IEEE Trans. on Automatic Control, Vol. AC-22, pp. 866-868, Oct. 1977

46. R. Ku, M. Athans and P. Varaiya, "The Effects of Discounted Cost on the Uncertainty Threshold Principle," ESL-P-749, MIT, May 1977

47. A. Abel, "A Comparison of Three Control Algorithms as Applied to the Monetarist-Fiscalist Debate," Annals of Economic and Social Measurement, Vol. 4/2, 1975

48. J.W. Forrester, "World Dynamics," Wright-Allen Press, Cambridge, Ma. 1971.

49. A. Wierzbicki and A. Dontchev, "Basic Relations in Performance Sensitivity Analysis of Optimal Control Systems," Control and Cybernetics, Vol. 3, No. 3/4, 1974

50. P. Vermenlen and De Jongh, "Growth in a Finite World - A Comprehensive Sensitivity Analysis," Automatica, Vol. 13, pp. 77-84, 1977

51. D. Meadows, J. Randers and W. Behreus, "The Limits to Growth," Potomac Associates - Universe Books, N.Y. 1972

52. R. Tomovic and R. Vukobratovic, "General Sensitivity Theory," Elsevier, N.Y. 1972

53. "Feedback Systems" edited by J. Cruz, Mc-Graw Hill, 1972

54. W. Gersch and F. Kozin, "Optimal Control of Multiplicatively Perturbed Stochastic Systems," Proc. 1963, Allerton Conference on Circuit and System Theory, Univ. of Illinois

55. Witsenhausen, "On the Sensitivity of Optimal Control Systems," IEEE, Trans. on Automatic Control, Vol. AC-10, 1965.

56. Bokrosky and Graupe, "Analysis of Optimal Cost Sensitivity to Parameter Changes," IEEE Trans. on Automatic Control, Vol. AC-16 Oct. 1971

57. Kurtaran and Sidar, "Analysis of Cost Sensitivity for LQ Stochastic Problems with instantaneous output feedback," IEEE Trans. on Automatic Control, Vol. AC-19, Oct. 1974

58. P. Dorato, "On Sensitivity in Optimal Control Systems," IEEE Trans. on Automatic Control, Vol. AC-8, July 1963.

59. B. Pagurek, "Sensitivity of the Performance of Linear Optimal Control, Systems to Parameter Variations," IEEE Trans. on Automatic Control, Vol. AC-10, April 1965

60. Sarma and Deekshatulu, "Sensitivity Design of Optimal Linear Systems," Int. J. of Control, Vol. 8, No. 6, 1968

61. Higginbotham, "Design of sensitivity and State variable feedback of Optimal Control Systems," Proc. 7th Allterton Conference on Circuit and System Theory, Oct. 1969.

62. J. Cruz and W. Perkins, "A New Approach to the Sensitivity Problem in Multivariable Feedback System Design," IEEE Trans. on Automatic Control, Vol. AC-9, July 1964

63. E. Kreindler, "Closed-loop Sensitivity Reduction of Linear Optimal Control Systems," IEEE Trans. on Automatic Control, Vol. AC-13, June 1968.

APPENDIX A

We solve here the optimal control problem posed in Chapter 2 using the method of stochastic dynamic programming.

We begin by stating the problem and the principle of optimality.

We have the following linear random parameter system :

$$x_{k+1} = A_k x_k + B_k u_k + c_k \qquad x_0 \text{ given} \tag{A.1}$$

where A_k, B_k and c_k are all white and Gaussian with known means, covariances and cross-covariances.

$$E\{A_k\} = \bar{A}$$

$$E\{B_k\} = \bar{B}$$

$$E\{c_k\} = \bar{c}$$

$$E\{[S(A_k) - S(\bar{A})][S(A_\ell) - S(\bar{A})]\} = \Sigma_A \delta_{k\ell}$$

$$E\{[S(B_k) - S(\bar{B})][S(B_\ell) - S(\bar{B})]\} = \Sigma_B \delta_{k\ell}$$

$$E\{[S(c_k) - S(\bar{c})][S(c_\ell) - S(\bar{c})]\} = \Sigma_c \delta_{k\ell}$$

$$E\{[S(B_k) - S(\bar{B})][S(A_\ell) - S(\bar{A})]\} = \Sigma_{BA} \delta_{k\ell}$$

$$E\{[S(B_k) - S(\bar{B})][S(c_\ell) - S(\bar{c})]\} = \Sigma_{Bc} \delta_{k\ell}$$

$$E\{[S(A_k) - S(\bar{A})][S(c_\ell) - S(\bar{c})]\} = \Sigma_{Ac} \delta_{k\ell}$$

Here we introduce some notation for convenience. For any matrices Y_k, Z_k let

$$E\{Y_k' Z_k' Y_k\} = \overline{Y_k' Z_k Y_k}$$

$$= \overline{Y' Z_k Y} \qquad \text{if } E\{Y_k\} = \bar{Y} \text{ constant}$$

The cost functional we choose to minimize is :

$$J = \frac{1}{2} E\{\sum_{k=0}^{N-1} [(x_k - \tilde{x}_k)'Q(x_k - \tilde{x}_k) + (u_k - \tilde{u}_k)'R(u_k - \tilde{u}_k)] + (x_N - \tilde{x}_N)'Q(x_N - \tilde{x}_N)\} \tag{A.2}$$

where Q, R are symmetric, positive semi-definite matrices and where x_k, u_k are given target trajectories.

The stochastic control problem is to find a control sequence $\{u_o, u_1, \ldots, u_{N-1}\}$ that minimizes the value of J. This problem is the stochastic tracking type of optimization problem and can be solved with either the discrete minimum principle or dynamic programming. We choose the second approach.

Let

$$J_k = \frac{1}{2} \sum_{i=k}^{N} [(x_i - \tilde{x}_i)'Q(x_i - \tilde{x}_i) + (u_{i-1} - \tilde{u}_{i-1})'R(u_{i-1} - \tilde{u}_{i-1})]$$

$$P_k = \frac{1}{2} [(x_k - \tilde{x}_k)'Q(x_k - \tilde{x}_k) + (u_{k-1}-\tilde{u}_{k-1})'R(u_{k-1}-\tilde{u}_{k-1})]$$

$$\gamma_k = E\{J_k\}$$

$$\lambda_k = E\{P_k\}$$

$$\gamma_k^* = \min_{u_{k-1}, \ldots, u_{N-1}} \gamma_k$$

where $k = 1, 2, \ldots, N$

We have

$$\gamma_k^* = \min_{u_{k-1}, \ldots, u_{N-1}} \gamma_k \qquad k = 1,2, \ldots, N$$

$$\gamma_k^* = \min_{u_{k-1},\ldots,u_{N-1}} EJ_k$$

$$= \min_{u_{k-1},\ \ldots\ ,\ u_{N-1}} E\,(P_k + P_{k+1} + \ldots + P_N)$$

$$= \min_{u_{k-1},\ldots,u_{N-1}} EP_k + \min_{u_{k-1},\ldots u_{N-1}} EJ_{k+1}$$

(Note : $EJ_{N+1} \equiv 0$)

$$= \min_{u_{k-1}} EP_k + \min_{u_{k-1}} \left(\min_{u_k,\ \ldots\ ,\ u_{N-1}} EJ_{k+1}\right)$$

$$= \min \lambda_k + \min \gamma_{k+1}^* \qquad \text{(Note : } \gamma_{N+1}^* \equiv 0)$$

$$\gamma_k^* = \min_{u_{k-1}} (\lambda_k + \gamma_{k+1}^*) \tag{A.3}$$

This is the functional recurrence relation that we shall use in our derivation.

We shall first calculate λ_k.

$$\lambda_k = EP_k$$

$$= \int P_k\, p(x_k)\, dx_k$$

$$= \frac{1}{2} \int [(x_k-\tilde{x}_k)'Q(x_k-\tilde{x}_k) + (u_{k-1}-\tilde{u}_{k-1})'R(u_{k-1}-\tilde{u}_{k-1})]\ .$$

$$p(x_k/A_{k-1},B_{k-1},c_{k-1},x_{k-1})p(A_{k-1},B_{k-1},c_{k-1})p(x_{k-1})$$

$$d(x_k,A_{k-1},B_{k-1},c_{k-1},x_{k-1})$$

using $p(x) = \int p(x/y)\, p(y)\, dy$.

Note that x_{k-1} is independent of A_{k-1}, B_{k-1}, c_{k-1} so we can write

$$p(A_{k-1},B_{k-1},c_{k-1},x_{k-1}) = p(A_{k-1},B_{k-1},c_{k-1})\ p\ (x_{k-1})$$

Also, $d(x_k,\ A_{k-1},\ B_{k-1},\ c_{k-1},\ x_{k-1})$ is merely an abbreviation for $dx_k dA_{k-1} dB_{k-1} dc_{k-1} dx_{k-1}$.

Therefore,

$$\begin{aligned}\lambda_k \quad = \quad & \frac{1}{2} \int [\ x'_{k-1}(A_{k-1}QA_{k-1})x_{k-1} + u'_{k-1}(R+B'_{k-1}QB_{k-1})\ u_{k-1} \\ & + c'_{k-1}Qc_{k-1} + 2u'_{k-1}(B'_{k-1}QA_{k-1})x_{k-1} + 2u'_{k-1}(B'_{k-1}Qc_{k-1}) \\ & + 2x'_{k-1}(A_{k-1}Qc_{k-1}) + \tilde{x}'_kQ\tilde{x}_k + \tilde{u}_{k-1}R\tilde{u}_{k-1} - 2\tilde{u}'_{k-1}Ru_{k-1} \\ & - 2\tilde{x}'_k(QA_{k-1})x_{k-1} - 2\tilde{x}'_k(QB_{k-1})u_{k-1} - 2\ \tilde{x}'_k(Qc_{k-1})\ ]\ . \\ & \qquad p(A_{k-1},B_{k-1},c_{k-1})p(x_{k-1})d(A_{k-1},B_{k-1},c_{k-1},x_{k-1})\end{aligned}$$

using $x_k = A_{k-1}x_{k-1} + B_{k-1}u_{k-1} + c_{k-1}$ and integrating out x_k.

Now, integrating with respect to A_{k-1}, B_{k-1}, and c_{k-1} we get

$$\begin{aligned}\lambda_k \quad = \quad & \frac{1}{2} \int [\ x'_{k-1}\overline{A'QA}\ x_{k-1} + u'_{k-1}(R + \overline{B'QB})\ u_{k-1} + \overline{c'Qc} \\ & + 2u'_{k-1}(\overline{B'QA})x_{k-1} + 2u'_{k-1}(\overline{B'Qc}) + 2x'_{k-1}(\overline{A'Qc}) \\ & + \tilde{x}'_kQ\tilde{x}_k + \tilde{u}'_{k-1}R\tilde{u}_{k-1} - 2\ \tilde{u}'_{k-1}Ru_{k-1} \\ & - 2\tilde{x}'_kQAx_{k-1} - 2\tilde{x}'_kQBu_{k-1} - 2\tilde{x}'_kQc\ ]\ p(x_{k-1})dx_{k-1}\end{aligned}$$

1. $k = N$

$$\gamma_N^* = \min_{u_{N-1}} \lambda_N \qquad (\gamma_{N+1}^* \equiv 0)$$

$$\frac{d\lambda_N}{du_{N-1}} = 0$$

$$2(R + \overline{B'QB})u_{N-1} + 2(\overline{B'QA})x_{N-1} + 2(\overline{B'Qc}) - 2R\tilde{u}_{N-1} - 2\bar{B}'Q\tilde{x}_N = 0$$

$$u_{N-1}^* = -(R + \overline{B'QB})^{-1}(\overline{B'QA})x_{N-1}$$
$$- (R + \overline{B'QB})^{-1}(\overline{B'Qc} - \bar{B}'Q\tilde{x}_N - R\tilde{u}_{N-1})$$

With this, we calculate γ_N^*.

$$\gamma_N^* = \frac{1}{2} \int [\, x_{N-1}'(\overline{A'QA})x_{N-1} + x_{N-1}'(\overline{A'QB})(R+\overline{B'QB})^{-1}(\overline{B'QA})x_{N-1}$$
$$+ (\overline{B'Qc} - \bar{B}'Q\tilde{x}_N - R\tilde{u}_{N-1})'(R+\overline{B'QB})^{-1}(\overline{B'Qc} - \bar{B}'Q\tilde{x}_N - R\tilde{u}_{N-1})$$
$$+ 2x_{N-1}(\overline{B'QA})'(R + \overline{B'QB})^{-1}(\overline{B'Qc} - \bar{B}'Q\tilde{x}_N - R\tilde{u}_{N-1})$$
$$+ \overline{c'Qc} - 2x_{N-1}'(\overline{A'QB})\ (R+\overline{B'QB})^{-1}(\overline{B'QA})x_{N-1}$$
$$- 2(\overline{B'Qc} - \bar{B}'Q\tilde{x}_N - R\tilde{u}_{N-1})'(R + \overline{B'QB})^{-1}(\overline{B'QA})x_{N-1}$$
$$- 2\ x_{N-1}'(\overline{A'QB})(R + \overline{B'QB})^{-1}(\overline{B'Qc})$$
$$- 2(\overline{B'Qc} - \bar{B}'Q\tilde{x}_N - R\tilde{u}_{N-1})'(R + \overline{B'QB})^{-1}(\overline{B'Qc})$$
$$+ 2\ x_{N-1}'(\overline{A'Qc}) + \tilde{x}_N'\ Q\tilde{x}_N + \tilde{u}_{N-1}'R\tilde{u}_{N-1}$$
$$+ 2\ \tilde{u}_{N-1}'\ R(R + \overline{B'QB})^{-1}(\overline{B'QA})\ x_{N-1}$$
$$+ 2\ \tilde{u}_{N-1}'R\ (R + \overline{B'QB})^{-1}(\overline{B'Qc} - \bar{B}'Q\tilde{x}_N - R\tilde{u}_{N-1})$$
$$- 2\ \tilde{x}_N'Q\bar{A}x_{N-1}$$

$$+ 2\,\tilde{x}_N'Q\bar{B}(R + \overline{B'QB})^{-1}(\overline{B'Qc} - \bar{B}'Q\tilde{x}_N - R\tilde{u}_{N-1})$$

$$- 2\,\tilde{x}_N'Q\bar{c}\ ]\quad p(x_{N-1})\ dx_{N-1}$$

On simplifying the above, we get,

$$\gamma_N^* = \frac{1}{2}\int [\ x_{N-1}'\{\overline{A'QA} - (\overline{A'QB})(R + \overline{B'QB})^{-1}(\overline{B'QA})\}\quad x_{N-1}$$

$$+ 2x_{N-1}'\ \{\overline{A'Qc} - \bar{A}'Qx_N - (\overline{A'QB})(R+\overline{B'QB})^{-1}(\overline{B'Qc} - \bar{B}'Q\tilde{x}_N - R\tilde{u}_{N-1})\}$$

$$- (\overline{B'Qc} - \bar{B}'Q\tilde{x}_N - R\tilde{u}_{N-1})'(R+\overline{B'QB})^{-1}(\overline{B'Qc} - \bar{B}'Q\tilde{x}_N - R\tilde{u}_{N-1})$$

$$+ \overline{c'Qc} + \tilde{x}_N'Q\tilde{x}_N + \tilde{u}_{N-1}'R\tilde{u}_{N-1} - 2\,\tilde{x}_N'Q\bar{c}\ ]\ p(x_{N-1})dx_{N-1}$$

Since we know the final answer, we can make some convenient definitions at this point.

Let

$$K_N = Q$$

$$p_N = -\,Q\tilde{x}_N$$

$$g_N = \frac{1}{2}\,\tilde{x}_N'Q\tilde{x}_N$$

Then,

$$\gamma_N^* = \frac{1}{2}\int [\ x_{N-1}'(\overline{A'K_NA} - \overline{A'K_NB}(R + \overline{B'K_NB})^{-1}\ \overline{B'K_NA})\ x_{N-1}$$

$$+ 2x_{N-1}'(\overline{A'K_Nc} + \bar{A}'p_N - (\overline{A'K_NB})(R+\overline{B'K_NB})^{-1}(\overline{B'K_Nc} + \bar{B}'p_N - R\tilde{u}_{N-1})$$

$$- (\overline{B'K_Nc} + \bar{B}'p_N - R\tilde{u}_{N-1})'(R+\overline{B'K_NB})^{-1}(\overline{B'K_Nc} + \bar{B}'p_N - R\tilde{u}_{N-1})$$

$$+ \overline{c'K_Nc} + 2g_N + \tilde{u}_{N-1}'R\tilde{u}_{N-1} + 2\bar{c}'p_N\ ]\ p(x_{N-1})dx_{N-1}$$

Now define

$$D_{N-i-1} = \overline{A'K_{N-i}A} - (\overline{A'K_{N-i}B})(R+\overline{B'K_{N-i}B})^{-1}(\overline{B'K_{N-i}A})$$

$$q_{N-i-1} = \overline{A'K_{N-i}c} + \bar{A}'p_{N-i} - \overline{(A'K_{N-i}B)}(R+\overline{B'K_{N-i}B})^{-1}$$
$$(\overline{B'K_{N-i}C} + \bar{B}'p_{N-i} - R\tilde{u}_{N-i-1})$$

$$K_i = Q + D_i$$

$$p_i = -Q\tilde{x}_i + q_i$$

$$L_i = -(R + \overline{B'K_{i+1}B})^{-1}\ (\overline{B'K_{i+1}A})$$

$$m_i = -(R + \overline{B'K_{i+1}B})^{-1}(\overline{B'K_{i+1}c} + \bar{B}'p_{i+1} - R\tilde{u}_i)$$

$$r_{N-i-1} = \frac{1}{2}(\overline{B'K_{N-i}c} + \bar{B}'p_{N-i} - R\tilde{u}_{N-i-1})'\ m_{N-i-1} + \frac{1}{2}\overline{c'K_{N-i}c}$$
$$+ \frac{1}{2}\tilde{u}'_{N-i-1}R\tilde{u}_{N-i-1} + \bar{c}'p_{N-i} + g_{N-i}$$

$$g_i = \frac{1}{2}\tilde{x}_i'Q\tilde{x}_i + r_i$$

Thus we can write

$$\gamma_N^* = \frac{1}{2}\int[x'_{N-1}D_{N-1}x_{N-1} + 2x'_{N-1}q_{N-1} + 2r_{N-1}]\ p(x_{N-1})dx_{N-1}$$

$$u_{N-1}^* = L_{N-1}\ x_{N-1} + m_{N-1}$$

From here we go on to the next step in our calculation.

2. $k = N - 1$

$$\gamma_{N-1}^* = \min_{u_{N-2}}\ (\lambda_{N-1} + \gamma_N^*)$$

We have, from the previous step,

$$\gamma_N^* = \frac{1}{2}\int[\ x'_{N-1}D_{N-1}x_{N-1} + 2x'_{N-1}q_{N-1} + 2r_{N-1}]\ p(x_{N-1})dx_{N-1}$$

$$= \frac{1}{2}\int[\ x'_{N-2}(A'_{N-2}D_{N-1}A_{N-2})x_{N-2} + u'_{N-2}B'_{N-2}D_{N-1}B_{N-2}u_{N-2}$$

$$+ c'_{N-2}D_{N-1}c_{N-2} + 2u'_{N-2}B'_{N-2}D_{N-1}A_{N-2}x_{N-2}$$

$$+ 2u'_{N-2}B'_{N-2}D_{N-1}c_{N-2} + 2x'_{N-2}A'_{N-2}D_{N-1}c_{N-2}$$

$$+ 2x'_{N-2}A'_{N-2}q_{N-1} + 2\ u'_{N-2}B'_{N-2}q_{N-1} + 2c'_{N-2}q_{N-1}$$

$$+ 2r_{N-1}]\ p(x_{N-1}|\ A_{N-2},B_{N-2},c_{N-2},x_{N-2})p(A_{N-2},B_{N-2},c_{N-2})$$

$$p(x_{N-2})d(x_{N-1},A_{N-2},B_{N-2},c_{N-2},x_{N-2})$$

$$= \frac{1}{2}\int[\ x'_{N-2}\overline{A'D_{N-1}A}\ x_{N-2} + u'_{N-2}\overline{B'D_{N-1}B}\ u_{N-2} + \overline{c'D_{N-1}c}$$

$$+ 2\ u'_{N-2}\overline{B'D_{N-1}A}\ x_{N-2} + 2\ u'_{N-2}\overline{B'D_{N-1}c} + 2\ x'_{N-2}\overline{A'D_{N-1}c}$$

$$+ 2\ x'_{N-2}\bar{A}'q_{N-1} + 2\ u'_{N-2}\bar{B}'q_{N-1} + 2\ \bar{c}'q_{N-1} + 2r_{N-1}]\ .$$

$$p(x_{N-2})dx_{N-2}$$

after integrating with respect to x_{N-1}, A_{N-2}, B_{N-2}, and c_{N-2}.

Therefore,

$$\lambda_{N-1} + \gamma_N^* = \frac{1}{2}\int\ [x'_{N-2}\ \{\ \overline{A'(Q + D_{N-1})A}\ \}\ x_{N-2}$$

$$+ u'_{N-2}\{R + \overline{B'(Q + D_{N-1})B}\}\ \ u_{N-2}$$

$$+ \overline{c'(Q + D_{N-1})c}$$

$$+ 2\ u'_{N-2}\ \overline{B'(Q + D_{N-1})A}\ x_{N-2}$$

$$+ 2\, u'_{N-2}\, \overline{B'(Q + D_{N-1})c}$$

$$+ 2\, x'_{N-2}\, \overline{A'(Q + D_{N-1})c}$$

$$+ 2\, x'_{N-2}\, \bar{A}'(q_{N-1} - Q\tilde{x}_{N-1})$$

$$+ 2\, u'_{N-2}\, \bar{B}'(q_{N-1} - Q\tilde{x}_{N-1})$$

$$+ 2\, \bar{c}'(q_{N-1} - Q\tilde{x}_{N-1})$$

$$+ 2\, r_{N-1}$$

$$+ \tilde{x}'_{N-1}Q\tilde{x}_{N-1} + \tilde{u}'_{N-2}R\tilde{u}_{N-2} - 2\, \tilde{u}'_{N-2}Ru_{N-2}] \,.$$

$$p(x_{N-2})dx_{N-2}$$

$$= \frac{1}{2} \int [\, x'_{N-2}(\overline{A'K_{N-1}A})x_{N-2} + u'_{N-2}(R+\overline{B'K_{N-1}B})u_{N-2}$$

$$+ \overline{c'K_{N-1}c} + 2\, u'_{N-2}(\overline{B'K_{N-1}A})x_{N-2} + 2\, u'_{N-2}\overline{B'K_{N-1}c}$$

$$+ 2x'_{N-2}\overline{A'K_{N-1}c} + 2\, x'_{N-2}\bar{A}'p_{N-1} + 2\, u'_{N-2}\bar{B}'p_{N-1}$$

$$+ 2\, \bar{c}'p_{N-1} + 2\, r_{N-1} + \tilde{x}'_{N-1}Q\tilde{x}_{N-1} + \tilde{u}'_{N-2}R\tilde{u}_{N-2}$$

$$- 2\, \tilde{u}'_{N-2}Ru_{N-2}\,]\; p(x_{N-2})dx_{N-2}$$

We can now minimise this expression w.r.t. u_{N-2}.

$$\frac{d}{du_{N-2}} \left(\lambda_{N-1} + \gamma_N^* \right) = 0$$

$$2(R + \overline{B'K_{N-1}B})u_{N-2} + 2\, \overline{B'K_{N-1}A}\; x_{N-2} + 2\, \overline{B'K_{N-1}c}$$

$$+ 2\, \bar{B}'p_{N-1} - 2\, R\tilde{u}_{N-2} = 0$$

$$u^*_{N-2} = -(R + \overline{B'K_{N-1}B})^{-1}(\overline{B'K_{N-1}A})\, x_{N-2}$$

$$- (R + \overline{B'K_{N-1}B})^{-1}(\overline{B'K_{N-1}c} + \bar{B}'p_{N-1} - R\tilde{u}_{N-2})$$

$$= L_{N-2}\, x_{N-2} + m_{N-2}$$

Let us now calculate γ^*_{N-1}

$$\gamma^*_{N-1} = \frac{1}{2}\int [\, x'_{N-2}(\overline{A'K_{N-1}A})x_{N-2} + x'_{N-2}L'_{N-2}(R + \overline{B'K_{N-1}B})L_{N-2}x_{N-2}$$

$$+ m'_{N-2}(R+\overline{B'K_{N-1}B})m_{N-2} + 2\, x'_{N-2}L'_{N-2}(R+\overline{B'K_{N-1}B})m_{N-2}$$

$$+ \overline{c'K_{N-1}c} + 2\, x'_{N-2}L'_{N-2}\overline{B'K_{N-1}A}\, x_{N-2} + 2\, m'_{N-2}\overline{B'K_{N-1}A}\, x_{N-2}$$

$$+ 2\, x'_{N-2}L'_{N-2}\overline{B'K_{N-1}c} + 2\, m'_{N-2}\overline{B'K_{N-1}c} + 2\, x'_{N-2}\overline{A'K_{N-1}c}$$

$$+ 2\, x'_{N-2}\bar{A}'p_{N-1} + 2\, x'_{N-2}L'_{N-2}\bar{B}'p_{N-1} + 2\, m'_{N-2}\bar{B}'p_{N-1}$$

$$+ 2\, \bar{c}'p_{N-1} + 2\, r_{N-1} + \tilde{x}'_{N-1}Q\tilde{x}_{N-1} + \tilde{u}'_{N-2}R\tilde{u}_{N-2}$$

$$- 2\, \tilde{u}'_{N-2}R\, L_{N-2}x_{N-2} - 2\, \tilde{u}'_{N-2}Rm_{N-2}]\; p(x_{N-2})dx_{N-2}$$

$$= \frac{1}{2}\int [x'_{N-2}D_{N-2}x_{N-2} + 2x'_{N-2}q_{N-2} + 2\, r_{N-2}]\; p(x_{N-2})dx_{N-2}$$

after some simplification and rearrangement of terms.

So we see that we get similar expressions for the control and optimal cost-to-go for the next period. This obviously carries through by a simple induction argument to all time periods. Thus we can write down the complete solution. Before we do this we eliminate some of the new variables we introduced earlier.

$$D_{i-1} = \overline{A'K_iA} - (\overline{A'K_iB})(R + \overline{B'K_iB})^{-1} (\overline{B'K_iA})$$

$$K_i = Q + D_i$$

$$= Q + \overline{A'K_{i+1}A} + (\overline{A'K_{i+1}B})\ L_i$$

$$q_{i-1} = \overline{A'K_ic} + \bar{A}'p_i - (\overline{A'K_iB})(R + \overline{B'K_iB})^{-1}(\overline{B'K_ic} + \bar{B}'p_i - R\tilde{u}_{i-1})$$

$$p_i = -Q\tilde{x}_i + q_i$$

$$= -Q\tilde{x}_i + \overline{A'K_{i+1}c} + \bar{A}'p_{i+1} + (\overline{A'K_{i+1}B})\ m_i$$

$$r_{i-1} = \frac{1}{2}(\overline{B'K_ic} + \bar{B}'p_i - R\tilde{u}_{i-1})'\ m_{i-1} + \frac{1}{2}\overline{c'K_ic} + \bar{c}'p_i + \frac{1}{2}\tilde{u}'_{i-1}R\ \tilde{u}_{i-1} + g_i$$

$$g_i = \frac{1}{2}\tilde{x}'_i\ Q\tilde{x}_i + r_i$$

$$= \frac{1}{2}\tilde{x}'_i\ Q\tilde{x}_i + \frac{1}{2}\tilde{u}'_i\ R\tilde{u}_i + \frac{1}{2}\overline{c'K_{i+1}c} + \bar{c}'p_{i+1} + \frac{1}{2}(\overline{B'K_{i+1}c} + \bar{B}'p_{i+1} - R\tilde{u}_i)'\ m_i + g_{i+1}$$

The complete solution to the optimization problem is therefore :

$$u_t^* = L_tx_t + m_t \tag{A.4}$$

$$L_t = -(R + \overline{B'K_{t+1}B})^{-1} (\overline{B'K_{t+1}A}) \tag{A.5}$$

$$m_t = -(R+\overline{B'K_{t+1}B})^{-1}(\overline{B'K_{t+1}c} + \bar{B}'p_{t+1} - R\tilde{u}_t) \tag{A.6}$$

$$p_t = -Q\tilde{x}_t - \overline{A'K_{t+1}c} + \bar{A}'p_{t+1} + (\overline{A'K_{t+1}B})\ m_t \tag{A.7}$$

$$K_t = Q + \overline{A'K_{t+1}A} + (\overline{A'K_{t+1}B})\ L_t \tag{A.8}$$

$$g_t = \frac{1}{2}\tilde{x}_t' Q\tilde{x}_t + \frac{1}{2}\tilde{u}_t' R\tilde{u}_t + \frac{1}{2}\overline{c' K_{t+1} c} + \bar{c}' p_{t+1}$$

$$+ \frac{1}{2}(\overline{B' K_{t+1} c} + \bar{B}' p_{t+1} - R\tilde{u}_t)' m_t + g_{t+1} \qquad (A.9)$$

$$K_N = Q \qquad (A.10)$$

$$p_N = -Q\tilde{x}_N \qquad (A.11)$$

$$g_N = \frac{1}{2}\tilde{x}_N' Q\tilde{x}_N \qquad (A.12)$$

and $t = 0,1,2, \ldots, N-1$

We can also calculate the value of the optimal cost-to-go and the optimal cost.

The optimal cost-to-go is given by :

$$\alpha_k = \gamma_k^* + \frac{1}{2} E [(x_{k-1} - \tilde{x}_{k-1})' Q (x_{k-1} - \tilde{x}_{k-1})]$$

$$= \frac{1}{2}\int [x_{k-1}' D_{k-1} x_{k-1} + 2 x_{k-1}' q_{k-1} + 2 r_{k-1}] p(x_{k-1}) dx_{k-1}$$

$$+ \frac{1}{2} [[x_{k-1}' Q x_{k-1} - 2 x_{k-1}' Q\tilde{x}_{k-1} + \tilde{x}_{k-1}' Q\tilde{x}_{k-1}] p(x_{k-1}) dx_{k-1}$$

$$= \frac{1}{2}\int [x_{k-1}'(Q + D_{k-1}) x_{k-1} + 2 x_{k-1}'(q_{k-1} - Q\tilde{x}_{k-1}) + 2 r_{k-1} + \tilde{x}_{k-1}' Q\tilde{x}_{k-1}] p(x_{k-1}) dx_{k-1}$$

$$= \frac{1}{2}\int (x_{k-1}' K_{k-1} x_{k-1} + 2 p_{k-1}' x_{k-1} + 2 g_{k-1}) p(x_{k-1}) dx_{k-1}$$

$$= E \{ \frac{1}{2} x_{k-1}' K_{k-1} x_{k-1} + p_{k-1}' x_{k-1} + g_{k-1} \} \quad k = 1,2, \ldots, N \qquad (A.13)$$

The optimal cost is given by $J^* = \alpha_1$

$$J^* = E \left\{ \frac{1}{2} x_o' K_o x_o + p_o' x_o + g_o \right\}$$

Since x_o is known with certainty we can write

$$J^* = \frac{1}{2} x_o' K_o x_o + p_o' x_o + g_o \tag{A.14}$$

APPENDIX B

COMPUTER SUBROUTINES

```
      INTEGER NA,NS,NNA,NPTS,N,M,NM,NN,IPVT(10),KIN,KOUT
      DOUBLE PRECISION A(10,2),B(10,2),C(10,2),Q(10,2),R(10,2),
     +                 SIGA(7,4),SIGB(12,4),SIGBA(12,4),SIGC(10,2),
     +                 SIGAC(7,2),SIGBC(12,2),PT(12),GT,XZERO(10),
     +                 XT(12),UT(12),D(10,2),EKT(10,2),EL(10,2),EM(10),
     +                 U(10,2),V(10,2),W(10,2),W1(10),W2(10),WORK(10),
     +                 VW(10,2),UVW(10,2),ARRAY(51,10),
     +                 BKB(10,2),BKA(10,2),BPB(10,2),
     +                 BPA(10,2),LCOST(20),COST(20),BPC(10),
     +                 BDP(10),DM(10),DP(10),BKC(10),DG
C
      COMMON/INOU/KIN,KOUT
C
      KIN=5
      KOUT=6
      NA=10
      NN=4
      NM=4
      MM=1
      NS=12
      NNA=7
      N=2
      M=2
      CALL MATIO(NA,N,N,Q,4)
      CALL MATIO(NA,M,M,R,4)
      CALL MATIO(NA,N,N,A,4)
      CALL MATIO(NA,N,M,B,4)
      CALL MATIO(N,N,MM,C,4)
      CALL MATIO(NNA,NM,NM,SIGA,4)
      CALL MATIO(NS,NM,NN,SIGB,4)
      CALL MATIO(NA,N,N,SIGC,4)
      UT(1)=0.0D0
      UT(2)=0.0D0
      XZERO(1)=362.0D0
      XZERO(2)=89.0D0
      NPTS=16
      XT(1)=((1.0125D0)**(NPTS-1))*XZERO(1)
      XT(2)=((1.0125D0)**(NPTS-1))*XZERO(2)
      CALL PAR(NA,NS,NNA,NPTS,N,M,NM,NN,A,B,C,Q,R,SIGA,SIGB,
     +         SIGBA,SIGC,SIGAC,SIGBC,XT,UT,PT,GT,XZERO,D,EKT,
     +         EM,EL,ARRAY,COST,LCOST,BKB,BKA,BPA,BPB,EM,DP,
     +         BPC,BDP,BKC,DG,U,V,W,VW,UVW,W1,W2,WORK,IPVT)
      WRITE(KOUT,15) GT
   15 FORMAT(1H0,7H  GT = ,D26.16)
      WRITE(KOUT,16)
   16 FORMAT(1H0,5H  PT )
      CALL MATIO(N,N,MM,PT,3)
      WRITE(KOUT,17)
   17 FORMAT(1H0,7H   M(T) )
      CALL MATIO(N,N,MM,EM,3)
      WRITE(KOUT,18)
   18 FORMAT(1H0,7H  L(T) )
      CALL MATIO(NA,M,N,EL,3)
      WRITE(KOUT,19)
   19 FORMAT(1H0,7H   K(T) )
```

```
      CALL MATIO(NA,N,N,EKT,3)                                          PAR00560
      STOP                                                              PAR00570
C                                                                       PAR00580
C     LAST LINE OF PARMN                                                PAR00590
C                                                                       PAR00600
      END                                                               PAR00610
```

```
      SUBROUTINE PAR(NA,NS,NNA,NPTS,N,M,NM,NN,A,B,C,Q,R,SIGA,SIGB,
     +               SIGBA,SIGC,SIGAC,SIGBC,XT,UT,PT,GT,XZERO,
     +               DK,EKT,EM,EL,ARRAY,COST,LCOST,BKB,BKA,BPA,
     +               BPB,DM,DP,BPC,BDP,BKC,DG,U,V,W,VW,UVW,W1,W2,
     +               WORK,IPVT)
C
C     *****PARAMETERS:
      INTEGER NA,NS,NNA,NPTS,N,M,NM,NN,IPVT(N)
      DOUBLE PRECISION A(NA,N),B(NA,M),C(NA,N),Q(NA,N),R(NA,M),
     +                 SIGA(NNA,NN),SIGB(NS,NM),SIGBA(NS,N),
     +                 SIGC(NA,N),SIGAC(NNA,N),SIGBC(NS,N),
     +                 XT(N),UT(M),PT(N),GT,XZERO(N),DK(NA,N),EKT(NA,N),
     +                 EL(NA,N),ARRAY(NPTS,1),COST(NPTS),LCOST(NPTS),
     +                 EM(M),BKB(NA,M),BKA(NA,N),BPA(NA,N),BPB(NA,M),
     +                 U(NA,N),V(NA,N),W(NA,N),VW(NA,N),UVW(NA,N),W1(N),
     +                 W2(N),DP(N),DM(M),DG,BPC(N),BDP(N),BKC(N),WORK(N)
C
C     *****LOCAL VARIABLES:
      INTEGER K,L,KK,LL,KIN,KOUT,ITOP(40,6),IN(9),NSYM(1),MSC,MAXES,
     +        IXY,JNDEX,II,II5,IEGY,MM,NLG,NGRIDH,INDEX,ICOUNT,ID,IL,
     +        I,IND1,IND2,IND3,IND4,IR,IS,IT,ITM1,IT1,IU,IV,I1,I2,I4,I5,
     +        J,JD,J1,KL
      DOUBLE PRECISION COND,TR,SUM,XMIN,XMAX,YMIN,YMAX,YSF(10),ZERO,
     +                 XM,XS1(2),XS(2),LTS(10,2),XSAVE(16,2),
     +                 LTSAVE(30,2),MTSAVE(30),MTS(2),XS2(2),
     +                 USAVE(15,2),DKSAVE(96,32),PTSAVE(32),GTSAVE(16),
     +                 W3(1,2),DPSAVE(96,16),DMSAVE(96,16),
     +                 DGSAVE(48,16),CSTSEN,LARRAY(15,4),MARRAY(15,2),
     +                 RELSEN(8),SC
C
C     *****FUNCTIONS:
      INTEGER MOD
      DOUBLE PRECISION DFLOAT
C
C     *****SUBROUTINES CALLED:
C     SAVE,MADD,MSUB,MMUL,MQF,MSCALE,TRNATB,TRACE,THPLT,LINEQ,MLINEQ
C
C     ----------------------------------------------------------------
C
C     *****PURPOSE:
C     THIS SUBROUTINE PERFORMS TWO FUNCTIONS:
C     (1)    IT SOLVES THE FOLLOWING DISCRETE TIME LINEAR QUADRATIC
C            OPTIMAL CONTROL PROBLEM FOR A LINEAR SYSTEM WITH PURELY
C            RANDOM PARAMETERS.
C            THE SYSTEM IS DESCRIBED BY
C
C            X(T+1) = A*X(T) + B*U(T) + C, X(0) = XZERO
C
C            A,B, AND C ARE WHITE AND RANDOM.
C
C            THE MEANS, COVARIANCES, AND CROSS COVARIANCES OF
C            A, B, AND C ARE SPECIFIED.
C
C            THE COST CRITERION IS
C
```

```
C          J = (1/2)*E(SUMMATION FROM T = 0 TO N-1 OF
C                                    T
C                  (X(T)-XTILDA(T))  *Q*(X(T)-XTILDA(T)) +
C                                    T
C                  (U(T)-UTILDA(T))  *R*(U(T)-UTILDA(T))) +
C                                    T
C                  (X(N)-XTILDA(N))  *Q*(X(N)-XTILDA(N))
C
C          THE TARGET SEQUENCES (XTILDA(T)), (UTILDA(T)), T=0,1, . .N
C          MUST BE SPECIFIED ALONG WITH Q AND R.
C
C     (2)  IT CALCULATES THE QUANTITIES
C          PARTIAL DERIVATIVE OF JSTAR WITH RESPECT TO SIGMA AND
C          THE PARTIAL DERIVATIVE OF JSTAR WITH RESPECT TO SIGMA *
C          (SIGMA/JSTAR)
C          WHERE JSTAR IS THE OPTIMAL COST (OBTAINED FROM(1)) AND
C          SIGMA IS AN ELEMENT OF ONE OF THE COVARIANCE MATRICES
C          SIGA, SIGB, OR SIGBA. THIS GIVES THE ABSOLUTE AND RELATIVE
C          SENSITIVIES OF THE OPTIMAL PERFORMANCE TO VARIATIONS
C          IN THE PARAMETER VARIANCES.
C
C     *****PARAMETER DESCRIPTION:
C     ON INPUT:
C        NA,NS,NNA          ROW DIMENSIONS OF THE ARRAYS CONTAINING A (AND
C                           B,C,Q,R,SIGC,DK,EKT,EL,BKB,BKA,BPA,BPB,U,V,W,
C                           VW,UVW), SIGB (AND SIGBA,SIGBC), AND SIGA (AND
C                           SIGAC), RESPECTIVELY, AS DECLARED IN THE
C                           CALLING PROGRAM DIMENSION STATEMENT;
C
C        NPTS               NUMBER OF POINTS TO BE PLOTTED;
C
C        N                  NUMBER OF STATES;
C
C        M                  NUMBER OF CONTROLS;
C
C        NM                 = N*M;
C
C        NN                 = N*N;
C
C        A                  N X N SYSTEM MATRIX;
C
C        B                  N X M INPUT MATRIX;
C
C        C                  N X 1 ADDITIVE NOISE VECTOR;
C
C        Q                  N X N STATE WEIGHTING MATRIX;
C
C        R                  M X M CONTROL WEIGHTING MATRIX;
C
C        SIGA               NN X NN COVARIANCE MATRIX OF A;
C
C        SIGB               NM X NM COVARIANCE MATRIX OF B;
C
C        SIGBA              NM X NN CROSS COVARIANCE MATRIX OF A
C                           AND B;
```

```
C
C       SIGC          N X N COVARIANCE MATRIX OF C;
C
C       SIGAC         NN X N CROSS COVARIANCE MATRIX OF A
C                     AND C;
C
C       SIGBC         NM X N CROSS COVARIANCE MATRIX OF B
C                     AND C;
C
C       XT            REAL VECTOR OF LENGTH N CONTAINING
C                     XTILDA(NPTS);
C
C       UT            REAL VECTOR OF LENGTH N CONTAINING
C                     UTILDA(NPTS);
C
C       PT            REAL VECTOR OF LENGTH N CONTAINING
C                     THE VALUES OF P(NPTS);
C
C       GT            REAL SCALAR CONTAINING THE VALUE OF
C                     G(NPTS);
C
C       XZERO         INITIAL CONDITION VECTOR.
C
C     ON OUTPUT:
C
C       EKT           N X N ARRAY CONTAINING THE RICCATI MATRIX;
C
C       EM            M X 1 REAL VECTOR CONTAINING THE CORRECTION
C                     CUM TRACKING TERM;
C
C       EL            M X N GAIN MATRIX;
C
C       ARRAY         NPTS X NN REAL SCRATCH ARRAY USED FOR
C                     PLOTTING;
C
C       COST          NPTS X 1 REAL VECTOR CONTAINING THE OPTIMAL
C                     COST TO GO;
C
C       DK            N X N ARRAY CONTAINING THE PARTIAL DERIVATIVE
C                     OF EKT WITH RESPECT TO SIGMA;
C
C       DM            REAL VECTOR OF LENGTH N CONTAINING THE
C                     PARTIAL DERIVATIVE OF EM WITH RESPECT TO
C                     SIGMA;
C
C       DP            REAL VECTOR OF LENGTH N CONTAINING THE
C                     PARTIAL DERIVATIVE OF PT WITH RESPECT TO
C                     SIGMA;
C
C       DG            REAL SCALAR EQUAL TO THE PARTIAL DERVIATIVE OF
C                     GT WITH RESPECT TO SIGMA;
C
C       BKA,BPA       M X N REAL SCRATCH ARRAYS;
C
C       BKB,BPB       M X M REAL SCRATCH ARRAYS;
```

```
C
C         BDP,BPC,BKC         REAL SCRATCH VECTORS OF LENGTH N;
C
C         U,V,W,VW,UVW        N X N REAL SCRATCH ARRAYS;
C
C         W1,W2,WORK          REAL SCRATCH VECTORS OF LENGTH N;
C
C         IPVT                INTEGER SCRATCH VECTOR OF LENGTH N.
C
C     *****HISTORY:
C     WRITTEN BY J.A.K. CARRIG,
C     (LAB. FOR INF. AND DEC. SYS., (L.I.D.S), M.I.T., RM. 35-427,
C     CAMBRIDGE, MA 02139, PH.: (617) - 253-7263),
C     JANUARY 1979.
C     MOST RECENT VERSION: JANUARY 11, 1979.
C
C     -------------------------------------------------------------------
C
      COMMON/INOU/KIN,KOUT
C
      DATA IBLANK/1H /
C
      DATA IN(1),IN(2),IN(3),IN(4)/1H1,1H2,1H3,1H4/
      DATA IN(5),IN(6),IN(7),IN(8),IN(9)/1H5,1H6,1H7,1H8,1H9/
C
      DATA ITOP(1,1),ITOP(2,1),ITOP(4,1),ITOP(5,1),ITOP(6,1),
     +ITOP(7,1),ITOP(8,1),ITOP(9,1),ITOP(10,1),ITOP(11,1),ITOP(12,1),
     +ITOP(13,1),ITOP(14,1),ITOP(15,1),ITOP(16,1),ITOP(17,1),
     +ITOP(18,1),ITOP(19,1),ITOP(20,1),ITOP(21,1),ITOP(22,1),ITOP(23,1)
     +/1H ,1HK,1H ,1HV,1HE,1HR,1HS,1HU,1HS,1H ,1HT,1HI,1HM,1HE,
     +1H ,1H ,1H ,1H ,1H ,1H ,1H ,1H /
C
      DATA ITOP(1,2),ITOP(2,2),ITOP(3,2),ITOP(4,2),ITOP(5,2),ITOP(6,2),
     +ITOP(7,2),ITOP(8,2),ITOP(10,2),ITOP(11,2),ITOP(12,2),
     +ITOP(13,2),ITOP(14,2),ITOP(15,2),ITOP(16,2),ITOP(17,2),ITOP(18,2),
     +ITOP(19,2),ITOP(20,2),ITOP(21,2),ITOP(22,1),ITOP(23,1)
     +/1H ,1HS,1HT,1HA,1HT,1HE,1H ,1HX,1H ,1HV,1HE,1HR,1HS,1HU,1HS,1H ,
     +1HT,1HI,1HM,1HE,1H ,1H /
C
      DATA ITOP(1,3),ITOP(2,3),ITOP(3,3),ITOP(4,3),ITOP(5,3),ITOP(6,3),
     +ITOP(7,3),ITOP(8,3),ITOP(9,3),ITOP(10,3),ITOP(12,3),ITOP(13,3),
     +ITOP(14,3),ITOP(15,3),ITOP(16,3),ITOP(17,3),ITOP(18,3),ITOP(19,3),
     +ITOP(20,3),ITOP(21,3),ITOP(22,3),ITOP(23,3)
     +/1H ,1HC,1HO,1HN,1HT,1HR,1HO,1HL,1H ,1HU,1H ,1HV,1HE,1HR,1HS,1HU,
     +1HS,1H ,1HT,1HI,1HM,1HE /
C
      DATA ITOP(1,4),ITOP(2,4),ITOP(3,4),ITOP(4,4),ITOP(5,4),ITOP(7,4),
     +ITOP(8,4),ITOP(9,4),ITOP(10,4),ITOP(11,4),ITOP(12,4),ITOP(13,4),
     +ITOP(14,4),ITOP(15,4),ITOP(16,4),ITOP(17,4),ITOP(18,4),ITOP(19,4),
     +ITOP(20,4),ITOP(21,4),ITOP(22,4),ITOP(23,4)
     +/1H ,1HG,1HA,1HI,1HN,1H ,1HV,1HE,1HR,1HS,1HU,1HS,1H ,1HT,
     +1HI,1HM,1HE,1H ,1H ,1H ,1H ,1H /
C
      DATA ITOP(1,5),ITOP(2,5),ITOP(4,5),ITOP(5,5),ITOP(6,5),
     +ITOP(7,5),ITOP(8,5),ITOP(9,5),ITOP(10,5),ITOP(11,5),ITOP(12,5),
```

```
     +ITOP(13,5),ITOP(14,5),ITOP(15,5),ITOP(16,5),ITOP(17,5),ITOP(18,5),
     +ITOP(19,5),ITOP(20,5),ITOP(21,5),ITOP(22,5),ITOP(23,5)
     +/1H ,1HM,1H ,1HV,1HE,1HR,1HS,1HU,1HS,1H ,1HT,1HI,1HM,1HE,1H ,
     + 1H ,1H ,1H ,1H ,1H ,1H ,1H /
C
      DATA ITOP(1,6),ITOP(2,6),ITOP(3,6),ITOP(4,6),ITOP(5,6),ITOP(6,6),
     +ITOP(7,6),ITOP(8,6),ITOP(9,6),ITOP(10,6),ITOP(11,6),ITOP(12,6),
     +ITOP(13,6),ITOP(14,6),ITOP(15,6),ITOP(16,6),ITOP(17,6),ITOP(18,6),
     +ITOP(19,6),ITOP(20,6),ITOP(21,6),ITOP(22,6),ITOP(23,6)
     +/1H ,1HC,1HO,1HS,1HT,1H ,1HV,1HE,1HR,1HS,1HU,1HS,1H ,1HT,1HI,
     +1HM,1HE,1H ,1H ,1H ,1H ,1H ,1H /
C
      MSC=1
      MAXES=0
      IXY=0
      IEGY=1
      ZERO=0.0D0
      XMIN=1.0D0
      NGRIDH=5
      MM=1
      NLG=0
      DO 10 I=1,10
          YSF(I)=1.0D0
   10 CONTINUE
      DO 20 I=24,40
          ITOP(I,1)=IBLANK
          ITOP(I,2)=IBLANK
          ITOP(I,3)=IBLANK
          ITOP(I,4)=IBLANK
          ITOP(I,5)=IBLANK
          ITOP(I,6)=IBLANK
   20 CONTINUE
      IT=NPTS
      XMAX=DFLOAT(IT)
      CALL SAVE(NA,NA,N,N,Q,EKT)
      CALL MMUL(NA,N,N,MM,N,N,Q,XT,PT)
      CALL MSCALE(N,N,MM,-1.0D0,PT)
      CALL MQF(NA,N,N,N,MM,Q,XT,W1,WORK)
      GT=W1(1)/2.0D0
      DO 30 L=1,N
          DO 30 K=1,N
              INDEX=K+(L-1)*N
              ARRAY(IT,INDEX)=Q(K,L)
   30 CONTINUE
      INDEX=IT*N
      PTSAVE(INDEX-1)=PT(1)
      PTSAVE(INDEX)=PT(2)
      GTSAVE(IT)=GT
      ITM1=IT-1
      DO 220 IL=1,ITM1
          IT1=IT-IL
          CALL TRNATE(NA,NA,N,M,B,U)
          CALL MMUL(NA,N,N,MM,M,N,U,PT,W1)
          CALL MMUL(NA,M,NA,MM,M,M,R,UT,V)
          CALL MSUB(N,NA,N,M,MM,W1,V,W1)
```

```
      CALL MMUL(NA,NA,NA,N,M,N,U,EKT,W)
      CALL MMUL(NA,NA,NA,M,M,N,W,B,BKB)
      CALL MMUL(NA,NA,NA,N,M,N,W,A,BKA)
      CALL MMUL(NA,NA,N,MM,M,N,W,C,BKC)
C
C     CALCULATE M(T),L(T)
C
      DO 60 K=1,M
         KK=1+(K-1)*N
         CALL MMUL(NA,NS,NA,N,N,N,EKT,SIGBC(KK,1),W)
         CALL TRACE(NA,N,W,TR)
         BKC(K)=BKC(K)+W1(K)+TR
         DO 40 L=1,M
            LL=1+(L-1)*N
            CALL MMUL(NA,NS,NA,N,N,N,EKT,SIGB(KK,LL),W)
            CALL TRACE(NA,N,W,TR)
            BKB(K,L)=-R(K,L)-BKB(K,L)-TR
   40    CONTINUE
         DO 50 L=1,N
            LL=1+(L-1)*N
            CALL MMUL(NA,NS,NA,N,N,N,EKT,SIGBA(KK,LL),W)
            CALL TRACE(NA,N,W,TR)
            BKA(K,L)=BKA(K,L)+TR
   50    CONTINUE
   60 CONTINUE
      CALL SAVE(NA,NA,M,M,BKB,W)
      CALL SAVE(N,M,M,MM,BKC,EM)
      CALL LINEQ(NA,M,W,EM,COND,IPVT,WORK)
      CALL SAVE(NA,NA,M,M,BKB,W)
      CALL SAVE(NA,NA,M,N,BKA,EL)
      CALL MLINEQ(NA,NA,M,N,W,EL,COND,IPVT,WORK)
C
C     SAVE LT AND MT
C
      DO 70 J=1,M
         I1=2*(ITM1-IL)+J
         LTSAVE(I1,1)=EL(J,1)
         LTSAVE(I1,2)=EL(J,2)
   70 CONTINUE
C
      DO 80 L=1,N
         DO 80 K=1,M
            INDEX=K+(L-1)*M
            LARRAY(IT1,INDEX)=EL(K,L)
   80 CONTINUE
C
      I2=2*(ITM1-IL)
      MTSAVE(I2+1)=EM(1)
      MTSAVE(I2+2)=EM(2)
C
      DO 90 K=1,M
         MARRAY(IT1,K)=EM(K)
   90 CONTINUE
C
C     CALCULATE DK, DM, DG, DP, COST SENSITIVITY
```

```
C
      DO 190 ICOUNT=1,3
         IND1=-1
         IND2=N*IL+1+(ICOUNT-1)*12
         IND3=0
         IND4=IL+1+(ICOUNT-1)*6
         DO 180 I=1,NN
            J=I
            IND1=IND1+N
            IND3=IND3+1
            DO 100 I1=1,2
               INDEX=IND2-N+I1-1
               DP(I1)=DPSAVE(INDEX,IND3)
               DM(I1)=DMSAVE(INDEX,IND3)
               DO 100 J1=1,2
                  JNDEX=IND1+J1-1
                  DK(I1,J1)=DKSAVE(INDEX,JNDEX)
 100        CONTINUE
            DG=DGSAVE(IND4-1,IND3)
            CALL TRNATB(NA,NA,N,M,B,U)
            CALL MMUL(NA,NA,NA,N,M,N,U,DK,W)
            CALL MMUL(NA,NA,NA,N,M,N,W,A,BPA)
            CALL MMUL(NA,NA,NA,M,M,N,W,B,BPB)
            CALL MMUL(NA,NA,NA,MM,M,M,W,C,BPC)
            CALL MMUL(NA,N,N,MM,M,N,U,DP,BDP)
            IR=1+(I-1)/N
            IS=1+(J-1)/N
            IU=1+MOD(I-1,N)
            IV=1+MOD(J-1,N)
C
C           CALCULATE DK
C
            DO 110 K=1,M
               DO 110 L=1,N
                  KK=1+(K-1)*N
                  LL=1+(L-1)*N
                  CALL MMUL(NA,NA,NA,N,N,N,DK,SIGBA(KK,LL),W)
                  CALL TRACE(NA,N,W,TR)
                  BPA(K,L)=BPA(K,L)+TR
  110       CONTINUE
            DO 120 K=1,M
               DO 120 L=1,M
                  KK=1+(K-1)*N
                  LL=1+(L-1)*N
                  CALL MMUL(NA,NA,NA,N,N,N,DK,SIGB(KK,LL),W)
                  CALL TRACE(NA,N,W,TR)
                  BPB(K,L)=BPB(K,L)+TR
  120       CONTINUE
            IF(ICOUNT.EQ.2) BPA(IR,IS)=BPA(IR,IS)+EKT(IV,IU)
            CALL TRNATB(NA,NA,M,N,BPA,W)
            CALL MMUL(NA,NA,NA,N,N,M,W,EL,VW)
C
            CALL SAVE(NA,NA,M,N,BPA,UVW)
            CALL SAVE(NA,NA,M,M,BKB,W)
            CALL MLINEQ(NA,NA,M,N,W,UVW,COND,IPVT,WORK)
```

```
      CALL TRNATB(NA,NA,M,N,BKA,W)
      CALL MMUL(NA,NA,NA,N,M,M,W,UVW,V)
      CALL MADD(NA,NA,NA,M,N,V,VW,VW)
C
      CALL SAVE(NA,NA,M,N,BKA,W)
      CALL SAVE(NA,NA,M,M,BKB,UVW)
      CALL MLINEQ(NA,NA,M,N,UVW,W,COND,IPVT,WORK)
      IF(ICOUNT.EQ.3) BPB(IR,IS)=BPB(IR,IS)+EKT(IV,IU)
      CALL MMUL(NA,NA,NA,N,M,M,BPB,W,UVW)
      CALL SAVE(NA,NA,M,M,BKB,W)
      CALL MLINEQ(NA,NA,M,N,W,UVW,COND,IPVT,WORK)
      CALL TRNATB(NA,NA,M,N,BKA,W)
      CALL MMUL(NA,NA,NA,N,M,M,W,UVW,V)
      CALL MADD(NA,NA,NA,M,N,VW,V,VW)
C
      CALL TRNATB(NA,NA,N,N,A,V)
      CALL MMUL(NA,NA,NA,N,N,N,DK,A,UVW)
      CALL MMUL(NA,NA,NA,N,N,N,V,UVW,W)
      CALL MADD(NA,NA,NA,N,N,W,VW,UVW)
      DO 130 K=1,N
         DO 130 L=1,N
            KK=1+(K-1)*N
            LL=1+(L-1)*N
            CALL MMUL(NA,NA,NA,N,N,N,DK,SIGA(KK,LL),W)
            CALL TRACE(NA,N,W,TR)
            UVW(K,L)=UVW(K,L)+TR
  130 CONTINUE
      IF(ICOUNT.EQ.1) UVW(IR,IS)=UVW(IR,IS)+EKT(IV,IU)
C
C     CALCULATE DM
C
      DO 140 K=1,M
         KK=1+(K-1)*N
         CALL MMUL(NA,NS,NA,N,N,N,DK,SIGBC(KK,1),W)
         CALL TRACE(NA,N,W,TR)
         BPC(K)=BPC(K)+BDP(K)+TR
  140 CONTINUE
C
      CALL SAVE(NA,NA,N,N,BKB,W)
      CALL SAVE(N,M,N,M,BPC,DM)
      CALL LINEQ(NA,M,W,DM,COND,IPVT,WORK)
C
      CALL SAVE(N,N,N,MM,BKC,W1)
      CALL SAVE(NA,NA,N,N,BKB,W)
      CALL LINEQ(NA,M,W,W1,COND,IPVT,WORK)
      CALL MMUL(NA,N,N,MM,N,N,BPB,W1,W2)
      CALL SAVE(NA,NA,N,N,BKB,W)
      CALL LINEQ(NA,N,W,W2,COND,IPVT,WORK)
      CALL MADD(N,N,N,N,MM,DM,W2,DM)
C
C     CALCULATE DG
C
      CALL MMUL(NA,NA,NA,N,N,N,DK,SIGC,W)
      CALL TRACE(NA,N,W,TR)
      CALL TRNATB(NA,MM,N,MM,C,W3)
```

```
      CALL MMUL(NA,NA,N,MM,N,N,DK,C,W1)
      CALL MMUL(MM,N,NA,MM,MM,N,W3,W1,W)
      CALL MMUL(MM,N,NA,MM,MM,N,W3,DP,V)
C
      DG=DG+(TR+W(1,1))/2.0D0 + V(1,1)
C
      CALL TRNATB(N,MM,N,MM,BPC,W3)
      CALL MMUL(MM,M,NA,MM,MM,N,W3,EM,W)
C
      CALL TRNATB(N,MM,N,MM,BKC,W3)
      CALL MMUL(MM,N,N,MM,MM,N,W3,DM,V)
C
      DG= DG+(W(1,1)+V(1,1))/2.0D0
C
C     CALCULATE DP
C
      CALL TRNATB(NA,NA,N,N,A,W)
      CALL MMUL(NA,NA,N,MM,N,N,DK,C,W1)
      CALL MADD(N,N,N,N,MM,W1,DP,W1)
      CALL MMUL(NA,N,N,MM,N,N,W,W1,W2)
C
      DO 150 K=1,N
         KK=1+(K-1)*N
         CALL MMUL(NA,NNA,NA,N,N,N,DK,SIGAC(KK,1),W)
         CALL TRACE(NA,N,W,TR)
         DP(K)=W2(K)+TR
  150 CONTINUE
C
      CALL TRNATB(NA,NA,N,N,BPA,W)
      CALL MMUL(NA,M,N,MM,N,N,W,EM,W1)
      CALL MADD(N,N,N,N,MM,DP,W1,DP)
      CALL TRNATB(NA,NA,N,N,BKA,W)
      CALL MMUL(NA,N,N,MM,N,N,W,DM,W1)
C
      CALL MADD(N,N,N,N,MM,DP,W1,DP)
C
      IF(IL.NE.ITM1) GO TO 160
C
C     CALCULATE COST SENSITIVITY
C
      CALL TRNATB(N,1,N,1,XZERO,W3)
      CALL MMUL(NA,N,N,MM,N,N,UVW,XZERO,WORK)
      CALL MMUL(1,N,NA,MM,MM,N,W3,WORK,W)
C
      CALL TRNATB(N,N,N,MM,DP,W3)
      CALL MMUL(N,N,N,MM,MM,N,W3,XZERO,V)
C
      CSTSEN=W(1,1)/2.0D0 + V(1,1) + DG
      WRITE(KOUT,900) CSTSEN
      IF(ICOUNT.EQ.1.AND.I.EQ.J) KL=KL+1
      IF(ICOUNT.EQ.3.AND.I.EQ.J) KL=KL+1
      IF(ICOUNT.EQ.1.AND.I.EQ.J) RELSEN(KL)=CSTSEN*SIGA(I,J)
      IF(ICOUNT.EQ.3.AND.I.EQ.J) RELSEN(KL)=CSTSEN*SIGB(I,J)
C
  160 CONTINUE
```

```
C
C             SAVE DK, DP, DM, DG
C
              DO 170 ID=1,2
                 INDEX=IND2+ID-1
                 DPSAVE(INDEX,IND3)=DP(ID)
                 DMSAVE(INDEX,IND3)=DM(ID)
                 DO 170 JD=1,2
                    JNDEX=IND1+JD-1
                    DKSAVE(INDEX,JNDEX)=UVW(ID,JD)
  170         CONTINUE
              DGSAVE(IND4,IND3)=DG
C
  180       CONTINUE
  190    CONTINUE
C
C        CALCULATE G(T), OVERWRITING G(T+1)
C
         SC=.98765432098765440D0
         CALL MSCALE(N,N,MM,SC,XT)
C
         CALL TRNATB(N,NA,M,MM,BKC,V)
         CALL MMUL(NA,M,N,MM,MM,M,V,EM,W2)
         CALL MMUL(NA,NA,NA,N,N,N,EKT,SIGC,W)
         CALL TRACE(NA,N,W,TR)
         CALL TRNATB(NA,NA,N,MM,C,W)
         CALL MMUL(NA,NA,NA,N,MM,N,W,EKT,V)
         CALL MMUL(NA,NA,N,MM,MM,N,V,C,W1)
         CALL MMUL(NA,N,NA,MM,MM,N,W,PT,V)
         GT=GT+V(1,1)+(W1(1)+W2(1)+TR)/2.0D0
         CALL MQF(NA,N,N,N,MM,Q,XT,W1,WORK)
         CALL MQF(NA,M,N,M,MM,R,UT,W2,WORK)
         GT=GT+(W1(1)+W2(1))/2.0D0
C
C        SAVE GT
C
         GTSAVE(IT1)=GT
C
C        CALCULATE P(T), OVERWRITING P(T+1)
C
         CALL TRNATB(NA,NA,N,N,A,W)
         CALL MMUL(NA,NA,NA,N,N,N,EKT,C,V)
         CALL MADD(NA,N,NA,N,MM,V,PT,V)
         CALL MMUL(NA,NA,N,MM,N,N,W,V,W1)
         CALL TRNATB(NA,NA,M,N,BKA,W)
         CALL MMUL(NA,M,NA,MM,N,M,W,EM,V)
         CALL MMUL(NA,N,N,MM,N,N,Q,XT,W2)
         CALL MSUB(NA,N,N,N,MM,V,W2,W2)
         DO 200 K=1,N
            KK=1+(K-1)*N
            CALL MMUL(NA,NNA,NA,N,N,N,EKT,SIGAC(KK,1),W)
            CALL TRACE(NA,N,W,TR)
            PT(K)=W1(K)+W2(K)+TR
  200    CONTINUE
C
```

```
C         SAVE PT
C
          I4=(IT-IL)*N
          PTSAVE(I4-1)=PT(1)
          PTSAVE(I4)=PT(2)
C
C         CALCULATE K(T), OVERWRITING K(T+1)
C
          CALL TRNATB(NA,NA,M,N,BKA,W)
          CALL MMUL(NA,NA,NA,N,N,M,W,EL,U)
          CALL MQF(NA,NA,NA,N,N,EKT,A,W,V)
          CALL MADD(NA,NA,NA,N,N,U,W,U)
          DO 210 L=1,N
             DO 210 K=1,N
                KK=1+(K-1)*N
                LL=1+(L-1)*N
                CALL MMUL(NA,NNA,NA,N,N,N,EKT,SIGA(KK,LL),V)
                CALL TRACE(NA,N,V,TR)
                W(K,L)=Q(K,L)+U(K,L)+TR
                INDEX=K+(L-1)*N
                ARRAY(IT1,INDEX)=W(K,L)
  210        CONTINUE
          CALL SAVE(NA,NA,N,N,W,EKT)
  220 CONTINUE
C
C     PLOT K
C
      DO 230 I=1,N
         DO 230 J=1,N
            INDEX=J+(I-1)*N
            IF(INDEX.LE.9) ITOP(3,1)=IN(INDEX)
            IF(INDEX.GT.9) ITOP(3,1)=IBLANK
            NSYM(1)=11
            CALL THPLT(NPTS,IESY,ARRAY(1,INDEX),NPTS,ITOP,NSYM,XMIN,
     +                 XMAX,YMIN,YMAX,YSF,NGRIDH,NLG,MSCALE,MAXES,IXY)
  230 CONTINUE
C
C     CALCULATE STATE XS
C
      XSAVE(1,1)=XZERO(1)
      XSAVE(1,2)=XZERO(2)
C
      XS(1)=XZERO(1)
      XS(2)=XZERO(2)
      DO 250 I=1,ITM1
         DO 240 J=1,M
            INDEX=2*I-2+J
            LTS(J,1)=LTSAVE(INDEX,1)
            LTS(J,2)=LTSAVE(INDEX,2)
  240    CONTINUE
         CALL MMUL(NA,N,M,MM,M,N,LTS,XS,XS1)
         II=I+I
         MTS(1)=MTSAVE(II-1)
         MTS(2)=MTSAVE(II)
         CALL MADD(M,M,M,M,MM,XS1,MTS,XS1)
```

```
         USAVE(I,1)=XS1(1)
         USAVE(I,2)=XS1(2)
         CALL MMUL(NA,M,M,MM,N,M,B,XS1,XS2)
         CALL MADD(N,N,N,N,MM,XS2,C,XS2)
         CALL MMUL(NA,N,N,MM,N,N,A,XS,XS1)
         CALL MADD(N,N,N,N,MM,XS1,XS2,XS)
         XSAVE(I+1,1)=XS(1)
         XSAVE(I+1,2)=XS(2)
  250 CONTINUE
C
C     PLOT STATE TRAJECTORY
C
      DO 260 J=1,N
         NSYM(1)=24
         IF(J.LE.9) ITOP(9,2)=IN(J)
         IF(J.GT.9) ITOP(9,2)=IBLANK
         CALL THPLT(NPTS,IEGY,XSAVE(1,J),NPTS,ITOP(1,2),NSYM,XMIN,XMAX,
     +              YMIN,YMAX,YSF,NGRIDH,NLG,MSCALE,MAXES,IXY)
  260 CONTINUE
C
C     PLOT CONTROL TRAJECTORY
C
      XM=DFLOAT(ITM1)
      DO 270 J=1,N
         NSYM(1)=21
         IF(J.LE.9) ITOP(11,3)=IN(J)
         IF(J.GT.9) ITOP(11,3)=IBLANK
         CALL THPLT(ITM1,IEGY,USAVE(1,J),ITM1,ITOP(1,3),NSYM,XMIN,XM,
     +              YMIN,YMAX,YSF,NGRIDH,NLG,MSCALE,MAXES,IXY)
  270 CONTINUE
C
C     PLOT GAINS
C
      DO 280 I=1,N
         DO 280 J=1,M
            NSYM(1)=12
            INDEX=J+(I-1)*M
            IF(INDEX.LE.9) ITOP(6,4)=IN(INDEX)
            IF(INDEX.GT.9) ITOP(6,4)=IBLANK
            CALL THPLT(ITM1,IEGY,LARRAY(1,INDEX),ITM1,ITOP(1,4),NSYM,
     +                 XMIN,XM,YMIN,YMAX,YSF,NGRIDH,NLG,MSCALE,MAXES,
     +                 IXY)
  280 CONTINUE
C
C     PLOT CORRECTION TERM M(T)
C
      DO 290 J=1,M
         NSYM(1)=13
         IF(J.LE.9) ITOP(3,5)=IN(J)
         IF(J.GT.9) ITOP(3,5)=IBLANK
         CALL THPLT(ITM1,IEGY,MARRAY(1,J),ITM1,ITOP(1,5),NSYM,XMIN,XM,
     +              YMIN,YMAX,YSF,NGRIDH,NLG,MSCALE,MAXES,IXY)
  290 CONTINUE
C
C     CALCULATE COST
```